February 2015, Vol. 53, No. 2

IEEE Communications MAGAZINE

www.comsoc.org

- *Spatially and Spectrally Flexible Elastic Optical Networking*
- *Optical Multi-Service Core and Access Networks*
- *Network and Service Virtualization*

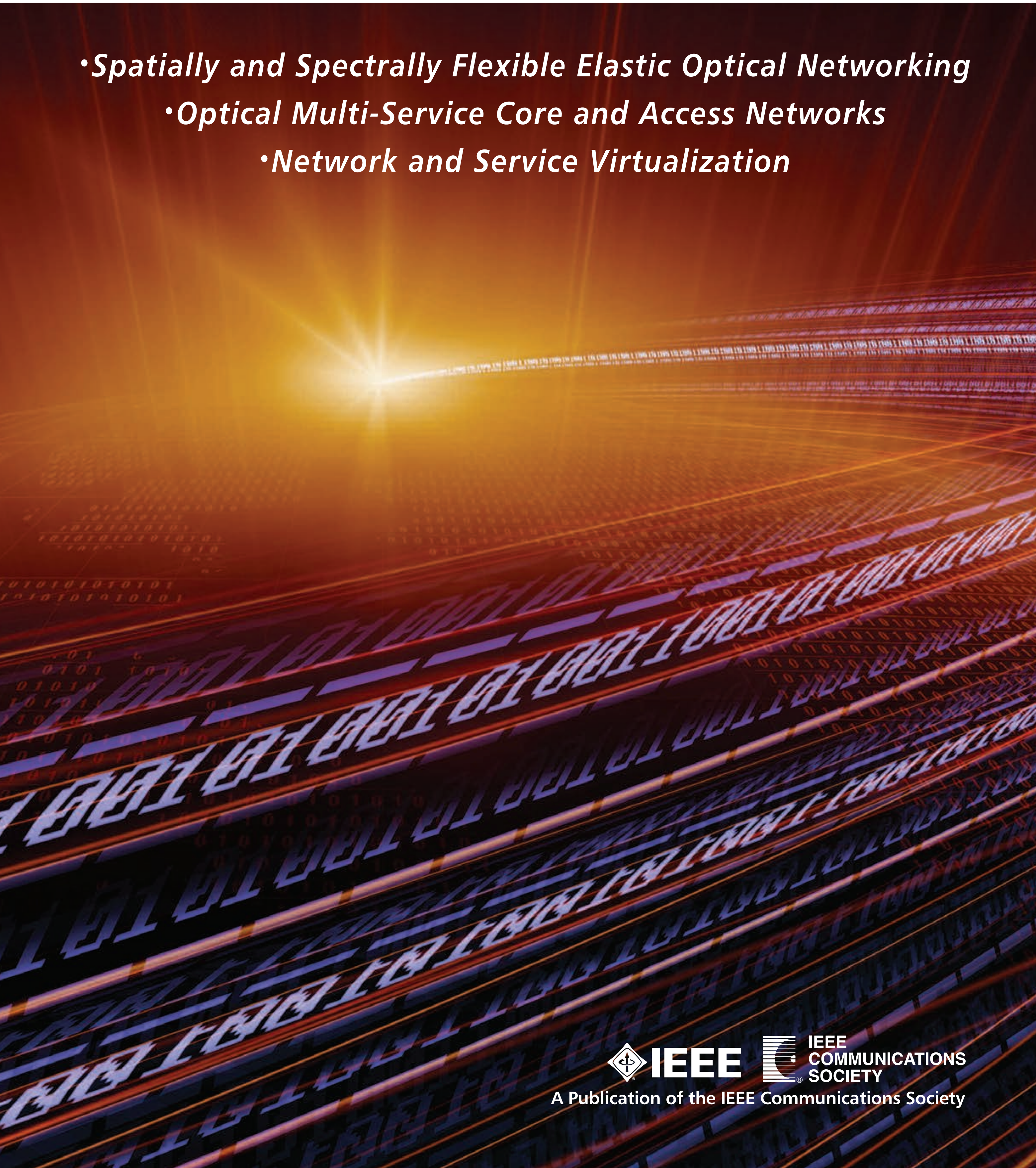
IEEE | IEEE COMMUNICATIONS SOCIETY

A Publication of the IEEE Communications Society

February 2015, Vol. 53, No. 2

IEEE Communications Magazine

www.comsoc.org



•Spatially and Spectrally Flexible Elastic Optical Networking

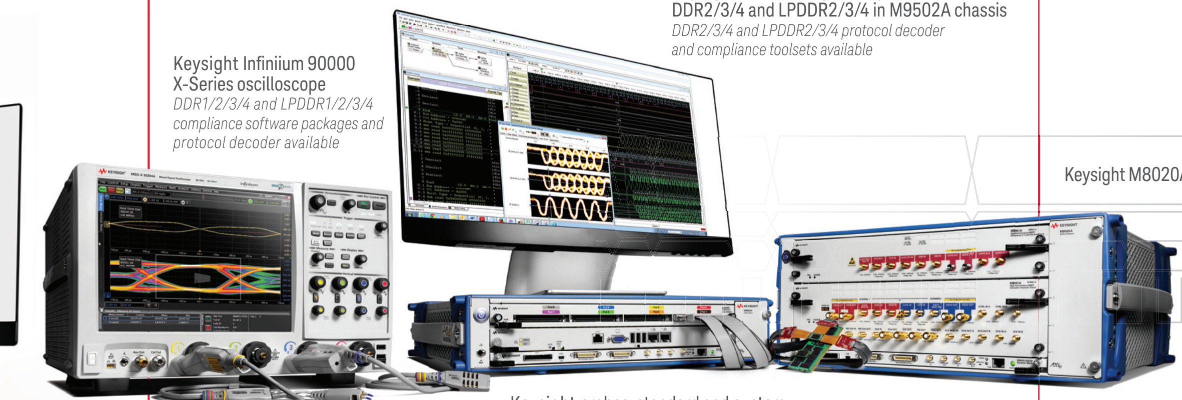
•Optical Multi-Service Core and Access Networks

•Network and Service Virtualization

IEEE

IEEE COMMUNICATIONS SOCIETY

A Publication of the IEEE Communications Society



Keysight Infiniium 90000 X-Series oscilloscope
DDR1/2/3/4 and LPDDR1/2/3/4 compliance software packages and protocol decoder available

Keysight U4154B logic analyzer module for DDR2/3/4 and LPDDR2/3/4 in M9502A chassis
DDR2/3/4 and LPDDR2/3/4 protocol decoder and compliance toolsets available

Keysight M8020A high-performance J-BERT

Keysight probes-standard and custom
Standard and custom DDR and LPDDR oscilloscope and logic analyzer BGA interposer solutions

HARDWARE + SOFTWARE

If you're an engineer on the leading edge of DDR memory design, chances are, you feel challenged to go faster, using less power and a smaller footprint. We can help. Keysight is the only test and measurement company that offers hardware and software solutions across all stages of DDR chip development. From simulation to debug, from validation to compliance, we've got you covered.

- *More than 4,000 electronic measurement tools*
- *Benchtop, modular and software solutions from simulation to compliance*
- *Full line of high-speed, high-density probes*

PEOPLE

Keysight engineers are leading the industry in the design of the next generation of DDR memory standards and solutions. This means that in the future, we can help you solve cutting-edge design challenges by sharing our expertise. It also means that we can rapidly integrate new DDR specs into our hardware and software. So they'll be fully functional the day you need them.

- *JEDEC and UFSA Board of Directors Chairman, JEDEC Digital Logic (JC40.5) and UFSA Compliance Committees*
- *Hundreds of applications engineers in 100 countries around the world*
- *Thousands of patents issued in Keysight's history*

KEYSIGHT TECHNOLOGIES

Unlocking Measurement Insights

IEEE Communications MAGAZINE

FEBRUARY 2015, Vol. 53, No. 2

www.comsoc.org/commag

Spatially and Spectrally Flexible Elastic Optical Networking

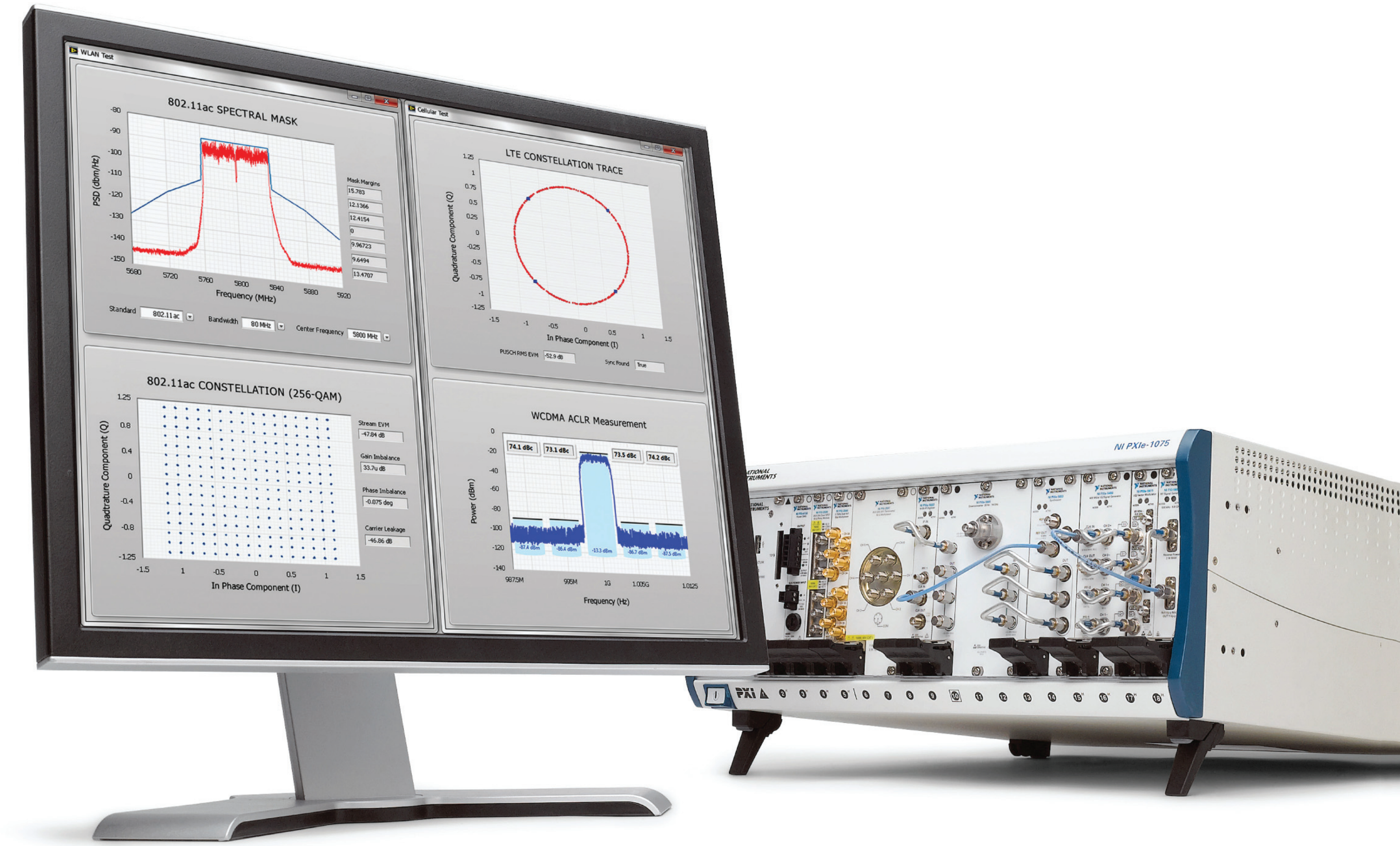
Guest Editors: Ioannis Tomkos, Yutaka Miyamoto, Glenn Wellbrock, and Peter J. Winzer

Network and Service Virtualization

Guest Editors: Kostas Pentikousis, Catalin Meirosu, Diego R. Lopez, Spyros Denazis, Kohei Shiomoto, and Fritz-Joachim Westphal

Redefining RF and Microwave Instrumentation

with open software and modular hardware



Achieve speed, accuracy, and flexibility in your RF and microwave test applications by combining National Instruments open software and modular hardware. Unlike rigid traditional instruments that quickly become obsolete by advancing technology, the system design software of NI LabVIEW coupled with NI PXI hardware puts the latest advances in PC buses, processors, and FPGAs at your fingertips.

WIRELESS TECHNOLOGIES

National Instruments supports a broad range of wireless standards including:

802.11a/b/g/n/ac	LTE
CDMA2000/EV-DO	GSM/EDGE
WCDMA/HSPA/HSPA+	Bluetooth

>> Learn more at ni.com/redefine

800 813 5078

NATIONAL INSTRUMENTS™

IEEE COMMUNICATIONS MAGAZINE (ISSN 0163-6804) is published monthly by The Institute of Electrical and Electronics Engineers, Inc. Headquarters address: IEEE, 3 Park Avenue, 17th Floor, New York, NY 10016-5997, USA; tel: +1 (212) 705-8900; http://www.comsoc.org/commag. Responsibility for the contents rests upon authors of signed articles and not the IEEE or its members. Unless otherwise specified, the IEEE neither endorses nor sanctions any positions or actions espoused in *IEEE Communications Magazine.*

ANNUAL SUBSCRIPTION: $27 per year print subscription. $16 per year digital subscription. Non-member print subscription: $400. Single copy price is $25.

EDITORIAL CORRESPONDENCE: Address to: Editor-in-Chief, Osman S. Gebizlioglu, Huawei Technologies, 400 Crossing Blvd., 2nd Floor, Bridgewater, NJ 08807, USA; tel: +1 (908) 541-3591, e-mail: Osman.Gebizlioglu@huawei.com.

POSTMASTER: Send address changes to *IEEE Communications Magazine*, IEEE, 445 Hoes Lane, Piscataway, NJ 08855-1331. GST Registration No. 125634188. Printed in USA. Periodicals postage paid at New York, NY and at additional mailing offices. Canadian Post International Publications Mail (Canadian Distribution) Sales Agreement No. 40030962. Return undeliverable Canadian addresses to: Frontier, PO Box 1051, 1031 Helena Street, Fort Eire, ON L2A 6C7.

SUBSCRIPTIONS: Orders, address changes — IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08855-1331, USA; tel: +1 (732) 981-0060; e-mail: address.change@ieee.org.

SUBMISSIONS: The magazine welcomes tutorial or survey articles that span the breadth of communications. Submissions will normally be approximately 4500 words, with few mathematical formulas, accompanied by up to six figures and/or tables, with up to 10 carefully selected references. Electronic submissions are preferred, and should be sumitted through Manuscript Central: http://mc.manuscriptcentral.com/commag-ieee. Submission instructions can be found at the following: http://www.comsoc.org/commag/paper-submission-guidelines. For further information contact Zoran Zvonar, Associate Editor-in-Chief (zoran.zvonar@mediatek.com). All submissions will be peer reviewed.

Optical Multi-Service Core and Access Networks

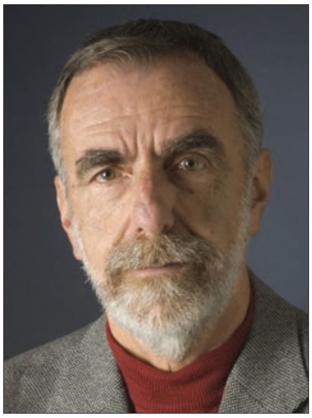
Series Editors: Osman S. Gebizlioglu and Vijay Jain

Accepted from Open Call

IOT INITIATIVE POISED FOR GROWTH IN COLLABORATION WITH COMSOC

The February President's page is devoted to the IEEE Internet of Things initiative. Because of its scope, and the many ComSoc volunteers involved in it, this initiative is a paradigmatic example of how ComSoc can (and should) contribute to multidisciplinary endeavours sponsored by IEEE.

I am happy to introduce the Chair of the IEEE Internet of Things initiative, Roberto Minerva, who will describe the initiative's past achievements and future goals.



SERGIO BENEDETTO

Roberto Minerva (ComSoc member from 1992) is the Chairman of the IEEE IoT Initiative. He is now in the Standardization Group of Telecom Italia. He has recently been the head of the Innovative Architectures group within the Future Centre in the Strategy Department of Telecom Italia. His job is to create advanced scenarios derived from the application of emerging ICT technologies with innovative business models, especially in the area of IoT, distributed computing, programmable networks, and personal data. Roberto is also a contract Professor at Turin's Polytechnics teaching a course on mobile services, and the author of several papers published in international conferences, books, and magazines. He has a Masters Degree in computer science from Bari University, Italy, and a Ph.D. in computer science and telecommunications from Telecom Sud Paris, France.

Since its launch in 2014, the IEEE Internet of Things (IoT) Initiative has passed several significant milestones, including the launch of its first international conference, creating a Web portal for disseminating information and connecting experts, launching a Newsletter in record time, and achieving greater visibility around the world, in part due to the efforts of the IEEE Communications Society (ComSoc). Moving forward, ComSoc will play a critical role in helping the IEEE IoT Initiative meet its future goals.

IEEE is better positioned than any other organization to foster cross-disciplinary collaboration on key aspects of the Internet of Things (IoT), and to help bring the various relevant technologies to maturity that will be the foundation of an ecosystem of products, applications, and services.

This past year has been a significant one for IEEE with the launch of the IEEE IoT Initiative, which has achieved many significant milestones such as new conferences, online resources, and publications, while raising IEEE's visibility as a source of expertise for IoT for media outlets, in part due to the participation of ComSoc. The first IEEE World Forum on IoT (WF-IoT) was held in Seoul, Korea, and it marked a significant moment in which eight IEEE Societies worked together in order to develop and deliver an integrated view of IoT. This effort has been synergistic with the further development of the *IEEE IoT Journal* that is sponsored and supported by ComSoc. In addition, the IoT Initiative, in strong cooperation with ComSoc, has supported an IoT Rapid Standardization Program whose effort has contributed to the consolidation of a new standards project, IEEE P2413™, "IEEE Standard for an Architectural Framework for the Internet of Things," which is a significant undertaking that addresses the growing intersection of smart technologies and high-speed communications that are starting to create profound, positive changes in nearly every aspect of our daily lives. [1]

Given the nature of the outreach and education planned for the IEEE IoT Initiative in 2015 and beyond, ComSoc will play a vital role as we work on a wide variety of activities that include information dissemination and events that will raise the visibility of IEEE and awareness of the IoT Initiative.

THE IEEE IOT INITIATIVE

In 2014 the IoT Initiative worked hard to establish IEEE as a thought leader, catalyst, and essential go-to resource for the global IoT community. The Initiative has connected technical and business communities to IEEE experts and resources, including the engineering and technology professionals in industry, academia, and government, as well as the general public, including consumers, and governmental bodies seeking non-biased and balanced information to help them better understand IoT developments.

SIGNIFICANT MILESTONES ACHIEVED IN 2014

One of the major objectives for 2014 was to ignite a lively IoT Technical Community. IoT is a particular example of what could happen in the future to Technical Communities. The bulk of IoT experts are still in academia and industry, but no longer exclusively there. Many experts are part of the Makers Communities (part of the wide Do-It-Yourself movement) and they are tackling IoT with a bottom up approach supported by the large availability of sensors and microcontrollers at a low price. This kind of expert is one additional target for the IoT Initiative. So far the IEEE IoT Technical Community includes more than 6,400 members. In order to maintain such a large community, the IEEE IoT Initiative has to provide updated content for different types of experts.

The IEEE IoT Initiative has set out to develop and promote valuable programs, products, and services for the IoT community on a number of fronts: the IEEE WF-IoT as the principal conference devoted to IoT; the *IEEE IoT*

Journal as the principal journal devoted to IoT; an online IoT newsletter with complimentary subscription as a benefit of joining the IEEE IoT Technical Community; IEEE IoT standards as the principal standards; and, tutorials, review articles, workshops, and courses as the principal IoT educational activities.

A key component of the IEEE IoT Initiative is the IoT Web Portal, which since its inception has had more than 15,000 visitors with an average of 1,670 visitors per month. What is particularly gratifying is that more than 70 percent of visitors return, and more than 63 percent of visitors are from outside the United States, indicating that the Portal is reaching a global audience interested in learning more about IoT standards and industry developments.

Part of the IEEE IoT Web Portal's success can most certainly be attributed to the fact that it has been refreshed more than 30 times with new content. Repeat visitors have become accustomed to seeing new information and resources posted on a consistent basis, including articles by industry representatives and IEEE IoT experts, as well new videos. ComSoc and its members have greatly contributed to that content. In 2015, the IEEE IoT Technical Community will continue to grow, and this has to be supported by the production of relevant content. In this respect ComSoc can play a relevant role.

INAUGURAL CONFERENCE WELL-ATTENDED AND SPONSORED

The launch of the IoT initiative's flagship conference, the IEEE World Forum on IoT, has provided a wealth of source material for video content on the Portal. The first conference was held March 6–8, 2014 in Seoul, Korea, attracting 237 attendees with Roberto De Marca, 2014 IEEE President, as one of the many speakers. Also featured were keynote speakers Dr. Kyungwhoon Cheun, senior vice president at the Samsung Electronics DMC R&D Center; Dr. Vida Ilderem, vice president of Intel Labs and director of Integrated Platform Research for Intel Corporation; and Chung-Sheng Li, director of commercial systems at IBM T.J. Watson Research Center. Of the overall attendees, 58 percent were IEEE members, 27 percent were non-members, and 15 percent were students, with representation from more than 60 global organizations. That we were able to attract such a broad mix of attendees for the inaugural conference is a great indicator of its initial and future success. The IEEE WF-IoT was also a financial success with broad sponsorship and leadership provided by ComSoc.

DISSEMINATING IEEE IOT INSIGHT AND EXPERTISE

Publications and newsletters have always been a critical vehicle for the IEEE's various initiatives, and as part of the IoT Initiative we launched a newsletter in September while gaining visibility in existing publications. The bimonthly *IEEE IoT Newsletter* was developed in record time, managing to produce two issues before the end of the year with four articles per issue. We participated in The Institute Special Report: The Internet of Things, March 2014, which attracted 45,000 online visits, with the IoT Tech News video from the issue featured on IEEE.tv receiving more than 6,000 views. Meanwhile, *IEEE IoT Journal* already boasts four issues with 33 papers and nearly 10,000 downloads within the first six months.

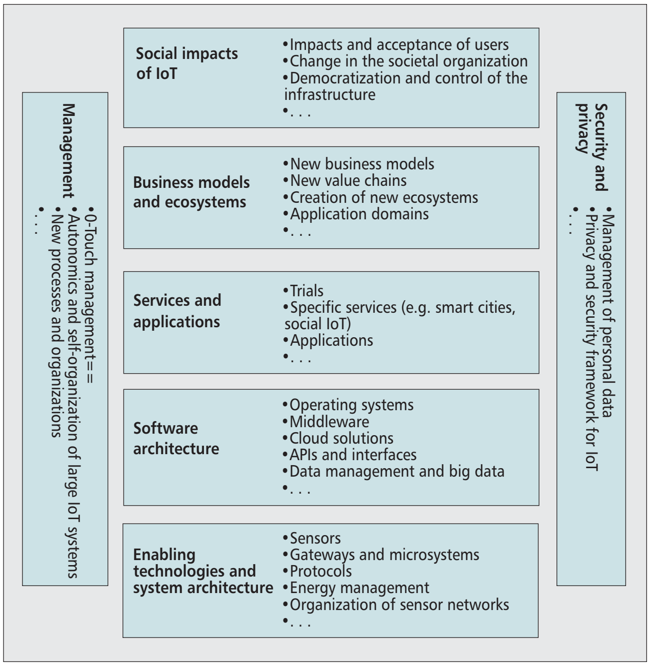


Figure 1. Technological facets of IoT.

Outside of publications produced by the IEEE, we've had 37 IoT expert bylines and articles published in industry magazines including *PC World*, *ComputerWorld*, *Intelligent Utility*, *Electronic Design*, and *Fierce SmartGrid*, encompassing topics such as our standards framework efforts, as well as our efforts toward building support for smart cities and smart grids. We have also received media coverage from CNN.

In addition, we've conducted a number of workshops that have gathered global IoT experts and other participants to explore new technologies, IEEE standards, applications, and future business models. So far workshops have been held in Silicon Valley, California; Shenzhen, China; and, Milan, Italy.

Guiding us forward in 2015 and beyond will be an IoT ecosystem study that will determine the connective areas and potential gaps in the concept of IoT that could be addressed through pre-standards and standards activities.

BUILDING THE IOT FRAMEWORK

A major aspect of the IEEE IoT Initiative is the definition of a framework for understanding what IoT is. This framework is based on the following pillars that will be developed in 2015 and will be consolidated in the future with the help of the IoT Technical Community: the IoT definition documents; the big data and cloud relationship with IoT; and the standardization effort toward a viable Architecture for IoT.

The set of documents comprising the IoT definition (to be published on the IEEE IoT Web Portal) will provide an analysis of the different technological facets of the Internet of Things.

Generally speaking, the IoT covers many areas (see Fig.

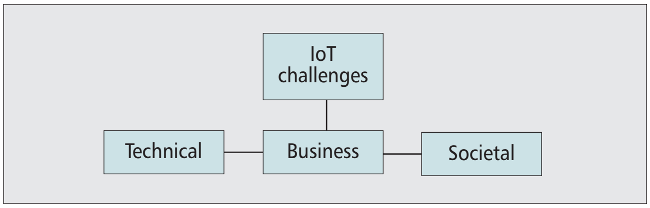


Figure 2. The several facets of IoT implications.

1) ranging from enabling technologies and components to several mechanisms to effectively integrate these low-level components. Software is then a discriminating factor for IoT systems. IoT operating systems are designed to run on small-scale components in the most efficient way possible, while at the same time providing basic functionalities to simplify and support the global IoT system in its objectives and purposes. Middleware, programmability — in terms of application programming interfaces (APIs) — and data management seem to be key factors for building a successful system in the IoT realm. Management capabilities are needed in order to properly handle systems that can potentially grow up to millions of different components. In this context, self-management and self-optimization of each individual component and/or subsystem may be strong requirements. In other words, autonomics behaviors could become the norm in large and complex IoT systems.

Data security and privacy will also play an important role in IoT deployments. Because IoT systems will produce and deal with personally identifiable information, data security and privacy will be critical from the very beginning. Services and applications will be built on top of this powerful and secure platform to satisfy business needs, so many applications are envisioned as well as generic and reusable services. This outcome will require new, viable business models for IoT and its related ecosystems of stakeholders. Finally, IoT can have an impact on people and the society they live in, and so it must be conceived and conducted within the constraints and regulations of each country. The technological analyses have led to the need to frame the IoT challenges into at least three categories (as depicted in Fig. 2) These challenges have been widely presented in several Conferences and in Webinars available on the IoT Initiative Portal by members of the IoT Initiative. [2]

These efforts to define the universe of discourse of IoT are synergic with the launch of a new standards project, IEEE P2413, "IEEE Standard for an Architectural Framework for the Internet of Things (IoT)," which will define an architectural framework for IoT, including descriptions of the various IoT domains, definitions of IoT domain abstractions, and identification of commonalities between different IoT domains.

IEEE P2413 will provide a robust architectural framework for IoT, reducing market fragmentation, improving interoperability, and serving as a catalyst for continued IoT growth and advancement. The first meeting of the IEEE P2413 Working Group took place in July 2014 in Munich, Germany. Its aim is to produce an IoT architectural framework and reference model, facilitating a unified approach to constructing critical IoT systems and infrastructures. The working group will also address key issues in security, privacy, and safety.

IoT is moving into its next stage of evolution, one that will require a data abstraction blueprint and a set of basic building blocks that can be easily integrated into multi-tiered systems. IEEE P2413 will provide the common elements needed to minimize industry and vertical market fragmentation, simplify implementation of cross-domain applications, and ensure that IoT achieves critical mass on a global scale. IoT has the potential to affect so many industries and people everywhere, with applications for smart cities, homes, and workplaces; e-health; resilient, self-healing power grids; and cleaner transportation. They will all benefit from the increased interoperability and portability that a standardized IoT architecture brings.[3]

IEEE will not be working alone. The IoT framework cannot be developed in a vacuum, so the IEEE P2413 working group will liaise with other standards bodies and standards development organizations in disciplines where IoT will have significant impact, such as manufacturing and healthcare. This collaboration with other key stakeholders will build industry consensus and accelerate development of underlying IoT frameworks and the overall innovation cycle.

LOOKING AHEAD

Creating an IoT architectural framework is obviously a significant project for the IEEE IoT Initiative in 2015 and beyond, but we will also continue forward with conferences and content development and community building initiatives. This includes a follow up to the 2014 WF-IoT, with ComSoc and the IEEE Computer Society acting as lead sponsors and organizers. Ultimately, we are looking to plan for versions of the conference for 2016 through to 2018 and cement it as the annual flagship IoT conference. We are also looking to participate and support eight to ten IEEE and non-IEEE IoT-focused and IoT-related conferences that include speaking engagements, panel sessions, and tutorials, while also getting involved as exhibitors, patrons, and sponsors.

On the content and community development front, it's full steam ahead as the IEEE IoT Initiative will leverage social media to further grow membership, visibility, and influence by making use of Twitter, LinkedIn, and Flipboard. Events will also make use of online tools, with webinars, Google Hangouts, and "Ask Me Anythings" (AMAs) planned. The Web Portal will continue to play a vital role with frequent updates that include interviews, articles, videos, and Q&As. We will also seek visibility through by-lined articles and columns and press interviews with non-IEEE publications and media outlets.

The IEEE IoT Initiative also has ambitious goals for publications in 2015 and beyond, including partnering with ComSoc and other IEEE Societies to explore the possibility of potential interest in an IoT magazine that will target computer, communications, and technology professionals and practitioners, as well as those with a business interest in IoT. This publication would cover topics such as social impacts, business models, services and applications, enabling technology and security and privacy. Market research is now underway.

Such a publication could be the foundation for bigger things, as the IEEE IoT Initiative will use it to integrate different contributions from IEEE societies into a meaningful stream of discussion, while we continue to publish and promote the *IoT eNewsletter* and migrate it from bimonthly to monthly, as well as continue to promote the *IEEE IoT Journal.*

Education will also continue to be an essential pillar of the IoT Initiative, and the IEEE will be developing a series of eLearning modules on IoT that make use of presentations from conferences and courses lectures within a university setting. We will also develop eLearning modules and white papers to inform policy makers of the scope and implications of IoT.

All of these activities under a single umbrella of the IoT Initiative and others across IEEE, provide a number of benefits to the IEEE. In the past, our IoT activities and offerings were siloed, making it difficult for members and non-members seeking a one-stop shop for IoT information and intelligence to navigate the various resources already available. The IoT Initiative activities fill the gaps in coverage and provide a unified IEEE view and will serve as the foundation to help build thought leadership that ties all of the elements together.

Ultimately, a major objective of the IEEE IoT Initiative is to nurture a community that can access the vast knowledge of the IEEE and the work that has already been done, so ComSoc is pivotal to that end. It is already strongly involved in activities that are focused on aspects of IoT or are closely related to it. As such, ComSoc has always been a major sponsor and contributor to the IoT Initiative.

Many of the key people involved in the IoT Initiative have come from ComSoc, including Dr. Roberto Minerva, the Leader of the IoT Initiative, Prof. Abbas Jamalipour, and Dr. Heinrich Stuettgen, serving on the Steering Committee and guiding the creation and development of the inaugural IEEE WF-IoT with many other ComSoc experts and professionals contributing heavily to its success. The Chair of ComSoc's IoT sub-committee, Prof. Latif Ladid, is also the General Chair of the 2015 conference, while Prof. Antonio Skarmeta and Prof. Antonio Jara are vice chairs of that sub-committee, and respectively, act as General Technical Committee Program Chair and Co-Chair of the Conference. ComSoc will also be the leading financial sponsor of the second edition to be held in Milan in December 2015.

Many other experts of ComSoc are actively involved in the IoT Initiative. An example of this cross-fertilization is the contribution of Prof. Luigi Atzori, Dr. Chung-sheng Li, and other colleagues who have provided content to the conference with a tutorial or speeches that have been transformed into articles for the *eNewsletter* or Webinars hosted in the IoT Portal and offered to the IoT Technical Community. ComSoc has also provided the editorial guidance for the *IoT eNewsletter*, with ComSoc member Dr. Raffaele Giaffreda serving as editor-in-chief. The *eNewsletter* now reaches a large number of experts worldwide. Creating high-quality content is critical to engaging more people in the community and increasing conference attendance, including use cases and examples of IoT implementations and systems deployment.

In addition, the scenarios definition as well as the documents related to the definition of IoT permeate the topics and challenges that Comsoc is tackling, and the list of contributions expected from the group will further expand with the progress of the IoT Initiative, covering several important IoT topics such as security, privacy, the softwarization trend, as well as the relationship and cross fertilization between IoT and software-defined networking, big data, cloud computing, and vehicular communication.

The contribution of ComSoc will help the Initiative reach its goal of being a thought leader for IoT by leveraging the wide range of knowledge of the IEEE Societies in this promising, world-changing research area.

OMBUDSMAN

COMSOC BYLAWS ARTICLE 3.8.10

The Ombudsman shall be the first point of contact for reporting a dispute or complaint related to Society activities and/or volunteers.

The Ombudsman will investigate, provide direction to the appropriate IEEE resources if necessary, and/or otherwise help settle these disputes at an appropriate level within the Society…

IEEE Communications Society Ombudsman
c/o Executive Director
3 Park Avenue
17 Floor
New York, NY 10017, USA

ombudsman@comsoc.org
www@comsoc.org "About Us" (bottom of page)

DATA-DRIVEN MODELING & SCIENTIFIC COMPUTATION: METHODS FOR COMPLEX SYSTEMS & BIG DATA

BY J. NATHAN KUTZ

OXFORD UNIVERSITY PRESS, 2013, ISBN 978-0-19-966033-9, PAPERBACK, 638 PAGES

REVIEWER: PIOTR GUZIK

Scientific computing is an extensive area that is hardly ever presented in a comprehensive and at the same time relatively simple way. This book is a notable exception. Though it contains vast theoretical knowledge, it is mainly focused on practical aspects. There are many valuable exemplary scripts written in MATLAB. They depict issues raised in this book, and can be used to implement a more sophisticated code.

The book is organized into four parts. In Part 1, the author starts with an introduction to MATLAB and continues with a basic course on scientific computing. The readers will learn how to solve linear systems or fit a curve to the data. The next chapters are dedicated to numerical integration and differentiation, the basics of optimization, that is methods that are crucial for numerous scientific problems, also met in the design and operation of communication networks. Finally, Part 1 is concluded with aspects of results visualization. All of the chapters contain simple yet instructive examples written in MATLAB.

In Part 2 the author expounds on variety of problems based on differential and partial differential equations. Numerous methods of solving such equations are thoroughly described, each accompanied by practical, easy to understand examples. Apart from the basics of differential and partial equations, the reader will also learn how to use some more sophisticated tools, namely: difference methods, spectral methods, and finite element methods. All of them are explained clearly and each eager student should not have much trouble understanding or implementing them in MATLAB.

Later, in Part 3 the author proceeds to data analysis aspects. Here the readers will learn methods that are currently used in communications and computer networks, especially the basics of machine learning, image recognition, time-frequency analysis, and much more, although most of these topics are covered only briefly since each of them constitutes such a broad field that they deserve another book. Finally, in the last part, scientific applications are analyzed. It is a pity that this part is the shortest one. The problems that are considered here are stated clearly and the solutions based on the material presented in the preceding chapters seems very straightforward, though they definitely are far from being simple.

In conclusion, the level of mathematics required to understand this book is not beyond the level of undergraduate student, and it may serve as a valuable introductory textbook to widely understood area of scientific computation and data (mainly multimedia data) analysis. Readers familiar with those topics will also find some inspiring issues, though they would probably not find this book a must have. The problems, although not easy to understand in each case, are presented in an approachable manner, so that even readers who are just at the beginning of their journey into scientific computation and data analysis are going to get through this book relatively easily.

INTERNET PROTOCOL-BASED EMERGENCY SERVICES

BY HANNES TSCHOFENIG AND HENNING SCHULZRINNE (EDS.)

WILEY, 2013, ISBN 978-0-470-68976-9, HARDCOVER, 367 PAGES

REVIEWER: MARCIN NIEMIEC

The Internet has changed almost all aspects of our daily lives. Therefore, it is a bit surprising that emergency services are still offered using mainly legacy communications. In many regions of the world, it is still impossible to establish video calls, transmit pictures, send GPS location, or use instant messaging to start real-time communication with a call-taker at a Public Safety Answering Point. This book, edited by Tschofenig and Schulzrinne, summarizes the efforts and presents ongoing developments to enable IP-based communication systems to offer emergency services.

Although the work is a collection of the texts of over 30 authors, the volume has a well-organized and readable form. It is organized in 10 chapters and each chapter can be read independently. The first chapter illustrates the history, explains the reasoning behind writing the book, and provides a short overview with the main building blocks: recognizing, locating, and routing emergency calls. The next chapter presents the location issues that are crucial to proper emergency call routing and precise dispatching of the first responders. Two types of location are considered: civic and geospatial. Additionally, location information encoding and the most important protocols are introduced. Chapter 3 focuses on emergency services architectures and illustrates how various building blocks work together. The authors describe NENA i2 and NENA i3 systems supporting emergency calls from VoIP devices. Also, the multimedia-based emergency services architecture developed by an IETF Working Group is presented. Finally, the emergency architectures in different types of networks are explained: WiFi, WiMAX, and 3GPP.

After an introduction to location protocols and architectures, the authors present the four deployment examples: VoIP-based emergency services in Sweden, the United Kingdom, Canada, and the U.S./Indiana. They illustrate the complexity of deployment of the next-generation emergency services architecture. Chapter 5 is devoted to security issues, which are a serious problem because attacks consume valuable resources of emergency service systems. Security threats are described, as are a number of techniques to mitigate these threats. Chapter 6 covers emergency services for people with communications-related disabilities. In Chapter 7 the relevant regulations in the U.S. and Europe are illustrated. Chapter 8 introduces the three state-funded projects that support the research work on IP-based emergency services: the US NG9-1-1 Initiative, the UE-funded REACH112, and the PEACE Projects. In Chapter 9 a number of organizations involved in standardization of emergency services are presented. The last chapter summarizes the book. The conclusion contains a brief summary as well as remarks on other important issues, such as the funding of emergency services.

The great advantage of this book is that it elaborates on many different aspects of IP-based emergency services: architecture, protocols, security, standardization, regulations, deployment examples, and research projects. Although the editors have written that this book focuses on the communication chain from an emergency caller to the call-taker who can dispatch emergency personnel, the scope is in fact much wider.

This work can be recommended for academic researchers and graduate students in communications and computer science. It is not only a useful state-of-the-art, but it also discovers information on organizations and communities involved in emergency services as well as allows the readers to uncover still existing gaps in this field. In addition, practitioners, product architects, and developers will find interesting and useful ideas. Many parts of the book can be recommended to experts working on standards and regulations.

CLOUD NETWORKING: UNDERSTANDING CLOUD-BASED DATA CENTER NETWORKS

BY GARY LEE

ELSEVIER, 2014, ISBN 978-0-12-800728-0, 240 PAGES

REVIEWER: RADOSLAW SMIGIELSKI

Cloud data center networking is a complex puzzle with many blocks to combine. The process of building a cost-effective, applicable data center network at manageable scale is what the author is trying to guide the reader through. This books covers all of the important aspects that need to be taken into consideration during the design of cloud data center networks.

Chapter 1 covers basic networking concepts, starting from the OSI network stack definition and finishing with a short TCP/IP recap. Data center and networking technology evolution is covered in Chapter 2. The author begins from early mainframes and ends with the latest shift toward clouds. In a similar way he leads a reader through the history of network standards, from ARPANET and finishes with today's fast Ethernet. Chapter 3 is a concise overview of the switch fabric architecture, plus flow control, traffic, and congestion management. Modern cloud data center networks and server technologies are the main subject of Chapter 4. It also explains the advantages and disadvantages of the most common architectures. Different types of switches and servers are presented from the perspectives of performance, scalability, and cost. A flat data center model is discussed with details, and the concept of Network Function Virtualization (NFV) is introduced. The next chapter elaborates on inter-switch and inter-fabrics communication, flow control mechanisms, and effective link bandwidth utilization standards, including ECMP, TRILL, and SPB. RDMA is presented as a way to improve latency and application performance. Chapter 6 contains a short overview of two virtualization platforms: VMware and Microsoft HyperV. It is followed by a review of software virtual switches and the method for ensuring network connectivity to VMs. The characteristics of traffic generated by virtual machines are presented with a deeper look at live migration. As network virtualization and multi-tenancy have become the mandatory requirements in today's data centers, an outlook on the most popular and most promising separation, tunnelling, and load balancing methods can be found in Chapter 7. Chapter 8 is dedicated to storage technologies used in modern data centers. The following technologies are characterized: (i)SCSI, Fibre Channel, FCoE, NFS, convergent networks, and distributed storage. New initiatives such as Software Defined Datacenter, Software Defined Network, Network Function Virtualization versus the existing multi-tenancy standards like Q-in-Q, MPLS, VXLAN, and NVGRE are discussed in the next chapter. Specifically, OpenStack and Open Daylight are taken into consideration as existing SDN examples. Chapter 10 leads a reader into high performance computing (HPC), describing the related networking, computing, and architectural challenges. There is also a discussion of the current status of the Ethernet versus Infiniband competition. Chapter 11 outlines a set of the most interesting and promising future technology trends in data center construction. The final chapter is the author's recap of the book's content.

This work is a comprehensive review of cloud network technologies and other related topics. Each chapter usually contains a historical introduction, architectural considerations, and quite often a section on the financial aspects of the presented matter. Lee covers the existing well established technologies we can find in the today's data centers, but he also deals with the bleeding edge trends (the mentioned Software Defined Datacenter, etc.) in a very comprehensive manner. The author's goal to provide the reader with a broad overview of cloud networking technologies is surely accomplished. The author's unique and lengthy experience in network silicon design enabled him to write this book in a genuine way. Lee shows how all the network technologies are closely linked together and how they depend on semiconductor technology and its progress. Moreover, he is able to explain quite complex issues in an easy to understand way, although basic networking knowledge is required from the reader. Lee keeps a well balanced opinion on the current status and future of cloud networking and cloud computing.

IEEE WCNC 2015 TO EXPLORE ENTIRE RANGE OF NEXT-STAGE WIRELESS TECHNOLOGIES MARCH 9–12, 2015 IN NEW ORLEANS, LOUISIANA

LEADING INTERNATIONAL RESEARCHERS AND EXPERTS TO SHARE EXPERIENCES AND EXPERTISE ON NEWEST WIRELESS COMMUNICATIONS AND NETWORKING TRENDS, BENEFITS, AND APPLICATIONS

The IEEE Wireless Communications & Networking Conference (WCNC 2015) will hold its next annual international event dedicated to the advancement of next generation wireless communications and networking technologies from March 9 – 12 at the Hilton New Orleans Riverside hotel in New Orleans, Louisiana. IEEE WCNC 2015 will include more than 500 presentations consisting of keynotes, workshops, tutorials, panels and technical presentations devoted to the entire range of next-stage digital cellular, PCS, multimedia, channel modeling, radio resource management, green networking and emerging mobile applications.

"IEEE WCNC offers global communications experts the opportunity to explore the full-spectrum of next wave wireless broadband, Internet, interconnected software and radio networking technologies in a setting designed to promote learning and establish professional relationships that will span decades," says Richard Miller, IEEE WCNC 2015 Executive Chair. "It will be our honor to once again feature the presentations of many of the finest researchers and scientists from industry and academia, while witnessing the ongoing progress of their efforts to further the quality of everyday lives worldwide."

IEEE WCNC 2015 will open on Monday, March 9 with eight tutorials and seven workshops offered by leading international researchers and industry experts representing the United States, China, the United Kingdom, Switzerland, Japan, Taiwan, France, Finland, Sweden and Germany. This will include half- and full-day workshops dedicated to topics like 5G architectures, next generation WiFi technologies, device-to-device and public safety communications, Internet of Things energy efficiencies, cooperative and heterogeneous cellular networks, electric vehicles in the smart grid and vehicular clouds from connected cars to smart cities.

On the following day, the conference will officially commence with nearly 1,000 industry professionals, academics and government officials expected to attend the presentation of hundreds of panel discussions, technical sessions, networking events and executive forums exploring the latest advances in the wireless field. This will be highlighted by the keynote addresses of numerous world-renowned researchers and corporate communications leaders such as:

• Victor Bahl, Principal Researcher and Director of the Mobility and Networking Research (MNR) group, Microsoft Research, who will speak on "Cloud 2020: The Emergence of Micro Datacenters for Mobile Computing" and his vision for a "new world of disaggregated clouds in which seamless cognitive assistance for wireless users can be delivered anyplace, anytime using the latest and greatest computer science technologies".

• Richard Gitlin, Agere Systems Chair of Electrical Engineering, University of South Florida, who will address the latest "In Vivo Wireless Communications and Networking" innovations and its potential for advancing health care delivery and improving the effectiveness of emerging cyberphysical biomedical systems focused on minimally invasive surgery and cardiology procedures.

• Lajos Hanzo, Faculty of Physical Sciences and Engineering, University of Southampton, United Kingdom, who will offer his views on "Aircraft Ad-Hoc Networking for The "Internet-Above-The-Clouds", for "Free-Flight" and All That..." This will include the review of radically-improved Aeronautical Ad-hoc Networks (AANET) designed to provide more dynamic topologies, greater variable geographical network sizes and stricter security for in- flight entertainment networking solutions used within high-velocity aircraft.

• Kevin Jou, Corporate Vice President and Chief Technology Officer, MediaTek Inc., who will take "A look at trends in mobile technologies" and the advance of the mobile communications systems needed to accommodate increasingly faster computing and data rate speeds.

• Seizo Onoe, CTO, EVP and Member of Board of Directors, NTT DOCOMO, who will discuss "5G and Beyond," the current status of LTE and LTE-Advanced, and the combination of technologies that will provide cost-effective cellular system solutions with wide coverage and broader bandwidth in 2020 and beyond.

• Jerry Pi, Chief Technology Officer, Straight Path Communications Inc., who will detail "Innovation to Transformation – Next Generation Mobile Broadband as the Infra- structure for a Connected World." Highlighted will be the development of the innovative technologies necessary to transform current cellular and Wi-Fi networks into a communication infrastructure earmarked by 100 billion connected devices and an 1000x increase in mobile broadband by the next decade.

In addition, the conference will be highlighted by several senior-level executive panels addressing the areas of "5G Testbeds, experiments, demonstrations" and "5G Vision and Requirements." For instance, various experts will demonstrate the new capabilities of advanced 5G infrastructures for meeting latency, reliability and spectrum requirements, while Werner Mohr, the Head of Research Alliances in Nokia, Munich, Germany and Chair of the board of 5G Infrastructure Association, Belgium, will lead the panel exploring future 5G communications networks from the European perspective. This includes potential future frequency ranges, global standardization and the seamless integration of different existing access systems for very wideband services, ultra-dense deployment and IoT systems.

Other IEEE WCNC 2015 highlights will be the presentation of 391 technical papers detailing the latest innovations in PHY and fundamentals, MAC and cross-layer design, mobile and wireless networks and services, applications and business programs. As an example, specific subjects addressed will be "Cloud and Health Networking," "Delay-oriented Dynamic Control for Interactive Mobile Cloud Services," "Multi-pair Device-to-Device Communications with Space-Time Analog Network Coding," "GPS Tethering for Energy Conservation," "A Mobility Management System for Mobile Cloud Computing," "A Novel Handover Scheme to Support Small-cell Users in a HetNet Environment" and "Relaxed Channel Polarization for Reduced Complexity Polar Coding."

For more information on IEEE WCNC 2015, including registration and patronage details as well as ongoing conference updates, please visit www.ieee-wcnc.org/2015. All visitors to the IEEE WCNC 2015 website are also invited to reach out to friends and colleagues through the conference's Twitter, Facebook and LinkedIn pages.

SIEMENS

E20001-F730-P820-X-7600

Industrial Ethernet Switches

Switch to the future: SCALANCE XM-400

Secure and reliable data transmission is indispensable for industrial communication networks – not only today, but in the future as well. That makes it all the more important to be equipped with technology capable of keeping up with increasing requirements – as it is then much simpler to adapt or expand a system.

Industrial Ethernet Switches SCALANCE XM-400 are our innovative answer to the challenges of tomorrow – their modular design, fast mobile diagnostics for smartphones/tablets via NFC and existing WLAN, and upgradeable Layer 3 functionality using KEY-PLUG will ensure that your automation network is powerfully and flexibly equipped for the future.

siemens.com/x-400

GCN GLOBAL COMMUNICATIONS NEWSLETTER

February 2015
ISSN 2374-1082

MEMBER RELATIONS

Europe, Africa and Middle East Region

Interview with Hanna Bogucka, Director of the EAME Region

By Stefano Bregni, Vice-President for Member Relations, and Hanna Bogucka, Director of the Europe, Africa and Middle East Region

This is the sixth article in the series started in September 2014 and published monthly in the *Global Communications Newsletter*, which covers all areas of IEEE ComSoc Member Relations. In this series of articles, I introduce the seven Member Relations Directors (namely: Sister and Related Societies; Membership Programs Development; AP, NA, LA, EAME Regions; Marketing and Industry Relations) and the Chair of the Women in Communications Engineering (WICE) Standing Committee. In each article, one by one they present their activities and plans.

In this issue, I interview Hanna Bogucka, Director of the Europe, Africa and Middle East Region. Hanna is a full professor on the Faculty of Electronics and Telecommunications at Poznan University of Technology in Poland. Her main research area is wireless communications, in particular flexible, adaptive, and cognitive radio. Hanna has been the Technical Program co-Chair or a committee member of a number of IEEE ComSoc conferences, and guest editor and reviewer of various ComSoc journals. She serves the IEEE ComSoc community as the Director of the EAME Region (Europe, Africa, Middle East) for the 2014-2015 term. She is a member of the IEEE Kiyo Tomiyasu Award Committee, and recently she has been elected the ComSoc Radio Communications Committee Chair for the term of 2015-2016. I am glad to interview Hanna and to present the organization and activities of the Europe, Africa and Middle East Region.


Stefano Bregni


Hanna Bogucka

Bregni: Hello Hanna. Let us begin by presenting the EAME Region and how your Board is organized.

Bogucka: Geographically, EMEA is the largest IEEE ComSoc region. It has 52 chapters and approximately 11,800 members, about one quarter of the ComSoc members worldwide. Since the beginning of my term in January 2014, on forming the EMEA Region Board, our operating plan was to promote the development of membership, expand the distinguished lecturers tours program, implement a young researchers program, stimulate academia-industry cooperation, increase the number of ComSoc conferences, and expand social networks in the EMEA region. The EMEA Region Board is organized to implement this plan, and consists of representatives from industry, academia, governmental institutions, from Europe and Africa, men and women, with four operational members and two advisors.

Stefano: You mentioned that among your priorities you targeted membership development in the EAME Region. What are the challenges to pursue this goal?

Bogucka: Development of membership is particularly challenging in Region 8. There is visible disproportion of geographical coverage and membership percentile in our region. This is mainly due to the particularly small number of members in African and developing countries, especially in Sub-Saharan Africa.

Bregni: I agree that Africa poses very peculiar issues not only related to membership development, but also about how to serve the specific needs of members in developing countries, while keeping membership fees low enough to be affordable in disadvantaged economies.

Bogucka: At the IEEE level, an Africa Committee is in place, which looks at this problem, negotiates prices, and offers subscription rates acceptable to African universities. In the Communications Society, we try to promote new memberships by focusing on information campaigns, by favoring Distinguished Lecturers Tours (DLTs), and by promoting the organization of conferences in Africa. It is important to identify and remove obstacles to the integration of African researchers into the IEEE ComSoc community.

Bregni: Overall, what is your general perception on the activity of Chapters in the EMEA Region?

Bogucka: In general, the situation in EMEA is very good in terms of Chapter activities and membership development. I appreciate seeing many activities at the Chapter level. Every year, we select a particularly active chapter to receive the Chapter Achievement Award. This year the award was given to the Republic of Macedonia Chapter and presented at Globecom 2014 in Austin, TX, USA.

Bregni: You highlighted the importance of the Distinguished Lecturer/Speaker Program in Africa and other disadvantaged areas. You know that I am very sensitive to this topic, having served as Expert/Distinguished Lecturer for seven years. What are your objectives and achievements in the EMEA Region in this regard?

Bogucka: The Distinguished Lecturers Program (DLP) and Distinguished Speakers Program (DSP) are important programs in the EMEA Region, allowing for cross-border knowledge and the spreading of expertise. In 2014, six DLTs and DSPs have taken place, while two more have been organized but postponed until 2015. As an example, the very successful DLT of Prof. Norman C. Beaulieu in Macedonia, Italy, Montenegro, and Serbia (March 2014) was reported in an article in the August 2014 issue of the *Global Communications Newsletter*. Every time the distinguished lecturers give their lectures, they promote ComSoc to the listeners by presenting few slides on IEEE ComSoc activities. I do hope this helps membership development, since the lectures gather a lot of interest from local researchers, not necessarily ComSoc members yet.

Bregni: Would you like to highlight any other program organized in the EMEA Region?

Bogucka: A rather unique activity in our Region is our EMEA Young

(Continued on Newsletter page 4)

DISTINGUISHED LECTURER TOUR

Distinguished Lecturer Tour of Fabrizio Granelli in Japan, 2014

By Fabrizio Granelli, University of Trento, Italy

The Distinguished Lecturer Tour was organized in Spring 2014, and it was supported by all three Communications Society Chapters in Japan (Asia/Pacific Region 10): the Japan Council, the Kansai Chapter, and the Sendai Chapter.

The DLT started in Osaka, in the Kansai Region of Japan, on July 15, 2014. The lecture was held in a conference center near the main Osaka train station. The seminar was held in the early afternoon, with a good participation (approximately 25–30 people). The chapter expressed interest for a seminar organized in two sections: first, a brief description of my research interests, and then a longer presentation on the subject of 'Networking and the Smart Grid'. After presenting my research interests and major projects, I focused on an introductory section about what is the Smart Grid and its main expected features. Then I presented an overview of the existing communication standards related to the Grid, with specific attention to IEEE P2030 sponsored by ComSoc, which provides an effective taxonomy of the major components of the communication infrastructure to support the operation of the Smart Grid. Finally, I decided to introduce examples of research activities where communications know-how can contribute to Smart Grid development, e.g. modeling micro-grids and studying charging station setup for electrical vehicles.

With the local organizers in Osaka.

With the lecture attendees in Sendai.

During the seminar in Osaka.

We had an interesting Q&A exchange at the end of the session, mainly focused on issues related to actual deployment issues of smart grid architecture and the relationship between Smart Grid and D2D communications.

The second seminar was held in Tokyo, at the Kinkai Shinko Kaikan Building, close to the Tokyo Tower landmark. The Japan Council organized a two-hour seminar on two topics: 'Green Wireless Networking' and 'Cognitive and Adaptive Networking'. I started the former presentation by explaining the high power consumption associated with communication equipment, supported by some figures depicting the contribution of the different modules. As of today, the major problem in green wireless networks is due to the high power consumption associated with power amplifiers and thus on the effective utilization of the RF interface. Then I described the major classes of solutions to reduce power consumption in wireless networks, and focused on some examples of how to effectively use the sleep state in WLANs and how to address power consumption in emerging technologies such as cognitive radios.

After a short break I moved to the second topic, where the focus was on bridging the gap between theory and practice in making cognitive networks a reality. Indeed, I introduced the audience to some recent advances of my research group on including reasoning and learning mechanisms able to introduce cognitive functionality in today's networks.

The seminar was attended by approximately 50 people. What was very interesting was the fact that on this occasion the audience was more mixed than usual, and included students, postdocs, and professors, as well as several researchers from industry (NEC, IBM, DoCoMo). The audience was extremely interested and they offered several comments. In particular, the latter topic spread interest among participants familiar with Artificial Intelligence and Automated Learning.

The last seminar was hosted by the Sendai Chapter, in the North Region of Japan, at Tohoku University. The topic was again 'Green Wireless Networks', and based on the audience feedback during the presentation, I provided additional details on greening WiFi extensions and some ideas on how to measure the actual power consumption of real systems.

There were approximately 20 attendees, mostly Ph.D. and M.Sc. students. The audience was extremely interested and the follow-up discussion was fruitful. I received several questions and requests for pointers to further details on several topics I covered.

REGIONAL REPORT

The Republic of Moldova is Preparing for the EU!

An Interview with Pavel Filip, the Minister of Information Technology and Communications of the Republic of Moldova

By Nicolae Oaca

Oaca: How is your ministry contributing to the EU integration process?

Filip: According to the Association agreement between Moldova and the European Union, signed on 27 June 2014 in Vilnius, our country has to harmonize its legislation in order to be in line with the EU telecommunications regulations. So we already have started to transpose EU's regulations for telecommunications to our legislation. However, there is a long way to go as we have to deal with several issues. We need to revise and complete the electronic communications services regulations according to the universal services Directive. Also, we have to implement the EU Regulations concerning international roaming, net neutrality, personal data protection, the level of confidentiality of electronic communications, etc.

To be in line with the EU, our Ministry, MITC, has developed a national strategy called 'Digital Moldova 2020' aiming to create a national broadband network, to provide our citizens and our businesses with high speed connections, at least 30 Mb/s, and therefore enable information society development. 'Digital Moldova 2020' was developed according to the Digital Agenda for Europe 2020, and the best world practices being publicly presented in 2013. In fact, 'Digital Moldova 2020' is based on three pillars: broadband infrastructure and access for all citizens; digital content and electronic services; and finally, people's capacity for general use of the Internet and information technology. In order to implement this strategy, MITC proposes a plan that aims to develop the national broadband network until 2020; to establish at least one presence point for optical fiber on the territory of every locality: to provide buildings with broadband infrastructures; and to update the existing cable network for broadband access. We hope 'Digital Moldova 2020' will be a driver of our national economy and welfare.

In 2013 we also developed a program for spectrum management that will last until 2020 aiming to ensure better management of the existing spectrum and to provide enough spectrum for high speed networks and new technologies. Another goal standing behind this program is the harmonization between our regulations and practices and those of the EU. Our program is based on three pillars: technology neutrality that allows the usage of any technology able to create business oriented networks; spectrum auctioning as a clear and transparent way to award frequency licenses; and crystal clear auction conditions on prices and terms. We already awarded frequency licenses in the 800, 900, 1800, 2100, 2600 and 3800 MHz bands in 2014 for the next 15 years, and it is time for the operators to use them and to comply with the new rules.

Oaca: What are your plans for the analogue TV switchover and digital dividend band usage?

Filip: We elaborated on a program regarding analogue TV switchover planned for 17 June, 2015. Its aim is to improve TV quality, and mainly to free up the digital dividend band, the 800 MHz band, to be allocated to mobile communications in 2014 after a competitive and transparent contest, aiming to attract a new competitor.

Oaca: Who are the main operators In your telecommunications market?

Filip: As in other countries in our region, our mobile communications and Internet markets are continuing to grow while fixed telephony is decreasing. The number of active SIMs reached approximately 3.7 million, and our mobile network operators, Orange, Moldcell, and Moldtelecom, cover about 99 percent of the territory. Also important to mention is that Internet usage reached 13.5 percent of the population and continues to grow, fueled by investments in optical fiber and 3G and LTE networks.

Pavel Filip, the Minister of Information Technology and Communications of the Republic of Moldova.

Oaca: Can you please speak about competencies of different bodies in telecommunications area?

Filip: MICT is the national strategy body for electronic communications in Moldova, in charge of developing political and strategic documents. It does not own or operate any telecommunications operator. NRFC is a state entity subordinated to MICT, in charge of frequency monitoring. Also, ANRCETI is an independent authority responsible for regulating, monitoring, and developing the telecommunications market. ANRCETI also issues spectrum licences. Our national telecommunications company, Moldtelecom, is owned by an entity subordinated to the Ministry of Economy.

Oaca: What are the MITC's medium-term challenges and plans?

Filip: We could face a complex and fragile political and economical context, which could hamper faster development of our market or delay investments. However, we are able to face all these challenges, to develop the informational society in Moldova via 'Digital Moldova 2020' and to regulate our market as efficiently as possible. Stable and predictable rules are very important for market development, so we are thinking of a medium-term strategy to attract investments, to protect end-users, and to use our scarce resources in a better way. Moldova has to find its own way to meet the European Union requirements, bearing in mind the ability of our national market and industry to adjust to dynamic changes. We have to accelerate the institutional reform process in order to increase confidence and to attract investments in our telecommunications market to accelerate our economic growth.

Nicolae Oaca's note: In the Republic of Moldova, a significant evolution is occurring with the Internet. Cable and optical fiber technologies provide Moldova's citizens with up to 1 Gbps speeds. In 2013, thecountry had a 13.2 percent penetration of Internet users. The mobile Internet is quickly expanding as well, reaching 7.29 percent penetration in December 2013. And finally, last year the country successfully launched fixed and mobile telephony number portability, aiming to increase competition, service quality, and end user churn. A total of 44,720 numbers were ported and changed networks, made up of 40,376 mobile numbers and 4,344 fixed telephone numbers.

REGIONAL REPORT

2014 IEEE Sections Congress in Amsterdam

By John F. Pape, Director of Marketing and Creative Services, IEEE Communications Society

The 2014 IEEE Sections Congress was held in Amsterdam, 22-24 August, 2014. The IEEE Member and Geographic Activities (MGA) Board partnered with Region 8 to create an ambitious program themed 'Inspiring Our Leaders of Tomorrow.' The three main program tracks covered enhancing member satisfaction, improving the volunteer experience, and reaching globally with the local touch.

Over 1,000 section officers, volunteer leadership, and staff convened to attend training sessions, network with colleagues, learn about new and existing IEEE programs, and make recommendations to guide the future of IEEE.

Thirty-five Learning Labs Sections were aimed at attendees wanting to learn more about a particular IEEE tool. These 10- to 25-minute sessions gave live demonstrations on resources from MGA, Technical Activities, Educational Activities, and other IEEE organizational units. These included How Your Section Can Benefit from IEEE-ETA KAPPA NU (IEEE-HKN)–the IEEE Student Honor Society, vTools, and the IEEE Awards Program.

The exhibit space offered attendees the opportunity to share a cup of coffee, investigate IEEE programs, chat with Society staff, explore career development resources, consider IEEE.tv, discuss publication and conference issues, meet old friends, and establish new relationships. There were over 40 IEEE exhibitors. If attendees were trying to locate the Young Professionals booth or a session room, the Sections Congress smartphone app made planning, scheduling, communicating, and connecting easy.

The exhibit floor also featured over 30 IGNITE sessions of five-minute presentations on specific topics intended to spark interest in the subject, such as Women in Engineering, Section/Chapter Sponsored Conferences, TAB/MGA Collaborations, and the IEEE Consultants Network.

The IEEE Communications Society staffed a booth and featured recorded presentations on the ComSoc society overview, the online training program, and social media activities. Many current ComSoc Chapter officers and members stopped to confer, from places as diverse as Kerala, Hyderabad, Bahrain, Ottawa, Cedar Rapids, and several from Region 5, including Scott Atkinson, who avidly promoted IEEE GLOBECOM 2014.

Among the many highlights, previewing the Professional Productivity and Collaboration Tools (PPCT) – now known as IEEE Collabratec? — was a favorite. It is an integrated suite of online features and productivity tools that facilitate communication, collaboration, and networking. Scheduled to launch in 2015, it will create new opportunities to engage and interact for sections, chapters, volunteers, members, and affinity groups.

At the closing ceremony, the results of voting for SC recommendations are presented. The top recommendations included: increase access to the IEEE Digital Library as a member benefit; develop an incentive program for companies that support employee membership dues; and introduce loyalty awards for continued membership. The next Sections Congress will be held in 2017.

GCN GLOBAL COMMUNICATIONS NEWSLETTER

STEFANO BREGNI
Editor
Politecnico di Milano — Dept. of Electronics and Information
Piazza Leonardo da Vinci 32, 20133 MILANO MI, Italy
Tel: +39-02-2399.3503 — Fax: +39-02-2399.3413
Email: bregni@elet.polimi.it, s.bregni@ieee.org

A publication of the IEEE Communications Society
www.comsoc.org/gcn
ISSN 2374-1082

MEMBER RELATIONS/*Continued from page 1*

Researchers Award program. Every year, the Young Researchers Award Program Committee evaluates the applications of many candidates and then assigns one IEEE ComSoc Young Researcher Award of EMEA Region (consisting of a plaque and 500 $US) and two runner-up Outstanding Young Researchers Awards (consisting of a plaque and 250 $US each). Awards are presented at one of the two flagship IEEE ComSoc conferences, ICC or Globecom. The applications are collected every year in November. The winners are selected in January of the following year. We believe this is a particularly valuable initiative, as we have the pleasure to recognize so many outstanding scientific achievements of young researchers in our region.

Bregni: Is there any particularly successful ComSoc conference in EMEA? And what about regional conferences?

Bogucka: We support the organization of many conferences in EMEA, especially conferences from the ComSoc portfolio, as well as dedicated events at major IEEE Region 8 conferences, for example Africon or Sibircon, which are an opportunity to present ComSoc programs. Moreover, IEEE BlackSeaCom (IEEE International Black Sea Conference on Communications and Networking) is a new series of ComSoc Regional Conferences held in the countries surrounding the Black Sea, with the goal of bringing visionaries in academia, research labs, and industry from all over the world to the shores of the Black Sea.

The 2015 edition of BlackSeaCom will take place on May 18-21, 2015 in Constanta, Romania. I believe it is important to have such a cyclically organized conference in that area. It can be a nice driver to increase Chapter activity and integrate scientific research in the region of South and Eastern Europe. Let me take this opportunity to invite everyone to this very special conference in EMEA, listed in the IEEE ComSoc conference portfolio.

Finally, there are many other important events happening in EMEA beyond scientific conferences. In 2013 at ICC in Budapest, Hungary, the EMEA Regional Chapters Chairs Congress was organized by the former EMEA Director, Fambirai Takawira. In August 2014, the Krakow Chapter, Poland, hosted the IEEE Region 8 Student & Young Professional Congress. Also in August 2014, a particularly nice ceremony took place in Warsaw: the IEEE Milestone in Electrical Engineering and Computing to commemorate the First Breaking of the Enigma Code by the Team of Polish Cipher Bureau, 1932-1939. These are just a few good examples.

Bregni: What are your plans for 2015?

Bogucka: Our operating plan for the next year remains the same as for the whole term of my directorship, mentioned at the beginning. It is an ambitious long-term plan that I hope will be continued by our successors. We just wish to achieve intermediate goals as much as possible on this long-term vision for ComSoc in EMEA.

Taking the opportunity of being invited to do this interview at the beginning of the New Year, let me wish the GCN readers and our ComSoc community a successful realization of their professional and private plans in 2015!

SPATIALLY AND SPECTRALLY FLEXIBLE ELASTIC OPTICAL NETWORKING


Ioannis Tomkos


Yutaka Miyamoto


Glenn Wellbrock


Peter J. Winzer

The traffic carried by core optical networks as well as the per-channel interface rates required by IP routers are growing at a remarkable pace year after year. This trend is due to the widespread deployment of fixed and wireless broadband access networks, the huge growth of video-based services supported by the Internet, and social media applications, as well as a substantially growing amount of machine-to-machine traffic supporting a variety of data-centric applications. Optical transmission and networking advancements have been able to satisfy this huge traffic growth so far by providing the necessary network infrastructure in a cost- and energy-efficient manner, utilizing to the maximum extent the capabilities of optoelectronic and photonic devices across the available bandwidth of deployed optical fibers. However, as we are rapidly approaching fundamental spectral efficiency limits of single-mode fibers, the scientific and industrial telecommunications community foresees that the growth capabilities of conventional WDM networks operating on a fixed frequency grid in conventional wavelength bands are quite limited.

To address such limitations, a number of significant innovations have appeared over the past years that offer a capacity increase of around 10–20× in the latest generation of commercial optical transport systems compared to legacy WDM systems operating 10 Gb/s wavelengths on 50 GHz spacing. Initial efforts have targeted innovative modulation and coding techniques and flexible frequency allocations to increase the spectral density in optical fiber links, eventually leading to the definition of spectrally flexible/elastic optical networks utilizing optical super-channels together with spectrally flexible/elastic multiplexing schemes (e.g., Nyquist WDM or optical OFDM). However, while the spectrally flexible/elastic super-channel approach can optimize network resources through increased spectral utilization and increase interface rates through the efficient bundling of subcarriers on a single transponder card, it has limited growth potential in terms of per-fiber capacities due to the nonlinear Shannon limit imposed on the transport capacity of single-mode optical fiber within the limited gain bandwidth of optical amplifiers. Multi-band amplification technologies (e.g., C+L+S-band amplifiers) may yield temporary relief, but the only evident long-term solution to extend the capacity of optical communication systems relies on the use of the spatial domain. The simplest way to achieve spatial multiplexing is to deploy multiple optical line systems in parallel. However, by simply increasing the number of systems, the cost and power consumption also increase linearly. In order to continue reducing the cost and energy consumption per bit in optical transport, component sharing and integration have to be introduced. To this end, significant research efforts have focused on the development and performance evaluation of few-mode fibers (FMF) and multi-core fibers (MCF), which can be seen as integrated fiber media, for space-division multiplexed (SDM) systems. This line of work is further supported by the development of integrated optical amplification systems, as well as the significant development efforts in the field of terabit-per-second integrated transponders for SDM systems. For such systems, the use of spatial super-channels, which are groups of same-wavelength subchannels that are transmitted on separate spatial modes routed together through the network, are being investigated. Finally, significant efforts are being made on the development of the proper control plane framework to orchestrate the operation of such spectrally and spatially flexible networks in order to bring out their full potential (i.e., besides capacity increase to also support other capabilities like network virtualization).

This *IEEE Communications Magazine* Feature Topic aims to provide a comprehensive overview of the state of the art in spatially and spectrally flexible elastic optical networking, with respect to network architectures, network design and operational issues, transmission systems, and subsystem/device technologies. The articles presented here provide a holistic view of the projected commercial needs and requirements for the transport network, the state-of-the-art achievements by various research groups across the globe investigating the relevant topics, as well as the current research challenges and opportunities in this emerging scientific area.

All articles that appear in this Feature Topic are written by world-renowned researchers in the field of spectrally and spatially flexible optical networks and present their different views on how this emerging topic may evolve in the future. Most of these articles, following the recommendation of the Guest Editors, are co-authored by experts from different organizations in an effort to capture and reconcile some of the conflicting views and to agree on some common perspectives. The topics covered in the articles include transceiver technology for spatially and spectrally flexible elastic optical networking, switching schemes/elements for spatially and spectrally flexible elastic networking, novel transmission fibers for spatially and spectrally efficient transport, system transmission issues, as well as network planning and control plane issues for spatially and spectrally flexible elastic optical networking. In the following, we present the manuscripts we have accepted for inclusion in this Feature Topic and provide a brief summary of their content.

In the first article, "Introduction of Spectrally and Spatially Flexible Optical Networks," authors from Verizon (Tiejun J. Xia), Xtera (Herve Fevrier), NEC (Ting Wang), and DTU Fotonik (Toshio Morioka) discuss how spectral and spatial flexibility should be considered in an effort to allow network operators to meet the expected requirements for capacity growth while making data transport more cost effective. Their analysis starts with a discussion on the forecast traffic growth and sets capacity targets for future networks. Then they present different types of technologies that can be used to scale the capacity of systems and enhance the overall network efficiency in order to meet the identified network targets. In their opinion, the issue of capacity scaling is not as pressing a problem at the moment as is networking flexibility. Adding spectral and/or spatial flexibility in the optical network is considered an important step to improve network efficiency in the near term.

In the second article, "Migration from Fixed Grid to Flexible Grid in Optical Networks," authors from Beijing University of Posts and Telecommunications (Xiaosong Yu, Jie Zhang, and Yongli Zhao), Politecnico di Milano (Massimo Tornatore), Ericsson Research (Ming Xia), and the University of California, Davis (Biswanath Mukherjee) discuss how the evolution from conventional fixed-grid-based optical networks toward flexible-grid operation could be realized. In their study they present how the deployment of future flexible-grid technologies can follow brownfield policies that preserve the investment already made on existing fixed-grid WDM networks, while minimizing the interruptions to deployed services and maximizing the cost effectiveness of the migration process. Different migration options are proposed and examined via extensive network simulations revealing the pros and cons of each approach. Based on their simulation results, the authors find that the migration strategies should be carefully chosen by considering different traffic profiles and traffic distributions, and conclude that network operators can benefit from gradually upgrading the network to flexible grid, especially when the traffic is non-uniformly distributed in the network.

In the third article, "Transmission Media for an SDM-Based Optical Communication System," authors from NTT (Kazuhide Nakajima), Prysmian (Pierre Sillard), the University of Southampton (David Richardson), Corning (Ming-Jun Li), Alcatel-Lucent Bell Labs (René Essiambre), and Fujikura (Shoichiro Matsuo) discuss possible transmission media based on single- and multi-core optical fiber concepts and single- or multi-mode transmission per core that may be used in future SDM optical communication systems. The authors conclude that the final choice of the optimum combination of fiber core count and mode count will not be made by considering only the transmission performance in terms of supported capacity and reach, but will also depend on splicing performance, the cost of fabrication, as well as the ease of developing supporting subsystems (e.g., optical amplifiers or optical switches).

In the fourth article, "Physical Layer Transmission and Switching Solutions in Support of Spectrally and Spatially Flexible Optical Networks," authors from Alcatel-Lucent Bell Labs (Roland Ryf, S. Chandrasekhar, Sebastian Randel, David T. Neilson, and Nicolas K. Fontaine) and the College of Staten Island (Mark Feuer) present recent advancements in the area of transmission systems for spectrally and spatially flexible elastic optical networking. In particular, they discuss how new transponder technologies enable support for larger capacities and optimal use of the optical transmission spectrum, how novel optical switches will be able to achieve flexible routing of wavelength and spatially multiplexed channels, and how optical amplifiers can evolve to support the amplification of SDM channels. Following their investigations, they conclude that the concept of the "spatial super-channel," where multiple spatial channels are bundled and treated as a single entity, is expected to be a key driver in the optical communications industry. As a result, they see many opportunities for research toward new high-capacity low-cost-per-bit devices like multi-channel transponders, wavelength selective switches, and optical amplifiers for SDM-based networks.

In the fifth article, "Switching Solutions for WDM-SDM Optical Networks," authors from Hebrew University (Dan Marom and Miri Blau) outline possible optical switching operations that can occur at network nodes and discuss the associated systems for a variety of spatially flexible optical networks. Different switching schemes that can operate on individual wavelengths and SDM channels (over multiple cores or modes), or even be performed jointly over all cores/modes or all wavelengths are described and compared at a high level in terms of their physical realization requirements, as well as their routing flexibility and scaling potential.

In the sixth article, "Spectrally and Spatially Flexible Optical Network Planning and Operation," authors from the Athens Information Technology Center — AIT (Dimitrios Klonidis), CNIT (Filippo Cugini), Sedona Systems (Ori Gerstel), Kagawa University (Masahiko Jinno), Telefonica (Victor Lopez), Cisco (Eleni Palkopoulou), Fujitsu (Motoyoshi Sekiya), CREATE-NET (Domenico Siracusa), and Orange Labs (Gilles Thouénon and Christophe Betoule) provide an overview of the latest developments and approaches with respect to flexible optical networking with a special focus on network planning and resource optimization, as well as other important issues that need to be considered like IP/optical layer integration and control plane solutions. The different channel allocation options over the available fiber dimensions (wavelength, bandwidth, space) are discussed, and for each case the latest relevant technology advances and research efforts are presented in order to identify their benefits, limitations, and synergies.

Finally, in the seventh article, "3D Elastic Optical Networking in the Temporal, Spectral, and Spatial Domains," authors from the University of California, Davis (Roberto Proietti, Lei Liu, Ryan P. Scott, Binbin Guan, Chuan Qin, Tiehui Su, Francesco Giannone, and S. J. Ben Yoo) present technologies, architectures, and network planning algorithms for the emerging class of elastic optical networks (EONs) that could potentially support elasticity in the temporal, spectral, and spatial domains. In their studies, they introduce new routing, spectral, spatial mode, and modulation format assignment (RSSMA) algorithms that consider the use of the spatial domain for routing, and take into account constraints on the allocation and handling of spectral and spatial channel slots given the possible use of MIMO processing, which may be needed to combat in-band crosstalk in SDM systems.

We certainly hope that the articles of this Feature Topic in *IEEE Communications Magazine* will serve as a useful resource for researchers working in the field of optical communication systems and networks, while offering an interesting perspective of state of the art and future development of optical transport networks to all readers of this magazine.

BIOGRAPHIES

IOANNIS TOMKOS [M] (B.Sc., M.Sc. Ph.D.) (itom@ait.gr) is researching next generation optical communication systems and networks at AIT. Together with his colleagues and students he has authored over 550 scientific articles that have received several thousand citations. For his scientific achievements he has received several research grants and best paper awards (including the Best Paper in the *IEEE/OSA Journal of Lightwave Technology* in 2014). He was elected a Distinguished Lecturer of the IEEE Communications Society (2007), and a Fellow of the IET (2010) and OSA (2012).

YUTAKA MIYAMOTO [M'93] (miyamoto.yutaka@lab.ntt.co.jp) is the director of the Innovative Photonic Network Research Center (IPC) of NTT Network Innovation Laboratories, where he has been investigating and promoting the scalable OTN with petabit-per-second-class capacity based on innovative transport technologies such as digital signal processing, space-division multiplexing, and cutting-edge integrated devices for photonic preprocessing. He is a Senior Distinguished Researcher at NTT Laboratories and a Fellow of IEICE.

GLENN WELLBROCK (glenn.wellbrock@verizon.com) is the director of optical transport network architecture, design and planning at Verizon, where he is responsible for the development and deployment of new technologies for both the metro and long-haul transport infrastructure. Prior to this position, he ran the advanced technology lab, established evaluation criteria, and set engineering guidelines for all backbone transport equipment as well as various positions within network operations and worked at two different telecom equipment manufacturing companies.

PETER J. WINZER [F] (peter.winzer@bell-labs.com) heads the Optical Transmission Systems Research Department at Bell Labs, Alcatel-Lucent, New Jersey, working on high-speed and high-capacity fiber optic networks, including space-division multiplexing to scale optical transport systems. He is currently serving as Editor-in-Chief of the *IEEE/OSA Journal of Lightwave Technology* and as a Technical Program Chair of the 2015 Optical Fiber Communications Conference. He is a Bell Labs Fellow and a Fellow of the OSA.


free
Tutorials
Now™
ONLINE TUTORIAL
from IEEE Communications Society
www.comsoc.org/freetutorials
5G
TAKING 5G
FROM VISION TO REALITY
This tutorial, presented by Moray Rumney of Keysight Technologies, takes a broad look at the emergence of the fifth generation of mobile communications. How did we get here? What can we learn from previous generations? Who is setting the agenda and what might 5G requirements look like? From this foundation the tutorial will look at some of the new technologies being considered for 5G and the associated design and measurement challenges that follow.

Introduction of Spectrally and Spatially Flexible Optical Networks

Tiejun J. Xia, Herve Fevrier, Ting Wang, and Toshio Morioka

Tiejun J. Xia is with Verizon.

Herve Fevrier is with Xtera Communications, Inc.

Ting Wang is with NEC Laboratories America, Inc.

Toshio Morioka is with the Technical University of Denmark.

ABSTRACT

Given the introduction of coherent 100G systems has provided enough fiber capacity to meet data traffic growth in the near term, enhancing network efficiency will be service providers' high priority. Adding flexibility at the optical layer is a key step to increasing network efficiency, and both spectral and spatial functionality will be considered in next generation optical networks along with advanced network management to effectively harness the new capabilities.

INTRODUCTION

Network traffic will undoubtedly continue to grow in the foreseeable future due to the ever increasing demands of emerging applications, such as peer-to-peer video sharing, machine-to-machine communications, ultra-high definition video, gaming, mobile data, and Internet of Things (IoT). The overall global end-user IP traffic annual growth rate has been reported as 21 percent from 2013 through 2018 on average [1]. To handle this continuous growth in traffic demand, telecom carriers must increase capacity in their networks to support the demands from end users, devices, and applications in a scalable and cost-effective manner.

Fortunately, since commercial 100G systems with polarization multiplexed quadrature phase shift keying (PM-QPSK) modulation format and coherent detection was introduced into carriers' networks, the urgency to provide enough network capacity abated. It is estimated that the existing fiber infrastructure is able to support traffic demand growth for at least another 10 years without costly larger-scale fiber infrastructure upgrades [2]. Carriers now face a more serious challenge in how to make their networks more efficient to lower the overall cost of transporting bits. Next generation networks are expected to have better equipment utilization, customer service, and application performance. To reach this goal, adding more flexibility is a key step. Networks have had a lot of flexibility in the upper layers because they are not sensitive to physical distance and fiber impairments; however, the optical layers (layers 0 and 1) are. Here both spectral flexibility and spatial flexibility need to be considered. Consequently, network control platforms need to be improved to harness this flexibility if we are to realize highly efficient networks.

In this article we review aspects of spectral and spatial flexibility studied in recent years and show benefits of network flexibility to telecom carriers. Expected benefits include less capital expenditures (CAPEX), higher equipment utilization, faster service provision, better traffic protection, and easier network management. This article also shows several experiments supporting network flexibility. Finally, space-division multiplexing (SDM), as an additional way to provide spatial flexibility, is discussed from the angle of its potential benefit to optical networking in the future.

SPECTRAL FLEXIBILITY AND SPATIAL FLEXIBILITY

When we consider flexibility in networks, there are many flavors. In this article we focus on spectral flexibility and spatial flexibility in optical transport networks. Figure 1 shows some examples of network flexibility in the spectral and spatial domains, where the double arrow represents that one scenario can be adjusted into another scenario, controlled by network management systems.

The concept of flexible grid was introduced several years ago such that the optical bandwidth, and the central wavelength of a channel can vary depending on its data rate, symbol rate, channel arrangement in spectrum, and overhead of the channel (Fig 1a1). The modulation level of a channel can be controlled according to a balance between signal-to-noise ratio at the receiver and maximized spectral efficiency, as shown in Fig 1a2. Historically, most commercial optical channels have only had one optical carrier. However, commercial channels with data rates beyond 200 Gb/s, such as 400G or 1T channels as proposed in the industry, will likely be super-channels, which contain multiple optical carriers. Single-carrier solutions for the high-data-rate channels are still being investigated in research laboratories with challenges in data processing speeds and signal-to-noise ratio. Flex-

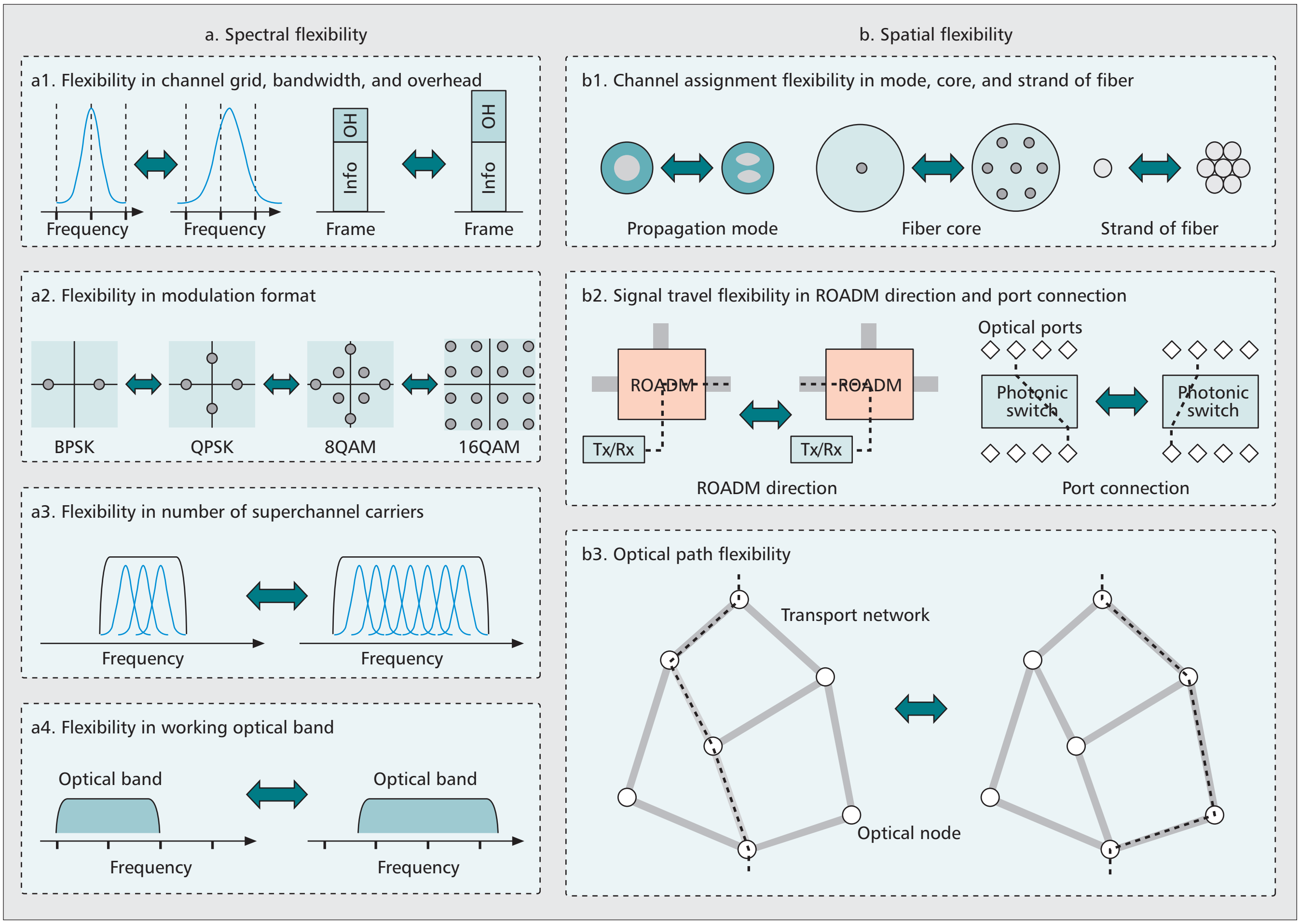


Figure 1. a) Scenarios showing spectral flexibility; b) spatial flexibility.

ibility in super-channel designs is reflected in the number of optical carriers and spacing between the carriers for a super-channel (Fig 1a3). Next generation optical transponders will have some or all of the attributes mentioned above. Using Raman amplification can increase network flexibility even further by allowing new optical channels to be assigned over wavelengths outside optical bands currently supported by Erbium-doped fiber amplifiers (EDFA). Raman amplification can allow new transmission bands to be created as required by traffic demand, as shown in Fig 1a4.

Spatial flexibility is referred to as the controllable arrangement of optical signals in the spatial domain. The introduction of SDM provides flexibility in the assignment of an optical channel to have spatial attributes, for example, different propagation modes, different fiber cores, or different strands of fiber in a fiber bundle, as shown in Fig 1b1. Next generation reconfigurable optical add/drop multiplexers (ROADMs) provide increased flexibility in optical path arrangement. A channel can be switched to different directions out of an optical node, controlled remotely by the network operation center (NOC) (Fig. 1b2). Spatial flexibility has also been introduced into port connections. Traditionally, port-to-port connections have been locked with fiber jumpers. Any change of port connections requires manual intervention. Adding photonic switches, which are wavelength- and data-rate-independent, will allow the NOC to be able to change equipment port connections remotely, as shown in Fig 1b2. At the network level, fast optical path reconfiguration is an ambitious but reachable target in the near future (Fig 1b3).

ROADM is a key element for optical networking. Colorless, directionless, contentionless, and flexible-grid ROADM (CDC-F ROADM) will play an important role in next generation flexible optical networks. ROADMs have been used in networks for more than a decade. However, only the CDC-F ROADM is a true flexible network enabler. Figure 2 shows a comparison of a basic ROADM design and a CDC-F ROADM design in a four-degree node as an example. A ROADM has an express core and add/drop modules for different degrees (directions). Bypassing traffic is switched to different directions in the express core, while added and dropped channels are switched to add/drop modules. In a basic ROADM each degree has its own add/drop modules, and each add/drop port has been assigned with a fixed wavelength. When a transponder (TRx) needs to change its wavelength or direction, it has to be physically moved to another port or another add/drop module, as shown in Fig. 2a. In a CDC-F ROADM, as shown in Fig 2b, a transponder does not need to move again once it is plugged into the node. The wavelength of the optical signal from the

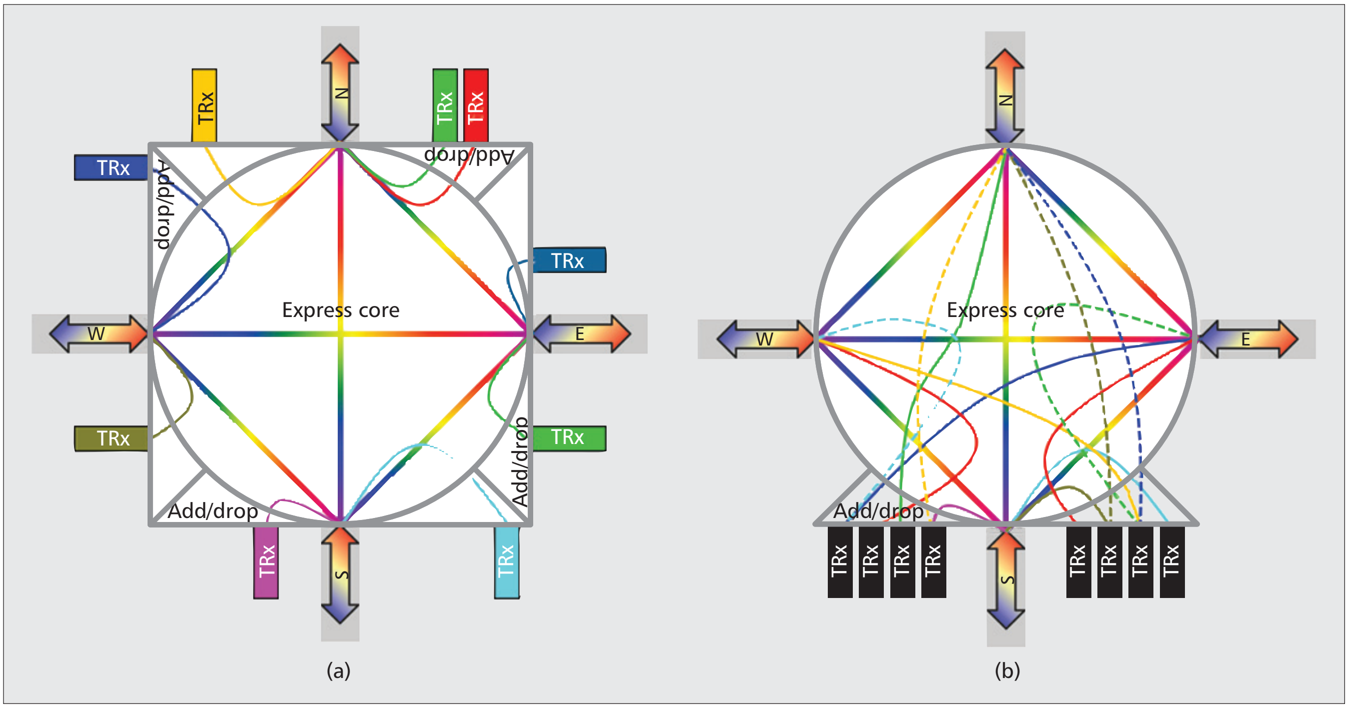


Figure 2. Schematic ROADM designs: a) basic design; b) colorless/directionless/contentionless and flexible grid (CDC-F) design for flexible optical networks (dashed lines represent potential connections).

transponder can be tuned and the direction of the signal can be changed without physically moving the hardware. In addition, the property of contentionless allows transponders sitting in the add/drop module to even have the same wavelength. Therefore, only one unified add/drop module is needed for a CDC-F ROADM. Using one unified add/drop module is particularly important since in this way the optical layer can be fully automatic and remotely managed by the NOC. Given that the flexible grid concept has been introduced into ROADM designs, it is also expected that the future CDC-F ROADM will be able to switch channels with variable optical bandwidths.

Equipped with flexible transponders and CDC-F ROADMs, a flexible optical transport node can also be constructed with electrical switch fabrics, as shown in Fig. 3. For global telecom carriers, the transport network must carry packet traffic from IP routers or multiprotocol label switched (MPLS) switches, time-division multiplexing (TDM) traffic from synchronous optical network/digital hierarchy (SONET/SDH) or optical transport network (OTN) switches, and packet or TDM traffic from customer routers or switches. Usually an optical node accepts the traffic via “grey optics,” which are short-reach pluggable devices. The traffic will be groomed and switched according to its destination and loaded into optical channels generated in transponders via switch fabrics. The switch fabrics are able to switch both packet and TDM traffic. As mentioned above, an optical channel may contain one optical carrier (single-carrier channel) or multiple optical carriers (super-channel). The channels leave the optical node in different directions via the CDC-F add/drop module and the express core. Ideally, any channel can be controlled by the NOC to adjust its traffic load, wavelength assignment, modulation format, number of optical carriers, overhead, direction, and optical path in the network. The flexibility opens up a whole new world of capabilities that promises to significantly improve network efficiency if managed properly.

With spectral and spatial flexibility it is critical to develop flexible and agile network control/management schemes to harness time-varying traffic flows in optical networks. In existing carriers’ networks, control systems are usually divided into domains, in which the equipment is usually supplied by different vendors and divided into layers. Network management across domains and layers often requires human intervention. The introduction of software-defined networking (SDN) into telecom networks is an effort to realize automation of network management and control. SDN has attracted a lot of attention recently in data center networks, enterprise networks, and academic networks. By decoupling the control plane from physical data forwarding devices, SDN is able to achieve simpler network management, faster traffic forwarding, and open development capability. The decoupling of the application layer, control layer, and infrastructure layer (Open Networking Foundation, ONF, terminology) is enabled by the northbound/southbound application programming interfaces (APIs) being standardized between different layers. There is growing interest in extending SDN to include optical transport layers [3]. Unified network controller software will control all network elements, including elements in the optical layer. Different from the upper layer control, the network controller software may not control optical elements directly, since physical layer perfor-

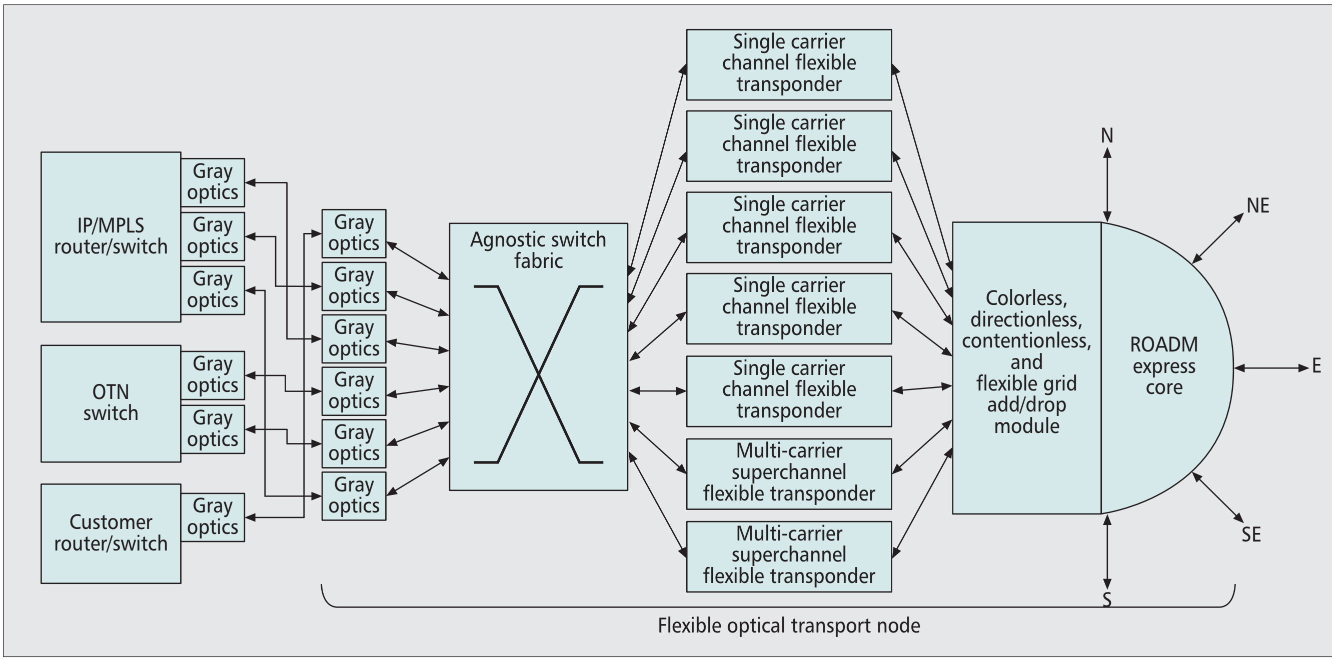


Figure 3. An example of a flexible optical transport node with agnostic switch fabrics, coherent channels, and CDC-F ROADM.

mance of optical signals is technology-dependent, and no standard in optical channel designs can be expected in the foreseeable future. A practical way to have unified control over optical elements is via control planes of optical domains. The control plane in each optical domain communicates with the unified controller software with a set of standard parameters and hides other attributes within the domain. With an SDN control/management platform, flexibilities in the optical layer are able to work with other layers in a coordinated manner.

Benefits of Network Flexibilities

Flexible optical networks can provide a lot of benefits to telecom carriers. Here we use a hypothetic optical network to show several such benefits. This hypothetic network has seven nodes, each equipped with a multi-degree CDC-F ROADM, transponders, switch fabrics, and grey optics, as shown in Fig. 4. An end-to-end data traffic connection is shown in the figure, from a router/switch in node A, via CDC-F ROADM A, D, F, and G, to another router/switch in node G. In the following several scenarios, we show how flexibility in the optical layer can help network operators.

Application Driving Network Reconfiguration

If a 400G super-channel going to node G originally via path A-D-F-G needs to be redirected to node B via path A-B according to the application riding on the channel, controller software can send commands to involved ROADMs to reconfigure the optical path of the channel. Assuming the 400G super-channel has four optical carriers, each of which has a PM-QPSK modulation format, to reach a long distance from node A to node G, after the path reconfiguration, the shorter distance between node A and node B allow the controller software to change the number of optical carriers to two and the modulation format to PM-16-quadrature amplitude modulation (QAM). The channel now still carries 400G data but uses only half the optical bandwidth. With unified controller software, optical path and channel bandwidth can be adjusted based on application needs.

Software Controllable Fast Network Restoration

Assuming there is a fiber cable cut between node A and node D, network controller software can immediately find an alternative optical path for the affected transponder pair. Since the optical signal of the channel is fully reconfigurable in wavelength and direction, the path of the channel can be configured to A-B-E-G. During path reconfiguration, the wavelength of the channel may be adjusted with a command from the controller software to fit into the spectral arrangement of the new path. If the path reconfiguration were fast enough, the router or switch would not even feel the cut. That is an important feature in an optical network to prevent unnecessary data rerouting or switching at higher layers. After the fiber cut is repaired, the controller software can shift the optical path of the channel back to the original one with minimal data traffic interruption.

Flexible Spare Cards Reduce Protection Costs

Since there is only one unified add/drop module in a CDC-F ROADM (logically), and all transponders are wavelength tunable, in principle, only one spare transponder card is needed to protect a failure in any transponder card of

In an elastic WDM network, a flexible transponder can dynamically adjust signal characteristics, such as data rate, modulation format, and error-correction coding scheme, for different channels in accordance with link conditions and quality of service requirements.

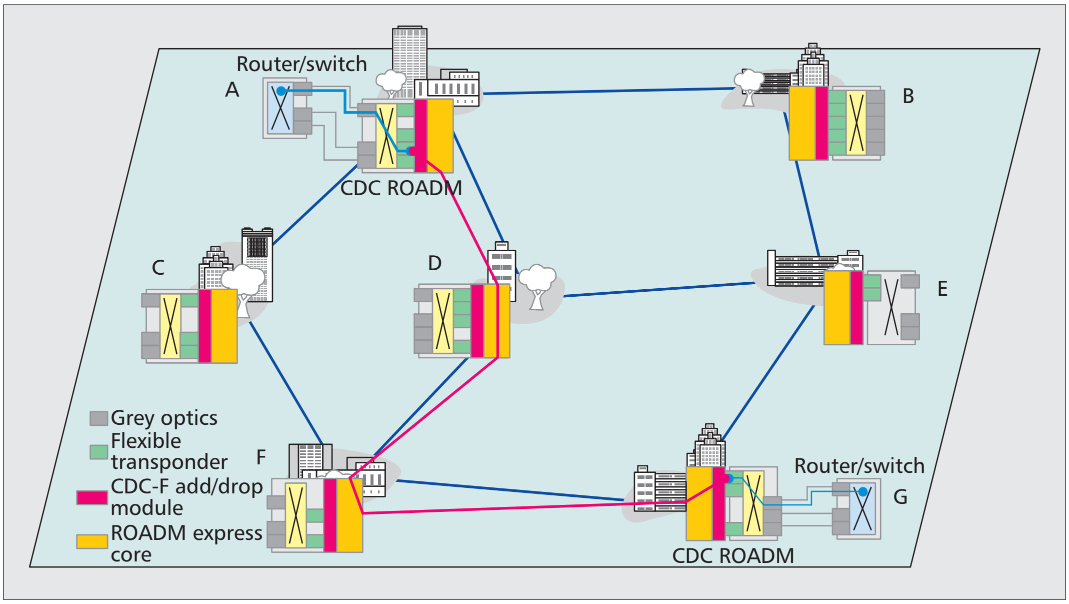


Figure 4. Hypothetic transport mesh network showing benefits of flexibility in the optical layer.

the same type in a CDC-F ROADM. For example, as shown in Fig. 4, assuming one of the transponder cards in node A is assigned as a floating card, if the network management system detects a failing working card in the ROADM, it can quickly issue commands to the node to shift the traffic flow from the failing card to the floating card via the switch fabrics, adjust the wavelength of the floating card to be the same as the failing card, and direct the signal from the floating card to the same destination as the failing card. Using floating cards to protect traffic is much more cost effective than traditional ways of traffic protection, but that can be realized only when the optical node has full flexibility.

Flexibility Helps Continuous Spectrum Arrangement Optimization

During installation of an optical channel it is very difficult, if not impossible, to assign the "right" wavelength to this channel by predicting what wavelengths other channels may take in future. After channels gradually populate a fiber network, it is almost inevitable that spectral fragmentation happens just like data fragmentation happens in hard disk drivers. With the freedom to tune wavelengths and optical paths, controller software in next generation networks will be able to defragment spectrum arrangement and free up spectrum for future use. Furthermore, controller software is able to perform such defragmentation periodically to continuously optimize the path and wavelength of optical channels.

Experiments Supporting Network Flexibility

Here we show a few examples of how spectral and spatial flexibility can be implemented in optical networks to help make optical networks more efficient.

Flexible Transponder Designs to Cope with Transmission Impairments

Of all elements able to present software controllable functions, transponders probably offer flexible functions the most since they serve as both source and destination of traffic flow. Therefore, technologies in transponder designs are key enablers for flexible optical networks. As mentioned above, by adjusting the modulation format, trade-offs can be achieved among system parameters such as spectral efficiency (SE), transmission distance, and system margin. Power consumption or capacity of transponders can also be adjusted through flexible forward error correction (FEC) designs with different coding gain options.

In an elastic WDM network, a flexible transponder can dynamically adjust signal characteristics, such as data rate, modulation format, and error correction coding scheme, for different channels in accordance with link conditions and quality of service requirements. This concept was first realized in a high-capacity field trial completed by Verizon and NEC that demonstrated 21.7 Tb/s fiber capacity over 1503 km [4]. Flexible modulation formats of 8QAM and QPSK were achieved by using a novel modulation unit. A total of 22 optical super-channels were transmitted over 19 dispersion-uncompensated field-installed fiber spans. A digital coherent receiver was used to receive the transmitted super-channels. For the first 18 super-channels the Q-factors of received signals (defined as digital SNR per symbol) are larger than 12.5 dB using PM-8QAM modulation. Their averaged spectral efficiency is 5.26 b/s/Hz. The last four super-channels on the short wavelength side have lower received optical signal-to-noise ratio (OSNR) due to lower gain and higher amplified spontaneous emission (ASE) noise and cannot support PM-8QAM transmission. With the flexible modulation unit, the transponder can still use the "inferior spectrum" by switching to

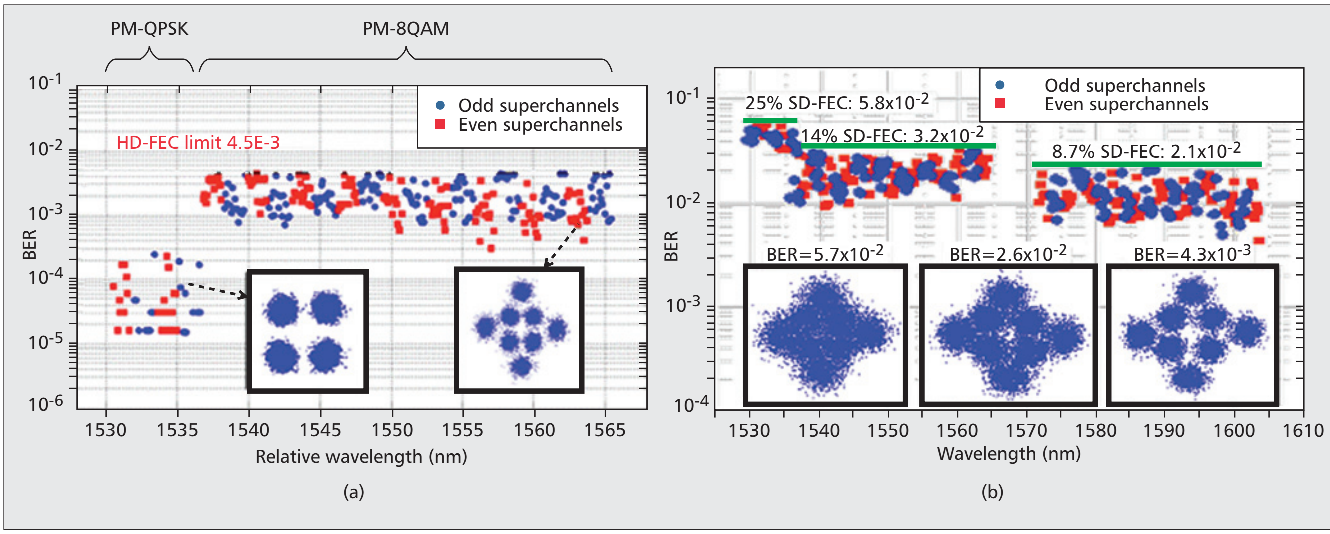


Figure 5. a) Field trial results of high capacity flexible transponders with adjustable modulation formats and b) adaptive FEC overheads.

QPSK and achieve error-free transmission with a reduced SE of 3.50 b/s/Hz. As a result, the total 4.4 THz WDM bandwidth was efficiently utilized with bit error ratio (BER) values of all optical subcarriers below the threshold of 4.5×10^{-3} using a product-ready CI-BCH FEC, as shown in Fig. 5a. Having the ability to administer different modulation formats, the transponders are able to use a larger portion of the spectrum with adjustable spectral efficiency, which is critical in > 20 Tb/s capacity at long-haul distances.

Other than changing modulation formats, adjustable FEC performance can also achieve trade-offs between spectral efficiency and transmission performance. In another record capacity field trial, an adaptive-rate low-density parity check (LDPC) code is employed to combat different received OSNR at different regions of C+L band WDM spectrum [5]. A total of 22 C-band super-channels and 19 L-band super-channels are combined to generate a total WDM bandwidth of ~9 THz. Each super-channel occupies 200 GHz and contains eight subcarriers with PM-8QAM modulation. The field trial uses 23 dispersion-uncompensated spans with a total distance of 1822 km and a hybrid amplification scheme including backward Raman pumping at 30 dBm averaged pump power with separate C- and L-band EDFAs. As a result of noise figure (NF) difference between C- and L-band EDFAs, unequal ASE noise distribution, and Raman energy shifting from C- to L-band, the 41 super-channels have as much as 2 dB difference in received OSNR. The adaptive LDPC coding with adaptive code rate and coding gain is designed to handle the non-uniform link condition. The pre-FEC BER results for all super-channel subcarriers are plotted in Fig. 5b, along with the BER threshold for the 25, 14, and 8.7 percent overhead used for the three spectral regions, respectively. In this trial, the adaptive FEC enables the transponder to maximize the data throughput based on the condition of its specific spectral region, as an example of many benefits in an elastic network. Using adaptive FEC, the field trial achieves highest long-haul field capacity of 40.5 Tb/s and field capacity-distance product of 73.7 Pb-km/s to date.

Flexible Optical Band Supported by All-Raman Amplification

Spectral flexibility also plays a role in selecting an optical band and the bandwidth of the band, as shown in Fig 1. In this area, Raman amplification shows more flexibility than EDFA amplification, since the spectrum of Raman amplification is quite flexible. All-Raman amplification technology, including discrete and distributed configurations, can also provide much broader spectrum compared to traditional EDFA. All-Raman systems with 100 nm spectrum were shown as commercial deployments as early as 2004. Distributed Raman amplification has been widely used in transmission systems to achieve better noise performance and reduction in channel power for mitigating nonlinear effects. In this section a field trial with flexible optical band provided by an all-distributed Raman system [6] is reviewed.

The field trial was carried out in a Verizon fiber network. The fiber plant is standard single-mode fiber with multiple splice points, which present discrete loss points as a result of construction activities in the metropolitan area. Each span is 79.2 km, and the average span loss is 21.8 dB. In the trial the span losses were mainly compensated by backward Raman pumping and occasionally helped by forward Raman pumping. The backward Raman pump module consists of five pump wavelengths from 1420 to 1500 nm and can deliver up to 1.9 W of pump power. The forward pump module includes three pump wavelengths (in the 1430 to 1480 nm range) and can deliver up to 0.85 W. It is a 150-channel 100G system, taking advantage of a 7.5 THz optical band supported by the all-distributed Raman system. Figure 6a shows the calculated power profiles for 150 channels along the transmission distance. Figures 6b and 6c depict input and output spectra of the transmission, respectively. Input channels are pre-emphasized

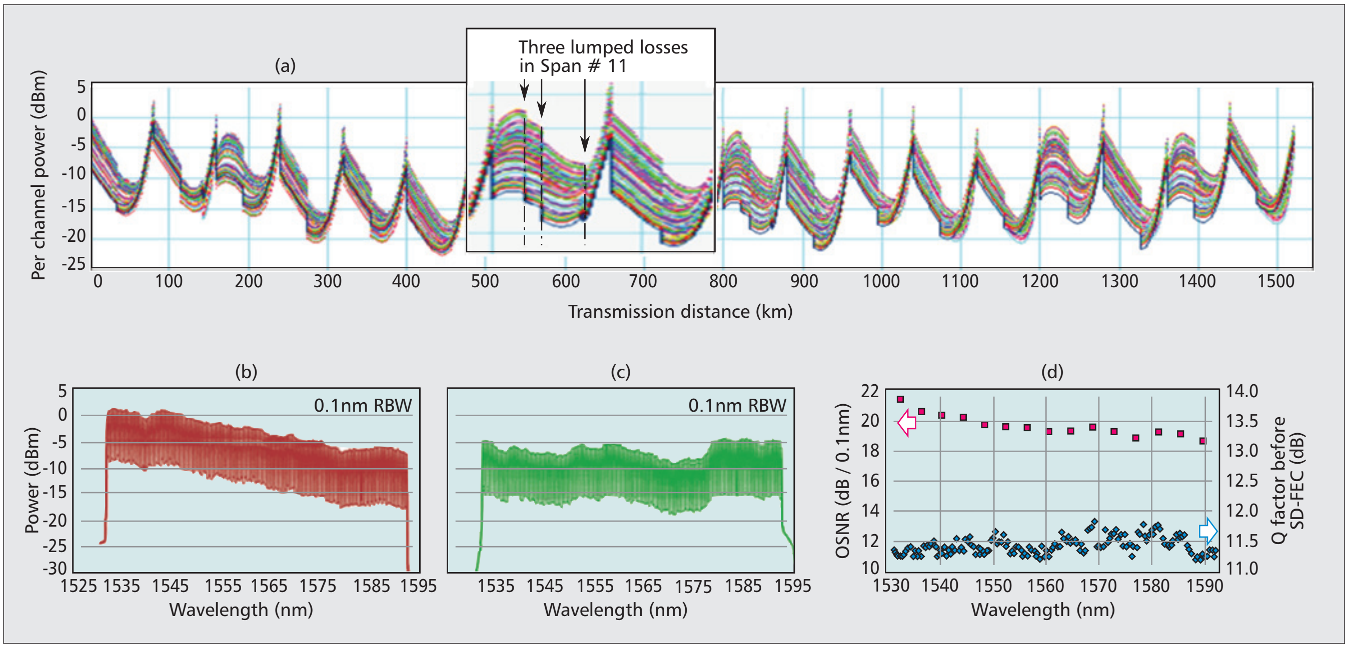


Figure 6. a) Results of the 150 × 100G all-Raman field trial, simulated per channel power profiles; b) spectrum at the booster output; c) spectrum after 1504 km transmission; d) measurements of OSNR and Q factor across the 150 × 100G channels spaced at 50 GHz.

to provide flat Q over spectrum at receive side. The gain ripple is smaller than 5 dB after the 19 spans across the 61 nm spectrum. Figure 6d represents the measurements of OSNR and Q factor performances across the 150 × 100G channels. The average OSNR is measured to be about 19.6 dB while Q factors of all channels are roughly uniform with an average value of 11.4 dB, which is 5 dB higher than the SD-FEC Q threshold of 6.4 dB.

The field trial demonstrated that properly designed Raman amplifiers can be deployed in fiber links with numerous optical connectors and discrete loss points. Given the nearly linear nature of the optical propagation (enabled by Raman distributed amplification), the high Q factor margin indicates that the 150 × 100G system is able to travel up to 4500 km in a real network environment. In the trial, channel spacing of 33.3 GHz was tested as well. Slight degradation in received signals due to the channel cross talk was observed, compared to 50 GHz channel spacing. However, the system still gives a 4 dBQ factor margin.

All-distributed Raman amplification also provides opportunity for PM-16QAM signals to be used for commercial long haul networks. PM-QPSK 100G channel is today's de facto standard long-haul transmission modulation format. PM-16QAM, on the other hand, thus far has been considered practical only for metro and regional networks due to higher OSNR requirements and sensitivities to nonlinearities along the line. The same setup as described above was used to demonstrate that PM-16QAM signals can be transmitted over long-haul distances in deployed networks with aged fibers [7]. In the trial, a total of eight dual-carrier 400G PM-16QAM channels, provided by NEC, were multiplexed in wavelengths with 134 100G PM-QPSK signals. All eight 400G PM-16QAM channels achieved transmission results above the SD-FEC Q-factor threshold of 4.95 dB, which corresponds to BER = 3.8×10^{-2}. With the above results we can estimate that 30 Tb/s capacity is feasible with 61 nm optical bandwidth for long-haul distances. Close to 50 Tb/s capacity is possible with usable optical band expanded to 100 nm with all-Raman amplification.

Unrepeatered subsea cable transmission represents another application of Raman amplification. Unrepeatered technology bridges long single-span submarine links connecting island to island or to mainland. Recent experiments show that Raman-powered remote optical amplifiers can support single span transmission with a full band of 100G channels up to 410 km with Corning's EX2000 ultra-low-loss fiber [8].

The optical band flexibility offered by Raman amplification provides a broader spectrum; therefore, larger transmission capacity and increased reach are possible. Components and subsystems for Raman systems, such as CDC ROADM, are available to provide optical networking functionality and wavelength tenability. The technical feasibility of similar flexibility for Raman systems with 100 nm gain spectrum has been demonstrated, and new deployments will happen again when they are required.

Software Defined Optics Experiments

As mentioned above, the emerging concept of SDN allows centralized control with a clear separation from the data plane. Several frameworks following this approach have been proposed, including, for example, the OpenFlow protocol developed by ONF and the Internet Engineering Task Force (IETF)-driven Network Configuration (NETCONF) protocol. Following the SDN paradigm, in principle all major hardware elements of an optical transport network can be controlled in a software-defined manner —

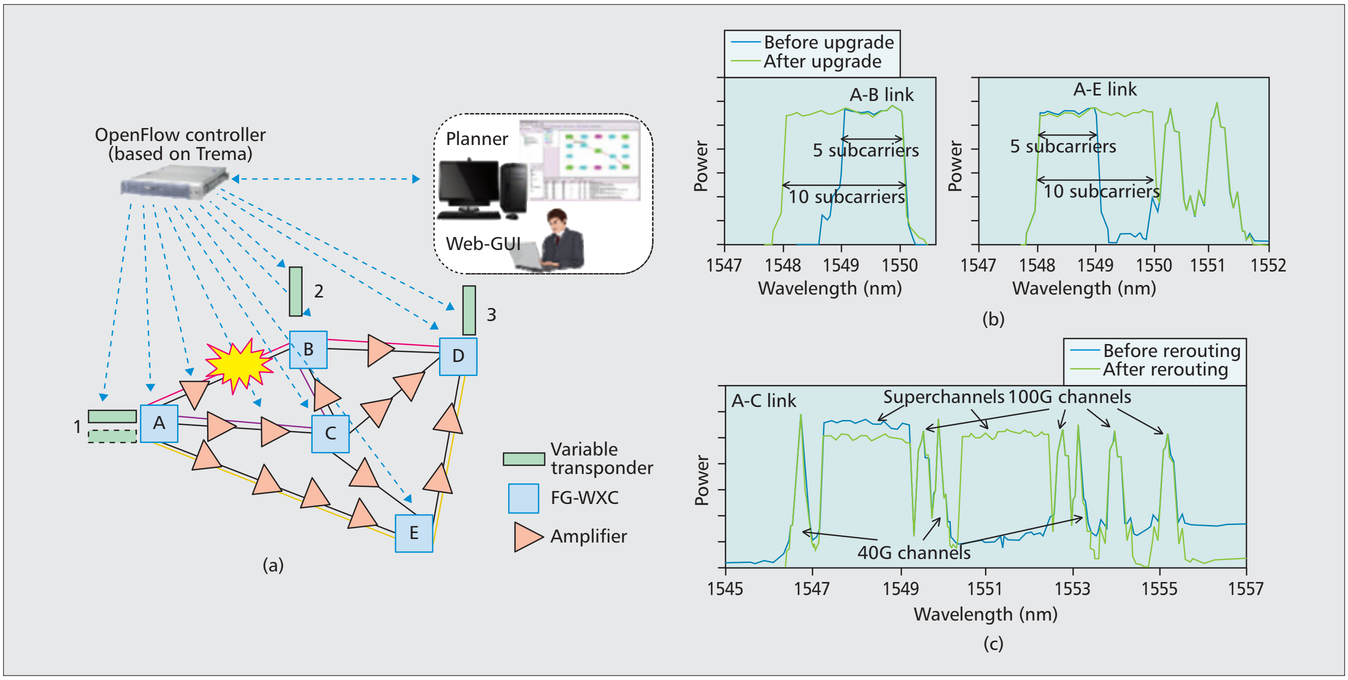


Figure 7. Demonstration of OpenFlow-enabled adaptive transport SDN.

including transponders, amplifiers, and switching nodes. Earlier, we presented how flexible trade-offs can be achieved by adjusting the modulation format or FEC of a transponder. In some cases, a transponder's power consumption can also be adjusted by flexible FEC design with different coding gain options [9]. By maintaining a global view of the optical network status and being able to interact with the network elements through a centralized controller using a standardized network protocol, operators can orchestrate the network much more efficiently. Although there is currently ongoing effort to allow SDN control of optical equipment using the aforementioned frameworks, the debate is still open regarding the extent of such control (i.e., the exact definition of abstractions that would still allow different vendors to employ and take advantage of their proprietary PHY technologies).

Use cases for SDN-controlled optical transport networks would include fast and automated optical path restoration, rearrangement, and spectral defragmentation, taking physical layer limitations into account. Moreover, the use of common SDN APIs for controlling equipment across different layers (e.g,. L2/L3 electrical packet switches, L1/L0 OTN switches, optical transponders, and ROADMs) would make it significantly easier for network operators to come up with cross-layer schemes for increasing network utilization and to intelligently implement other beneficial features such as IP traffic offloading through optical paths.

Recently, such synergy between diverse network elements of a flexible optical network under OpenFlow control was experimentally demonstrated [3] in order to showcase the benefits and feasibility of applying SDN to optical networking. As shown in Fig. 7a, the testbed consisted of five NEC SpectralWave DW7000 FlexGrid optical wavelength cross-connect (FG-WXC) nodes that use FlexGrid wavelength-selective switch (WSS) modules. Three bandwidth-variable transponders were employed at the optical network edges, each able to generate 10 optical subcarriers spaced at 25 GHz. Each subcarrier can be individually modulated with PM-QPSK or PM-16QAM signal with digital Nyquist shaping, delivering effective data rates of 180 and 90 Gb/s respectively. The adaptive optical amplifiers used were hybrid EDFA Raman amplifiers.

In this particular work all important network parameters like the number of subcarriers and modulation of bandwidth-variable super-channel transponders, the switching configuration of FG-WXC, as well as the gain of optical amplifiers were all controlled by a custom Trema-based OpenFlow controller through OpenFlow agents and extended OpenFlow messages. It was demonstrated that capacity upgrades via the use of an overlying network planner can easily be performed in a pay-as-you-grow fashion (Fig. 7b). It is noted that due to the software-defined nature of the system, bandwidth-on-demand services through automated processes could also be supported. In a second experiment (Fig. 7c), when link A—>B failed, the affected path was automatically rerouted through links A—>C and C—>B, while amplifiers' gain was remotely readjusted to account for the longer restoration path.

Discussion on Spatial Flexibility with SDM

For single-mode/single-core fiber, the total fiber capacity achievable has almost reached its limit. One approach to increase link capacity is to use spatially parallel transmission, or SDM. SDM provides a new approach to design channels and

Adding spectral and spatial flexibility has proven to be an important step toward improving network efficiency. Next generation optical network management platforms, such as extended SDN controller software, are a key element to harness the flexibility toward optimization of network resource utilization.

high-capacity links in a flexible way [10]. SDM can be broadly categorized into two categories:
- The parallel optical channels do not couple to each other; hence, existing transponders for single-mode fibers can be reused.
- The parallel channels couple to each other; thus, multiple-input multiple-output (MIMO) signal processing is required to untangle crosstalk.

SDM transmission using uncoupled multicore fibers (MCFs) has seen rapid progress since 2011. MCF can provide capacity more than 100 Tb/s relatively easily. So far the highest fiber capacity, 1 Pb/s, was reported with 12-core single-mode MCF [11] and 14-core hybrid-core MCF [12]. These experiments achieved spectral efficiencies of 91.4 and 109 b/s/Hz, respectively, in a single fiber. Multicore fiber design, fan-in fan-out devices, amplifier design, splicing techniques, mating sleeves, and cabling technologies are now sufficiently advanced. It is possible to combine few-mode-fiber (FMF) technology with uncoupled MCF to achieve even higher spatial multiplicity (SM). It has been reported that a 12-core MCF carrying three spatially non-degenerate modes per core (SM = 36) achieved a record spectral efficiency of 247.9 b/s/Hz [13]. If all the C- and L-band channels were fully utilized, capacity as high as 3 Pb/s may be achieved. Although the maximum number of cores achieved so far is 19, MCFs with > 30 cores can be well anticipated with the most advanced fiber designs. Since MCFs with uncoupled core behave like parallel fibers, they can enable the same networking functions performed on their spatial and wavelength channels as in current SMF-based systems.

The second type of SDM systems use coupled parallel channels, including coupled MCF and FMF. While these SDM fibers can achieve higher spectral efficiency per unit area than uncoupled MCF, they also face more engineering challenges. For example, mode coupling will generally require the use of MIMO processing at the receiver. To recover a particular spatial channel, the receiver must have access to the full field of all the parallel channels that couple into it. In terms of networking, we expect the use of spatial super-channels, with switching and networking functions restricted only to the wavelength dimension. Another challenge in strongly coupled parallel transmission is mode-dependent loss (MDL), which needs to be overcome to reach desired transmission performance. To support multi-mode transmission in FMF for long distance, few-mode EDFAs (FM-EDFAs) with low mode-dependent gain (MDG) are needed, along with a variety of spatial multiplexing solutions for FMF such as phase plates, spot couplers, or photonic lanterns. To date, long-haul transmission using FMF has reached 500 to 1500 km depending on the number of WDM channels.

In terms of network flexibility, the spatial dimension in SDM also allows flexible allocation of power and spectrum in the spatial domain. For example, channels with different spatial characteristics can support different constellation sizes. For single-mode cores, low core-to-core crosstalk (< –30 dB) may allow the use of high-level modulation formats for all cores. For few-mode fiber, mode-dependent loss may cause large variation in channel performance.

In addition to providing spatial flexibility, using SDM also achieves cost per bit reduction as total fiber capacity increases. It has been shown that cost reduction can be achieved in in-line components and at the transponders. Parallelization enables equipment overhead sharing, leading to better power efficiency, more efficient use of chip area, as well as reducing component counts and connector counts. For example, it is feasible to integrate parallel 100G transceivers in a single line card. Such a parallel transponder can be used to form spectral or spatial super-channels (i.e., parallel channels in frequency or space), leaving to higher-layer management how the super-channels may be transmitted flexibly. At this point, SDM fiber is likely to be deployed first in data centers with uncoupled MCF, where the benefit of increased spatial information density has already been demonstrated. There is still much debate on the optimal SDM solutions for access, metro, and long-haul networks.

Conclusions

While providing capacity in backbone and metro networks to meet end users' traffic demands seems not to be an issue to most telecom carriers, enhancing network efficiency is a very urgent task for the same carriers to ease cost reduction pressure and provide better services for new applications. Adding spectral and spatial flexibility has proven to be an important step toward improving network efficiency. Next generation optical network management platforms, such as extended SDN controller software, is a key element to harness the flexibility toward optimization of network resource utilization. Several experimental results, including SDO, are reviewed to show that an optical network is able to not only provide high capacity but also further increase network efficiency.

References

[1] Cisco, "Cisco Visual Networking Index: Forecast and Methodology, 2013–2018," white paper, June 10, 2014.

[2] G. A. Wellbrock and T. J. Xia, "How Will Optical Transport Deal With Future Network Traffic Growth?," *ECOC '14*, Th.1.2.1.

[3] P. Ji *et al.*, "Demonstration of OpenFlow-Enabled Traffic and Network Adaptive Transport SDN," *OFC '14*, W2A.20.

[4] Y.-K. Huang *et al.*, "High Capacity Fiber Field Trial Using Terabit/s All-Optical OFDM Superchannels with DPQPSK and DP-8QAM/DP-QPSK Modulation," *IEEE JLT*, vol. 31, no. 4, 2013, pp. 546–53.

[5] M.-F. Huang *et al.*, "Terabit/s Nyquist Superchannels in High Capacity Fiber Field Trials Using DP-16QAM and DP-8QAM Modulation for Regional and Long-Haul Applications," *IEEE JLT*, vol. 32, no. 4, 2014, pp. 776–82.

[6] D.-I. Chang *et al.*, "15T Unrepeatered Transmission over 409.6 km Using Distributed Raman Amplification and ROPA," ACP 2014, paper AF4B.4.

[7] T. J. Xia *et al.*, "Transmission of 400G PM-16QAM Channels over Long-Haul Distance with Commercial All-Distributed Raman Amplification System and Aged Standard SMF in Field," *OFC '14*, Tu2B.1.

[8] D.-I. Chang *et al.*, ""15T Unrepeatered Transmission over 409.6 km Using Distributed Raman Amplification and ROPA," ACP 2014, AF4B.4.

[9] S. Yamanaka *et al.*, "100 Gb/s CFP coherent transceiver enabled by power-optimized DSP," *OECC '14*, Sydney, Australia, July 2014, Paper THPDP1-4.

[10] T. Morioka, "Recent Progress in Space-Division Multiplexed Transmission Technologies," *OFC/NFOEC '13*, OW4F.2.

[11] H. Takara *et al.*, "1.01-Pb/s (12 SDM/222 WDM/456 Gb/s) Crosstalk-managed Transmission with 91.4-b/s/Hz Aggregate Spectral Efficiency," *ECOC '12*, Th.3.C.1.
[12] D. Qian *et al.*, "1.05 Pb/s Transmission with 109 b/s/Hz Spectral Efficiency Using Hybrid Single- and Few-Mode Cores," *Proc. Frontiers in Optics*, Rochester, NY, 2012, Paper FW6C.
[13] T. Mizano *et al.*, "12-core ? 3-Mode Dense Space Division Multiplexed Transmission voer 40 km Employing Multi-Carrier Signal with Parallel MIMO Equalization," *OFC '14*, Paper Th5B.2.

Biographies

TIEJUN J. (TJ) XIA is a key optics and photonics expert at Verizon, a global telecommunication service provider. He is currently a Distinguished Member of Technical Staff, and leads technology development for optical signal transmission and transport network architecture. He is also an adjunct professor at Miami University, Ohio, and was an adjunct professor at the University of Texas at Dallas from 2005 through 2009. In recent years he has led a series of high-speed, high-capacity, and flexible optical transmission studies in fiber networks, and set several field channel speed and transmission capacity records for the industry. He has served on conference Technical Program Committees (OFC/NFOEC, APOC, ACP, OECC, and LOIS). He is also a reviewer of technical publications of IEEE and OSA. He serves on the Executive Board of the IEEE Dallas Section and is active in promotion of innovation in the local community. He has a Ph.D. degree in physics from the Center for Research and Education in Optics and Lasers (CREOL) at the University of Central Florida, an M.S. degree from Zhejiang University, China, and a B.S. degree from the University of Science and Technology of China. He has written more than 20 invited papers, contributed more than 100 peer reviewed technical papers, published three book chapters, given more than 30 invited talks, and holds more than 80 granted or pending U.S. patents. He is a Fellow of OSA. He won technology Leadership Award of MCI in 1999 and Telecom Leaders Circle Award of Verizon in 2008. In 2011, he was featured as "Verizon Innovator" on YouTube. He is the winner of the CREOL Professional Achievement Award of UCF in 2013.

HERVE FEVRIER (herve.fevrier@xtera.com) holds an engineering degree in physics from the Ecole Centrale des Arts et Manufactures, Paris, France, and a doctorate in physics from the University of Paris XIII, France. He is currently the chief strategy officer of Xtera Communications. Prior to this position, he held different positions with Xtera Communications, among those VP of Photonics and more recently chief operating officer. He has led for the past 14 years the product development, product marketing, and sales efforts of the Xtera All-Raman technology applied to the different segments of the core networks: terrestrial long-haul networks, unrepeatered submarine cable systems, and submarine repeated cable systems.

TING WANG [M'88] (ting@nec-labs.com) serves as department head of Optical Networking at NEC Laboratories America Inc., Princeton, New Jersey. He joined NEC Research Institute in 1991, conducting research on optical interconnections and optical communication subsystems. He currently manages advanced R&D for next-generation optical transport technologies, including beyond 100 Gb/s transmission systems and transport SDN. He received his M.S. and Ph.D. degrees, both in electrical engineering, in 1991 and 1997. He has published more than 100 technical papers and holds over 90 U.S. patents.

TOSHIO MORIOKA (tomo@fotonik.dtu.dk) received M.Sc. and Ph.D. degrees from the University of Arizona in 1984 and from Waseda University, Japan in 1995, respectively. Since 1985, he has been engaged in the R&D of high-capacity optical communication systems at NTT Research Labs in Japan, and now he is a professor in high-speed optical communications at Technical University of Denmark (DTU), DTU Fotonik since 2011. He has been involved in leading pioneering research on novel all-optical signal processing technologies and ultra-high-capacity transmission technologies including new optical fibers for space-division multiplexing (SDM). He was also involved in the first Tb/s and Pb/s optical fiber transmission in 1996 and 2012, respectively. He is a Fellow of the Optical Society of America (OSA) and the Institute of Electronics, Information and Communication Engineers of Japan (IEICE). He received IEE Electronics Letters Premium Award in 1997 for his world's first Tb/s optical fiber transmission.

Migration from Fixed Grid to Flexible Grid in Optical Networks

Xiaosong Yu, Massimo Tornatore, Ming Xia, Jianping Wang, Jiawei Zhang, Yongli Zhao, Jie Zhang, and Biswanath Mukherjee

Xiaosong Yu is with Beijing University of Posts and Telecommunications and the University of California, Davis.

Massimo Tornatore is with Politecnico di Milano and the University of California, Davis.

Ming Xia is with Ericsson Research.

Jianping Wang is with the University of Science and Technology Beijing.

Jiawei Zhang, Yongli Zhao, and Jie Zhang are with Beijing University of Posts and Telecommunications.

Biswanath Mukherjee is with the University of California, Davis.

ABSTRACT

Optical WDM backbone networks based on fixed spectrum grid have limitations such as low spectrum utilization and rigidity in provisioning for heterogeneous rates. Flexible-grid technologies can alleviate these limitations for on-demand provisioning. These technologies represent promising candidates for future optical networks supporting beyond-100-Gb/s signals. However, a one-time green-field deployment of flexible-grid technologies may not be practical, as the already-made investment in existing fixed-grid WDM networks needs to be preserved, and interruptions to ongoing services need to be minimized. Therefore, we envision that fixed- and flexible-grid technologies will coexist, which will bring the challenge of interoperating fixed- and flexible-grid equipment. It is also important to design the optimum migration strategy to maximize cost effectiveness and minimize service interruption. In this article, we discuss the key aspects of network architectures supporting coexistence of fixed and flexible grid technologies, and outline the challenges of network operations. We also propose and evaluate different migration strategies from fixed grid to flexible grid under different network scenarios.

INTRODUCTION

The vision of 50 billion connected devices by 2020 is constantly pushing the traffic volume carried by our networks to new heights. Although wavelength-division multiplexing (WDM) technology has already provided high bandwidth using parallel wavelength channels, the overall network spectrum efficiency is severely discounted by the fixed grid definition and standard rate transmission (e.g., 10 Gb/s and 40 Gb/s). Recently, the concept of flexible grid has been introduced into optical transport networks [1–3]. The flexible grid technology evolves the traditional International Telecommunication Union (ITU) grid toward high flexibility with fine-grained spectrum slots (e.g., 12.5 GHz vs. 50 GHz or 100 GHz) [4]. Advanced optical transmission technologies, such as coherent optical orthogonal frequency-division multiplexing (OFDM) [5], Nyquist WDM (N-WDM) [6], and optical arbitrary waveform generation (OAWG) [7] are identified as the enabling technologies for flexible-grid optical networks. By using on-demand spectrum assignment and adaptive modulation formats, flexible grid can significantly improve the spectrum efficiency and increase the overall network capacity [8]. In addition, super channels (i.e., channels spanning multiple slots) can be set up to support high-bandwidth demands (e.g., 400 Gb/s and 1 Tb/s [2]).

In light of these advantages, a flexible-grid network is regarded as a promising candidate for future transport infrastructure. In [9], various migration options from fixed grid to flexible grid are investigated, and the impacts on flexibility and cost are discussed. In [10], the authors review the main drivers during the migration toward flexible grid, and introduce a planning tool to optimize the migration process. However, they leave an open but important question: how should the fixed-grid network be migrated toward the flexible grid? In other words, the problem of devising the most effective migration path toward flexible grid is still underinvestigated. Our recent work addresses the static routing and spectrum allocation (RSA) problem in fixed- and flexible-grid networks, and proposes several migration strategies with the goal of reducing the bandwidth blocking ratio of the network [11, 12]. This article investigates this problem in more detail (with a special focus on quantifying how much benefit we can get by gradually migrating from the current fixed grid to the flexible grid), and compares the performance of several migration strategies under different traffic profiles.

The remainder of this article is organized as follows. We discuss the network architecture with coexisting fixed-grid and flexible-grid nodes, and study the lightpath routing and spectrum allocation problem. We discuss how to perform effective and gradual migration from fixed grid to flexible grid. A case study is provided to compare the performance of the migration strategies. We then conclude the article.

 0163-6804/15/$25.00 © 2015 IEEE

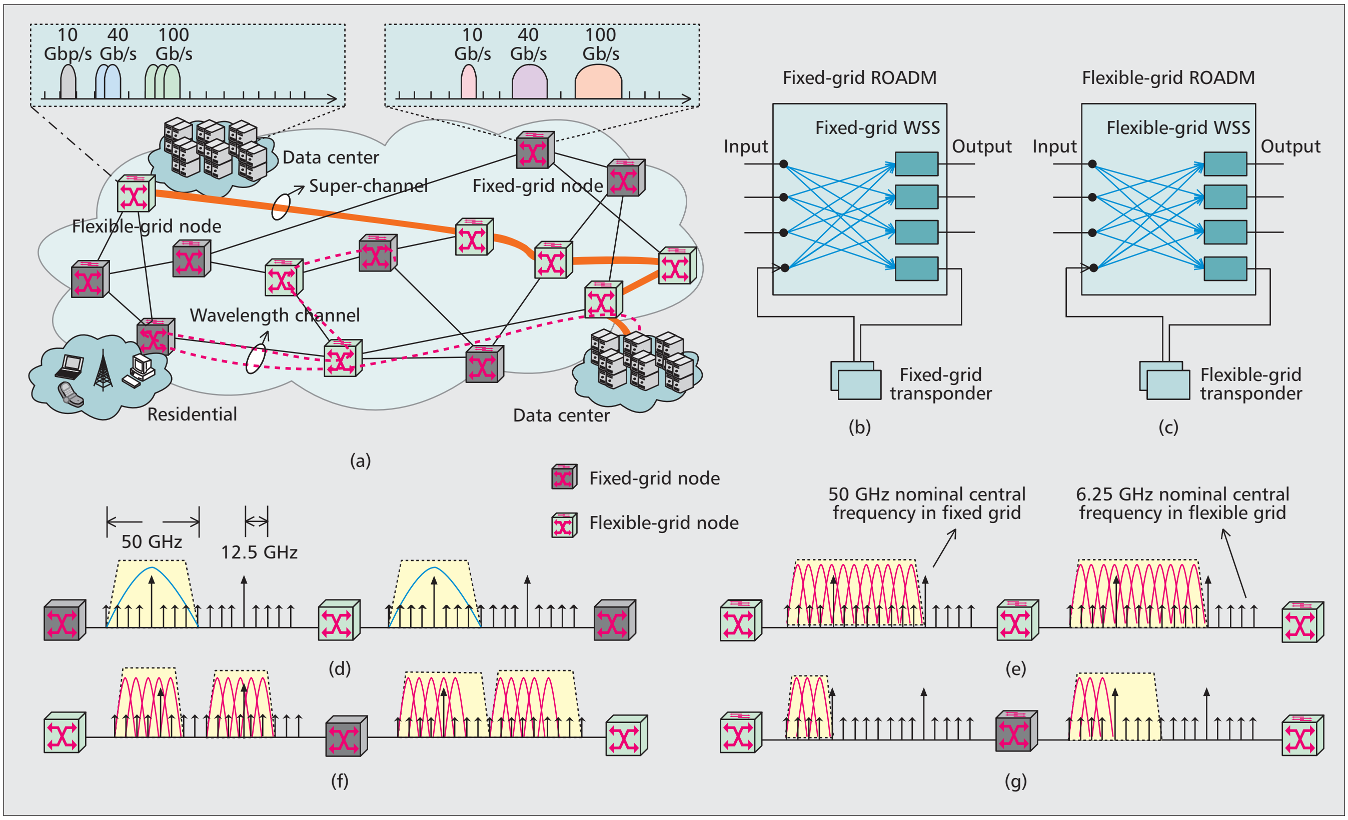


Figure 1. Optical network with co-existing fixed-grid and flexible-grid technologies: a) network architecture; b) fixed-grid ROADM; c) flexible-grid ROADM; d) wavelength channel; e) 200-Gb/s super-channel; f) two 100-Gb/s channels; g) 40-Gb/s subchannel.

Brown-Field Migration

Due to the increasing pressure on network operators to provide higher bandwidth with more efficient resource utilization, replacing the legacy fixed-grid equipment with flexible-grid equipment in their transport networks is just a matter of time. However, the operator's decision to migrate to flexible-grid technology will be influenced by key factors such as trade-off between benefit and equipment cost, compatibility with legacy systems, and complexity of network management. On one hand, the key enabling equipment (e.g., bandwidth variable wavelength-selective switches [BV-WSSs] supporting different grid definitions) has not yet reached a price point that allows massive deployment. It may not be economically viable to make a one-time complete upgrade to full flexible-grid technology for the entire network. On the other hand, before the current optical transport network capacity is exhausted, the current fixed-grid network could be kept maximally operational during the migration to preserve the already-made investment.

Network Architecture

In practice, a likely scenario is that traffic loads on some nodes/links are significantly higher than others, so they become bottlenecks [10]. For example, a common scenario concerns nodes associated with data centers, which tend to generate a large amount of traffic and can benefit from high-bandwidth super-channels interconnecting them. In these situations, the equipment causing the bottleneck could be replaced with flexible-grid equipment. As a result, brownfield flexible-grid deployment on top of the existing fixed-grid network could happen as shown in Fig. 1a.

While the sparse deployment of flexible-grid nodes can cost-effectively increase the capacity of only selected nodes/links, one challenge we will face is in the operational issues due to the coexistence of fixed-grid and flexible-grid technologies. Fixed- and flexible-grid nodes require different technologies. In particular, reconfigurable optical add/drop multiplexers (ROADM) are the key equipment to perform wavelength switching at intermediate nodes. Fixed-grid ROADMs follow the traditional rigid ITU-Telecommunication Standardization Sector (ITU-T)-defined central frequencies and spectrum grids (e.g., 50 or 100 GHz) regardless of the actual bit rate carried by each individual channel. Network devices (e.g., optical switches, multiplexers, and transponders) have to comply with this grid, as shown in Fig. 1b. The flexible-grid ROADM is different, as shown in Fig. 1c. Embedded wavelength-selective switches (WSSs) in flexible-grid ROADM do not need to strictly follow the ITU-T fixed grid, and can switch multiple concatenated slices as a single entity, where each slice may be 6.25 or 12.5 GHz.

These fixed-grid and flexible-grid nodes would need to interoperate before all nodes are upgraded to flexible grid. So a question is: how can newly added flexible-grid nodes be operated

Nodes generating the largest number of low-bandwidth traffic (e.g., 40 Gb/s) will be upgraded first. The intuition for this strategy lies in the fact that flexible-grid technology is spectrum-efficient for low-bandwidth traffic due to its on-demand spectrum provisioning instead of the rigid provisioning in fixed-grid technology.

in a network with other legacy fixed-grid nodes? Below, we discuss the relevant challenges in terms of lightpath routing, wavelength assignment, and spectrum allocation.

Interoperation between Fixed-Grid and Flexible-Grid Nodes

When a request arrives, we need to establish an optical path between its source and destination by determining a route through the network, and assigning/allocating a wavelength/frequency slot for this path. Here, a frequency slot is a spectrum allocation dedicated to a certain connection, and is specified by its nominal central frequency and slot width. Suppose a route is selected for a lightpath in an optical network with both fixed-grid and flexible-grid technologies, so there are several situations for wavelength assignment (WA)/spectrum allocation (SA):

- When the source is a fixed-grid node, we have the traditional WA problem. If the traffic demand is larger than 100 Gb/s, it can be served by several lightpaths, each accommodating 100 Gb/s or less (all following the same path, if possible).
- When the source is a flexible-grid node, there are two cases:
  - If the nodes on the path are flexible nodes, we have the SA problem, where a single-carrier channel or a super-channel with multiple subcarriers can be set up to accommodate the demands.
  - If there are both fixed- and flexible-grid nodes on the path, the spectrum is shared as common resources between fixed- and flexible-grid technologies, and the corresponding WA and SA problem becomes complex, different from WA in fixed grid and SA in flexible grid. On the path from the flexible-grid node to the fixed-grid node, we have the SA problem; but from the fixed-grid node to the flexible-grid node, we have the WA problem. If the traffic demand is larger than 100 Gb/s, we set up several lightpaths, each offering up to 100 Gb/s rate.

Figures 1d–1g illustrate four possible cases in networks with fixed-/flexible-grid coexistence. We consider the spectral granularity of fixed-grid nodes to be 50 GHz and that of flexible-grid nodes to be 12.5 GHz. Figure 1d shows the spectrum utilization of links for a 100-Gb/s lightpath that originates from a fixed-grid node and goes through a flexible-grid node. It occupies 50 GHz on both a fixed-grid link (i.e., a link originating from a fixed-grid node) and a flexible-grid link (i.e., a link originating from a flexible-grid node); Fig. 1e shows a 200-Gb/s lightpath that originates from a flexible-grid node and then goes through a flexible-grid node. Since we can set up a super-channel that comprises six 12.5 GHz slots, only 75 GHz of spectrum will be used instead of two 50 GHz channels in a fixed-grid network. However, when the path of a 200 Gb/s demand originates from a flexible-grid node but goes through a fixed-grid node, as shown in Fig. 1f, two lightpaths are set up, with each offering up to 100 Gb/s. Figure 1g shows a 40-Gb/s lightpath originating from a flexible-grid node and going through a fixed-grid node. Here, 25 GHz spectrum will be assigned to the optical path on the flexible-grid link, and 50 GHz will be assigned on the fixed-grid link, since the switching granularity of the fixed node cannot be smaller than 50 GHz.

Migration Strategies

As discussed in the previous section, migration to flexible-grid technologies may not be done at one time; instead, a network operator may choose to first upgrade network equipment where a bottleneck occurs. Around this scheme, various interesting questions arise as discussed below.

Question 1: Which node should be upgraded first?

When choosing a node (or nodes) to upgrade, many factors should be considered, such as network topology, traffic profile, network load, and network bottlenecks. Following are the strategies considered in our study, which are numerically evaluated in the next section.

Highest degree first (HDF): Nodes with the highest node degree will be chosen first to be upgraded. High node connectivity may have a positive impact on the upgrade performance, as a node with a higher degree connects to a larger number of other nodes in the network, thereby facilitating traffic provisioning options.

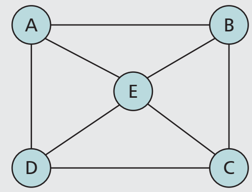
Highest generated traffic first (HGTF): Nodes that generate more traffic will be upgraded first so that more traffic might benefit from the upgrade.

Highest carried traffic first (HCTF): Nodes that carry the most traffic will be upgraded first. This is similar to the previous case, but it also considers transit traffic. Here, transit traffic includes all generated traffic as well as pass-through traffic.

Most high bandwidth traffic first (MHTF): Nodes generating the largest amount of high-bandwidth traffic (e.g., 400 Gb/s or 1 Tb/s) will be upgraded first. The argument for this strategy is that flexible-grid nodes enable super-channels for high-bandwidth requests, thus saving spectrum resources.

Most low bandwidth traffic first (MLTF): Nodes generating the largest number of low-bandwidth traffic (e.g., 40 Gb/s) will be upgraded first. The intuition for this strategy lies in the fact that flexible-grid technology is spectrum-efficient for low-bandwidth traffic due to its on-demand spectrum provisioning instead of rigid provisioning in fixed-grid technology. For example, a flexible-grid node uses only 25 GHz spectrum resources instead of 50 GHz to transmit a 40 Gb/s signal, which saves spectrum resources.

Figures 2a–2c illustrate the above migration strategies. Figure 2a shows a small five-node topology, with the traffic matrix shown in Fig. 2b. Total carried traffic by each node is shown in Fig. 2c. Thus, for HDF, node E will be upgraded first, since it has the largest node degree (i.e., 4); for HGTF, node D will be chosen, since it generates the highest traffic load (i.e., 500 Gb/s in total); for HCTF, node B will be chosen, since it



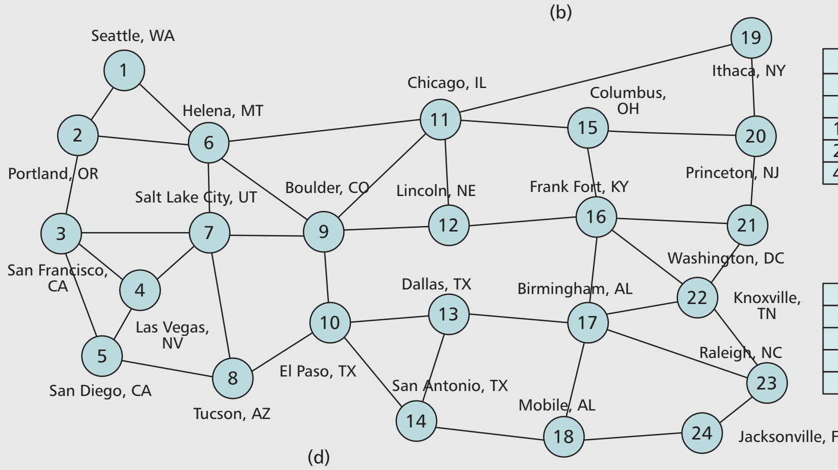
(a)

Source	Destination	Bit rate (Gbps)	Route
A	B	40	A-B
	E	40	A-E
	C	200	A-B-C
D	E	100	D-E
	B	400	D-C-B

(b)

Node	Carried traffic (Gbps)
A	280
B	640
C	600
D	500
E	140

(c)



(d)

	Fixed grid		Flexible grid	
Channel	Spectrum	#wavelength	Spectrum	# slots
40 Gbps	50 GHz	1	25 GHz	2
100 Gbps	50 GHz	1	37.5GHz	3
200 Gbps	100 GHz	2	75 GHz	6
400 Gbps	200 GHz	4	125 GHz	10

(e)

	Profile 1	Profile 2	Profile 3
40 Gbps	50%	20%	0%
100 Gbps	30%	50%	40%
200 Gbps	15%	20%	40%
400 Gbps	5%	10%	20%

(f)

Node	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Traffic distribution	63	60	83	60	134	3	19	52	10	67	272	27	124	138	81	3	21	20	3	3	63	18	42	84

(g)

Figure 2. a) 5-node topology; b) traffic matrix; c) carried traffic by each node; d) U.S. network topology; e) optical channels in fixed-grid and flexible-grid technologies; f) connection demand ratios in different traffic profiles; g) traffic distribution in the non-uniform case.

carries the highest traffic (i.e., 640 Gb/s); for MHTF, node D will be chosen, since it generates the largest number of high-bandwidth traffic demands (400 Gb/s); for MLTF, node A will be chosen, since it generates the largest number of low-bandwidth traffic demands (40 Gb/s).

Question 2: Should we create "flexible island(s)"?

If we simply follow the above policies, we would upgrade the nodes one by one, without considering the influence of the already upgraded nodes. For example, if we upgrade a node with a neighbor that is already a flexible-grid node, a high-bandwidth and spectrum-efficient super-channel can be set up between them. Thus, we can state that these two flexible-grid nodes have formed a "flexible-grid island." More rigorously, an island is a subset of network nodes where any two nodes of the subset can be connected to each other directly or through the node(s) that is (are) also in the same subset; a flexible-grid island means every node in this subset supports flexible-grid technology. In general, trying to form an island during a gradual migration process seems to be an effective way to maximize carried traffic, and some possible considerations can be drawn if we want to create flexible-grid islands.

Enlarging a single island. In this case, we could start by upgrading the first node according to, say, HCTF, and then choose as the second node the one with the highest carried traffic, but only among nodes adjacent to those already upgraded. This policy leads to the formation of an island that will keep growing during the migration until a complete migration is done.

Enlarging multiple islands. Since a traffic pattern in the network may have several centers (e.g., the east and west coasts of the United States may be observed with higher traffic volume than other places), a further improvement would be to have multiple islands growing independently. An idea is to choose nodes to be upgraded using metrics that can capture the *locality of traffic*.

Different migration strategies can also be devised if an operator decides to upgrade more than one node at a time. More optimized approaches could be explored to identify the most efficient migration strategy when the number of nodes upgraded at each step can be more than one.

Question 3: How many nodes should be upgraded?

While the ultimate goal is to migrate the entire network to support flexible-grid technology, upgrading only a subset of the nodes might be enough to remove current network bottlenecks. This may lead to different numbers of

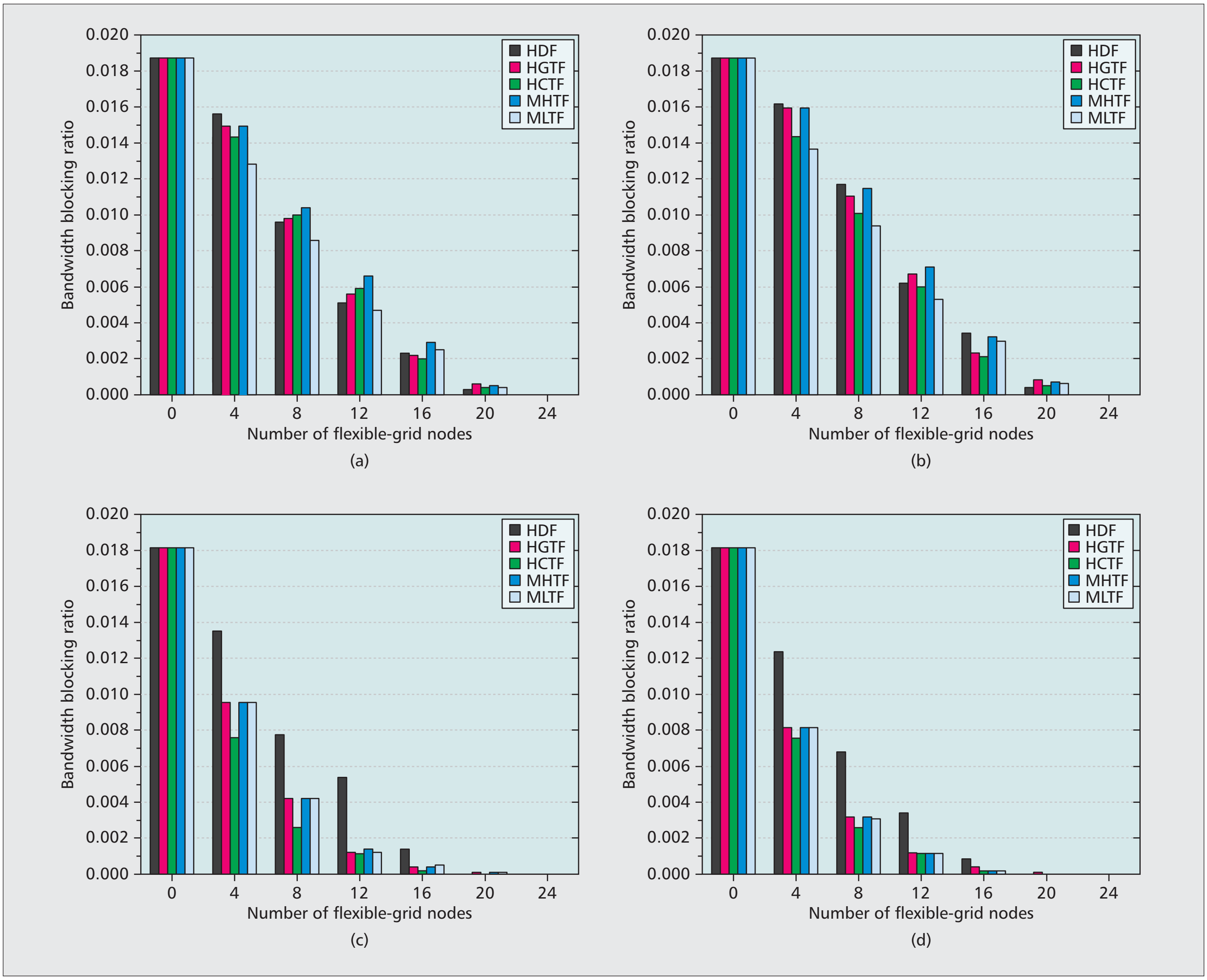


Figure 3. Bandwidth blocking ratio for traffic profile 1: a) forming one island under uniform traffic; b) forming two islands under uniform traffic; c) forming one island under non-uniform traffic; d) forming two islands under non-uniform traffic.

nodes to be upgraded under a given scenario with a predetermined objective (e.g., to lower bandwidth blocking probability under some predefined target). Also, for a network planning scenario, the number of upgraded nodes has an important effect on the node selection decision. For example, in Fig. 2a, when only two nodes are upgraded, nodes A and B may be selected; it may be the case that nodes A, C, and E need to be upgraded when the quota is three.

Case Study

We compare the migration strategies described earlier on a 24-node U.S.-wide network shown in Fig. 2d. Each link is bidirectional with 4 THz spectrum in each direction. For the fixed-grid technology, we consider a 50 GHz frequency grid, so each link has 80 wavelengths; for flexible-grid technology, the frequency grid is 12.5 GHz, so each link has 320 frequency slots. We generate 500,000 any-pair connection demands following Poisson arrival, and their bandwidth requirements are uniformly chosen among [40, 100, 200, 400] Gb/s. We use Fig. 2e to map a bandwidth demand to a spectrum allocation using fixed-grid and flexible-grid technologies, respectively. We use dual-polarization quadrature phase-shift keying (DP-QPSK) for 40 and 100 Gb/s rates in both the fixed-grid and flexible-grid scenarios, while we use orthogonal frequency-division multiplexing (OFDM) version of DP-QPSK for 200 and 400 Gb/s in the flexible-grid scenario. Here, we suppose the guard band is included in the required spectrum, and the maximum optical reaches for DP-QPSK and OFDM-DP-QPSK are 2800 km and 3500 km [13], respectively. Connection requests are handled sequentially; and for each connection, k-shortest path routing (with $k = 5$) and first-fit spectrum assignment are used. For a flexible-grid lightpath, if there are not enough spectrum resources to set up a super-channel, we also try to split the high-bandwidth request into small lightpaths [14]. For example, one 400 Gb/s connection can be split into two 200 Gb/s channels or four 100 Gb/s channels.

Since traffic may influence the migration strategies, we consider three traffic profiles as shown in Fig. 2f. For example, in *profile 1*, the

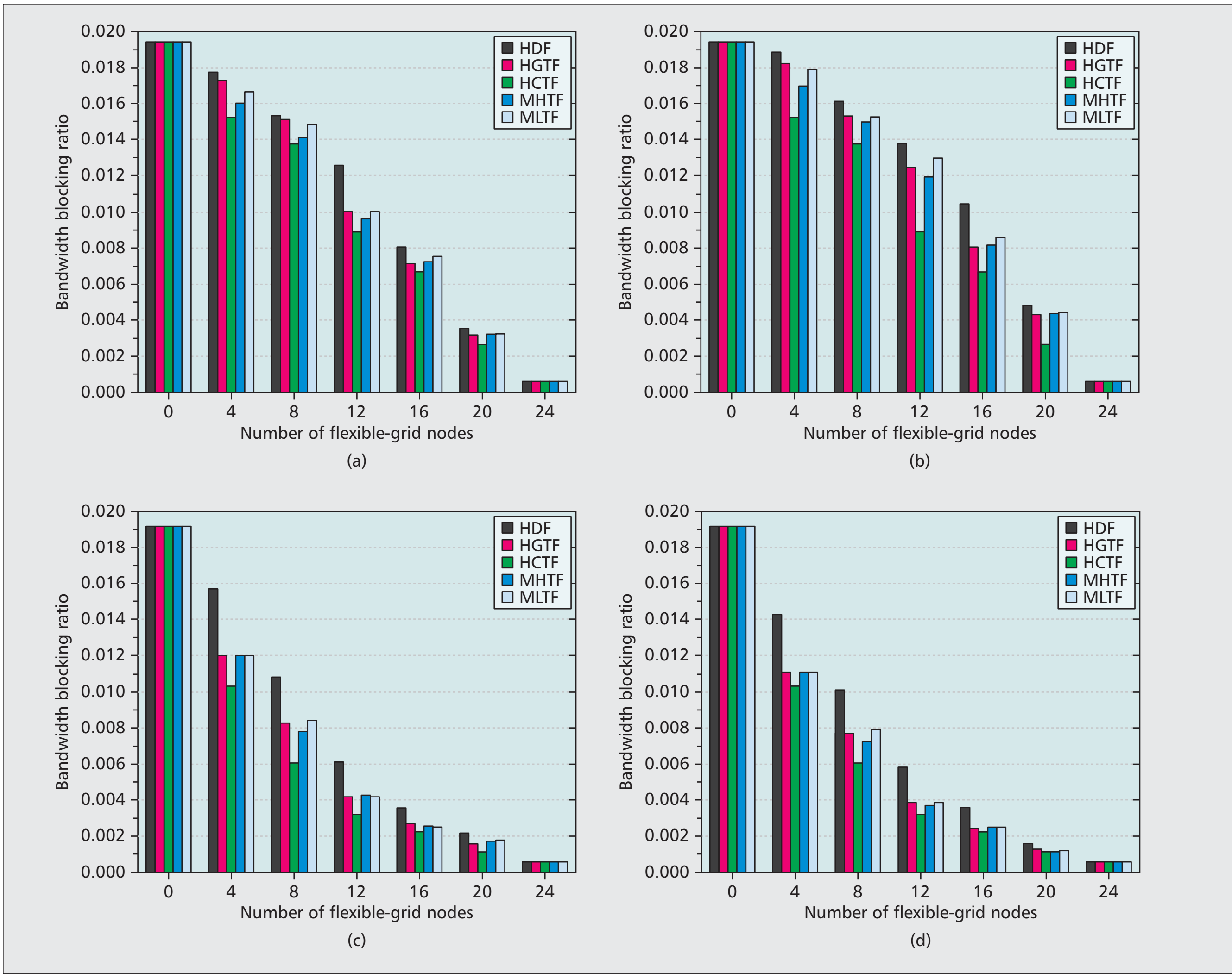


Figure 4. Bandwidth blocking ratio for traffic profile 2: a) forming one island under uniform traffic; b) forming two islands under uniform traffic; c) forming one island under non-uniform traffic; and d) forming two islands under non-uniform traffic.

ratios of 40, 100, 200, and 400 Gb/s are 50, 30, 15, and 5 percent, respectively (i.e., low-bandwidth traffic is predominant); while in *profile 2*, 100 Gb/s traffic is predominant, with only 20 percent 40 Gb/s traffic; in *profile 3*, all the traffic are 100G and beyond, with 400 Gb/s traffic as high as 40 percent. Please refer to Fig. 2f for the values of these ratios in all traffic profiles. The migration strategies are applied on all three profiles. For each profile, traffic can be either uniformly or non-uniformly distributed among all the nodes. For the second case, our study assumes that traffic is distributed according to the population of the city at the corresponding node as shown in Fig. 2g. Strategies' performances in terms of bandwidth blocking ratio (BBR), that is, the rejected bandwidth over the total bandwidth, are compared. The reason for using BBR is that for a network with non-uniform bandwidth requests, BBR is a weighted metric showing the capability of the network to admit traffic volume rather than only number of requests.

For *traffic profile 1*, the BBR for the different migration strategies is shown in Figs. 3a and 3b for uniformly distributed traffic, and Figs. 3c and 3d for non-uniformly distributed traffic. The traffic load is set to 900 Erlang. From Fig. 3a we see that as the number of upgraded nodes increases, the BBR of all the migration strategies decreases. When we upgrade fewer than 12 nodes, MLTF gives the best performance whether we form one or two islands. The reason is that in profile 1, the percentage of 40 Gb/s traffic is 50 percent, and if we choose the node that has the most 40 Gb/s demands to upgrade, more benefit can be achieved. However, if we upgrade more than 12 nodes, HCTF gives better performance. If we compare the performance of forming one island in Fig. 3a and two islands in Fig. 3b, we find that the BBR of HCTF is always the same. The reason is that, using HCTF, the updating sequence of the nodes for the one-island or two-island cases is the same. For other strategies, forming one island is better than forming two islands. This is due to the fact that in uniform traffic, there is not much difference in traffic intensity in the network, so forming one island brings more benefit than forming two islands. For non-uniform traffic, BBR drops more rapidly, as shown in Figs. 3c and 3d. For example, when we upgrade four nodes (forming

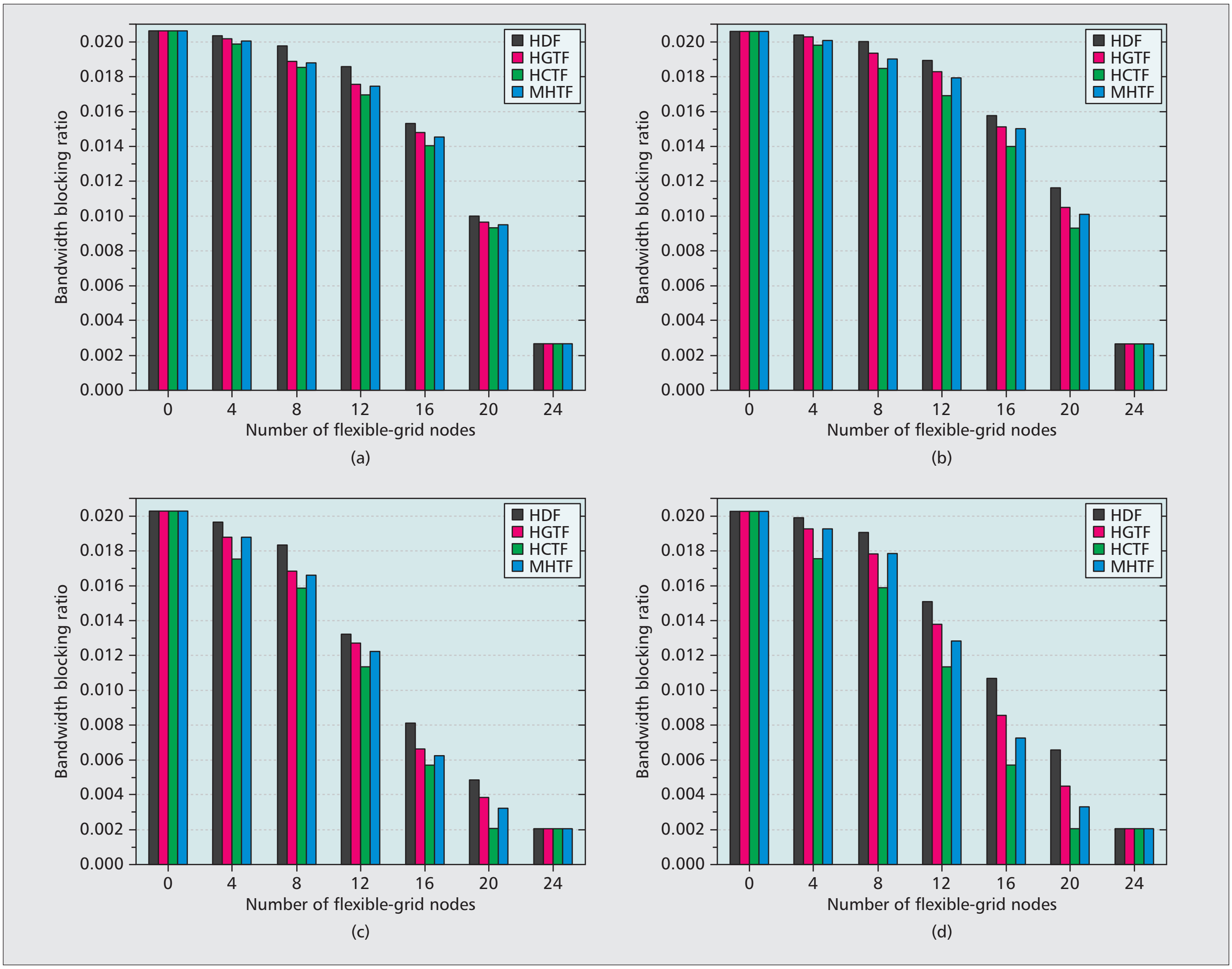


Figure 5. Bandwidth blocking ratio for traffic profile 3: a) forming one island under uniform traffic; b) forming two islands under uniform traffic; c) forming one island under non-uniform traffic; and d) forming two islands under non-uniform traffic.

one island) under uniform traffic distribution by HCTF, the BBR is reduced by only 17.6 percent; while under non-uniform traffic, it can be reduced by up to 55.6 percent. The reason is that in the non-uniform case, traffic is unevenly distributed, and by choosing a small number of nodes with the most traffic, the BBR can be reduced quickly. In addition, HCTF always gives the best performance whether forming one or two islands. Different from uniform traffic, BBR is lower if we form two islands for other migration strategies, as shown in Fig. 3d.

Figures 4a–4d show the BBR of different migration strategies for *traffic profile 2*. Figures 4a and 4b are for uniform traffic, and Figs. 4c and 4d are for non-uniform traffic. In order to have similar BBR values to profile 1, the traffic load is decreased to 760 Erlang. From Fig. 4a, we can see that HCTF gives the lowest BBR no matter how many nodes are upgraded. By comparing Figs. 4a and 4b, we find that the performance of HCTF is the same with one or two islands. The reason is that when using HCTF, the updating sequence of the nodes for the one-island and two-island cases is still the same. For other strategies, forming one island is more efficient than forming two islands, confirming that for uniform traffic, the one-island strategy is better. For non-uniform traffic, as seen before, the BBR again decreases more rapidly than in the uniform case. For example, by upgrading only four nodes, the BBR of HCTF can be reduced by up to 47 percent under non-uniform traffic instead of only 21 percent in the uniform case. In addition, in the non-uniform case, HCTF again gives the same performance when forming one or two islands, but now for other strategies, forming two islands gives a lower BBR compared to forming only one island, which confirms the trend in Figs. 3c and 3d.

Figures 5a–5d show the BBR of different migration strategies for *traffic profile 3*. Figures 5a and 5b are for uniform traffic, and Figs. 5c and 5d are for non-uniform traffic. Also, in this case, to have similar BBR values to profiles 1 and 2, we decrease the traffic load to 580 Erlang. From Figs. 5a and 5b, we see that under uniform traffic, all migration policies achieve comparable performance, with HCTF a little better. For non-uniform traffic in Figs. 5c and 5d, we see that if we upgrade only a small portion of nodes to flexible grid, less benefit is obtained com-

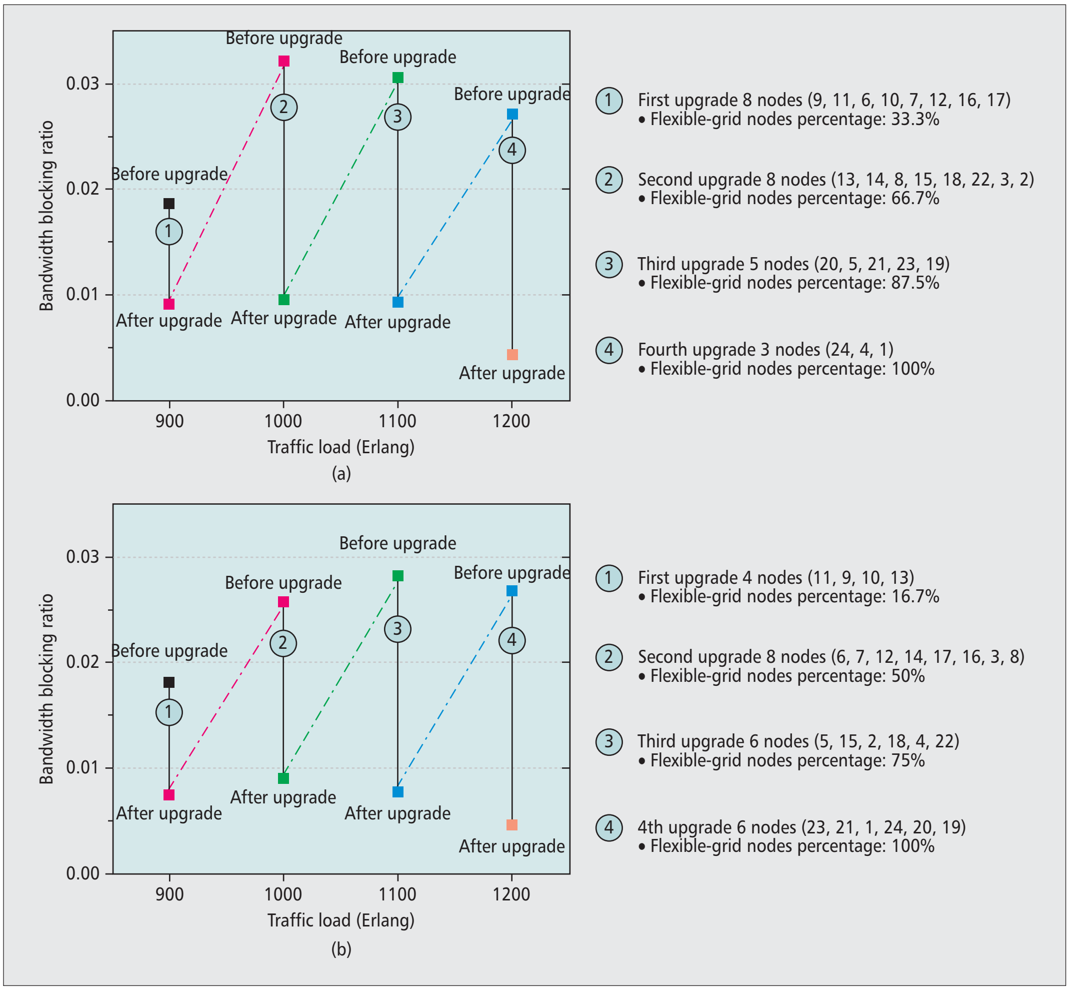


Figure 6. Network's gradual migration to flexible-grid technology: a) uniform traffic; b) non-uniform traffic.

A general conclusion is that migrating to flexible-grid technology can improve network capacity and lead to lower BBR. This benefit can be achieved even if we only upgrade a small portion of the network to flexible grid, especially for non-uniform traffic.

pared to profiles 1 and 2. For example, if four nodes are upgraded (i.e., 17 percent of the total nodes), BBR is reduced by only 10 percent. Only if we upgrade to 12 nodes (i.e., 50 percent of total nodes) can we reduce the BBR by up to 45 percent. The reason is that in profile 3, the percentage of 200 Gb/s and 400 Gb/s traffic (i.e., requests served by super-channels) is much higher than in profiles 1 and 2, so the network needs a much larger number of flexible nodes to avoid blocking these super-channels, and cannot quickly benefit by upgrading only a small proportion of nodes. By comparing Figs. 5c and 5d, we see that forming one island is better than forming two islands, which is different from profiles 1 and 2.

Finally, to investigate the benefits of gradual migration, we consider scenarios with increasing traffic volume. As an example, we consider traffic profile 1 and the HCTF migration strategy, and the results are shown in Figs. 6a and 6b. Our objective is to keep BBR less than a threshold (1 percent in this study). First, we consider uniform traffic. In our initial setting, traffic load is 900 Erlang and the BBR of the network is 1.87 percent, which is higher than 1 percent. This means the network should be upgraded. We start to upgrade the nodes one by one. After upgrading eight nodes (i.e., nodes 9, 11, 6, 10, 7, 16, 12, and 17), BBR is reduced to 0.91 percent. Since the BBR is now lower than the 1 percent performance target, we stop the upgrade process. But as traffic load increases to 1000 Erlang, the BBR increases to 3.23 percent, which means another upgrade process is needed. An additional eight nodes (i.e., nodes 13, 14, 8, 15, 18, 22, 3, and 2) are selected to be upgraded such that the BBR decreases to 0.97 percent. As traffic load increases to 1100 Erlang, BBR increases to 3.05 percent, and triggers another upgrade operation. Now, we upgrade another five nodes (i.e., nodes 20, 5, 21, 23, and 19) and BBR drops to 0.92 percent. When the traffic loads increases to 1200 Erlang, the BBR increases to 2.7 percent, triggering another upgrade operation. After upgrading the rest of the nodes in the network (i.e., nodes 24, 4, and 1), the BBR drops to 0.43 percent, which meets our objective again. When traffic increases to 1300 Erlang, the BBR reaches 1.76 percent, and we would need additional nodes/links or other traffic/network engineering (TE/NE) approaches to satisfy the 1 percent BBR objective since all the nodes in the network have already been upgraded. Note that for non-uniform traffic (Fig. 6b), considering the same settings, the sets of nodes upgraded for the different traffic loads are {11, 9, 10, 13}; {6, 7, 12, 14, 17, 16, 3, 8}; {5, 15, 2, 18, 4, 22}; and {23, We found that the migration strategies should be carefully chosen by considering traffic profiles and traffic distributions. Network operators can benefit from gradually upgrading their network to flexible grid, especially when traffic is non-uniformly distributed in the network.

21, 1, 24, 20, 19}, respectively. Also, when traffic load is 1300 Erlang, the network needs to deploy additional resources or use other TE/NE strategies to achieve the BBR objective. This example indicates that the network can benefit from gradual migration to flexible grid, especially under non-uniform traffic.

From the case study, we see that the performance of a migration strategy depends partly on the traffic profile. For example, if low-bandwidth demands (e.g., 40 Gb/s) are dominant, MLTF may give better performance if only a few nodes are upgraded; but if high-bandwidth demands (e.g., 400 Gb/s) are dominant, HCTF may perform better no matter how many nodes are upgraded. Also, traffic distribution has an important effect on the performance of a migration strategy. For example, under uniform traffic, forming one flexible-grid island gives more benefit; however, under non-uniform traffic, forming more than one island is a better choice. A general conclusion is that migrating to flexible-grid technology can improve network capacity and lead to lower BBR. This benefit can be achieved even if we only upgrade a small portion of the network to flexible grid, especially for non-uniform traffic.

Conclusion

Compared to fixed-grid technology, flexible-grid technology has many advantages such as higher capacity, more flexibility, and better spectrum efficiency, which make it a promising candidate for future optical transport networks. This article investigates various strategies for migrating from fixed grid to flexible grid, and studies the problem of interoperation between fixed-grid and flexible-grid technologies. Migration strategies from fixed grid to flexible grid are discussed, with their performance compared in a number of case studies. From the results, we find that the migration strategies should be carefully chosen by considering traffic profiles and traffic distributions. Network operators can benefit from gradually upgrading their network to flexible grid, especially when traffic is non-uniformly distributed in the network.

Acknowledgment

Preliminary and summarized versions of this work appeared in the proceedings of the OFC 2014 and ECOC 2014 conferences. This work has been supported by China's 863 program (2012AA011301), NSFC project (61201154, 60932004, and 61201260), Ministry of Education-China Mobile Research Foundation (MCM20130132), and by the Networks Lab at the University of California, Davis in the United States.

References

[1] M. Jinno *et al.*, "Spectrum-Efficient and Scalable Elastic Optical Path Network: Architecture, Benefits, and Enabling Technologies," *IEEE Commun. Mag.*, vol. 47, no. 11, Nov. 2009, pp. 66–73.

[2] O. Gerstel *et al.*, "Elastic optical Networking: A New Dawn for the Optical Layer?," *IEEE Commun. Mag.*, vol. 50, no. 2, Feb. 2012, pp. S12–S20.

[3] I. Tomkos *et al.*, "A Tutorial on the Flexible Optical Networking Paradigm: State of the Art, Trends, and Research Challenges," *Proc. IEEE*, vol. PP, no.99, pp. 1–21.

[4] ITU-T SG-15, "Spectrum Grids for WDM Applications: DWDM Frequency Grid," ITU-T G.694.1, 2012.

[5] G. Zhang *et al.*, ''A Survey on OFDM-based Elastic Core Optical Networking,'' *IEEE Commun. Surveys & Tutorials*, vol. 15, no. 1, 2013, pp. 65–87.

[6] G. Bosco *et al.*, ''On the Performance of Nyquist-WDM Terabit Superchannels Based on PM-BPSK, PM-QPSK, PM-8QAM or PM-16QAM Subcarriers," *J. Lightwave Tech.*, vol. 29, no. 1, 2011, pp. 53–61.

[7] D. J. Geisler *et al.*, ''Bandwidth Scalable, Coherent Transmitter based on the Parallel Synthesis of Multiple Spectral Slices Using Optical Arbitrary Waveform Generation,'' *Opt. Exp.*, vol. 19, no. 9, 2011, pp. 8242–53.

[8] E. Palkopoulou *et al.*, "Quantifying Spectrum, Cost, and Energy Efficiency in Fixed-Grid and Flex-Grid Networks [Invited]," *J. Opt. Commun. Net.*, vol. 4, no. 11, 2012, pp. B42–B51.

[9] M. Tahon *et al.*, "Valuing Flexibility in the Migration to Flexgrid Networks," *Proc. OFC/NFOEC '13*, 2013.

[10] M. Ruiz *et al.*, "Planning Fixed to flexgrid Gradual Migration: Drivers and Open Issues," *IEEE Commun. Mag.*, vol. 52, no. 1, 2014, pp. 70–76.

[11] X. Yu *et al.*, "Static Routing and Spectrum Assignment in Co-existing Fixed/Flex Grid Optical Networks," *Proc. OFC/NFOEC '14*, 2014.

[12] X. Yu *et al.*, "When and How Should the Optical Network Be Upgraded to Flexible Grid?," *Proc. ECOC'14*, 2014.

[13] Y. Huang *et al.*, "Mixed Line-Rate Transmission (112-Gb/s, 450-Gb/s, and 1.15-Tb/s) over 3560 km of Field-Installed Fiber With Filterless Coherent Receiver," *J. Lightwave Tech.*, vol. 30, no. 4, 2012, pp. 609–17.

[14] M. Xia *et al.*, "Split Spectrum: A Multi-Channel Approach to Elastic Optical Networking," *Opt. Exp.*, vol. 20, no. 28, 2012, pp. 29,143–48.

Biographies

Xiaosong Yu is currently working toward a Ph.D. degree in the State Key Lab of Information Photonics and Optical Communication at Beijing University of Posts and Telecommunications (BUPT), China. He was a visiting student at the University of California, Davis under the supervision of Prof. Biswanath Mukherjee. His research interests include elastic optical networks, network virtualization, and management and control of multi-layer optical transport networks.

Massimo Tornatore is an associate professor at Politecnico di Milano, Italy, and an adjunct associate professor at the University of California, Davis. He is an author of about 200 technical papers (with six best paper awards), and his research interests include the design and energy efficiency of telecom and cloud networks, through optimization and simulation. He received a Ph.D. degree in 2006 from Politecnico di Milano.

Ming Xia received a Ph.D. degree in computer science from the University of California, Davis, in 2009. He is currently a senior researcher in the IP Transport and Cloud team at Ericsson Research Silicon Valley, San Jose, California. He was an expert researcher at the National Institute of Information and Communications Technology, Japan. He serves as Associate Editor for Springer *Telecommunication Systems and Photonic Network Communications*, and a Guest Editor of the *Journal of Computers & Electrical Engineering* Special Issue on Ubiquitous Computing and Communications. His research interests include computer networks and cloud/data centers.

Jianping Wang is currently a professor and vice dean at the School of Computer and Communication Engineering (CCE), University of Science and Technology Beijing (USTB). She received her B.S. and Ph.D. degrees at Tianjin University in 1995 and 2000, respectively. From May 2000 to July 2005, she worked as a research fellow in the Department of Electronic Engineering at Tsinghua University. Her research interests include optimization of next generation optical networks and planning in multi-layer networks.

Jiawei Zhang received his Ph.D. degree from the State Key Lab of Information Photonics and Optical Communication at BUPT in 2014. He was a visiting student at the University of California, Davis from September 2012 to January 2014 under the supervision of Prof. Biswanath Mukherjee. He currently holds a postdoctoral position with the School of

Electronic Engineering, BUPT. He served on the Technical Program Committees (TPCs) for workshops on cloud computing systems, networks, and applications at IEEE GLOBECOM 2014 and IEEE ICC 2015. His research interests include software-defined elastic optical networks, network function virtualization, and intelligent radio optical access network.

YONGLI ZHAO is currently an associate professor at the Institute of Information Photonics and Optical Communications at BUPT. He received his B.S. degree in communication engineering and Ph.D. degree in electromagnetic field and microwave technology from BUPT in 2005 and 2010, respectively. He has had more than 150 journal and conference articles published. His research focuses on software defined optical networks, flexi-grid optical networks, network virtualization, and other areas.

JIE ZHANG is a professor and vice dean of the Institute of Information Photonics and Optical Communications at BUPT. He is sponsored by over 10 projects of Chinese government. He has had eight books and more than 200 articles published. Seven patents have also been granted. He has served as Co-Chair for ACP '12, '13, and '14, ICOCN '14, ChinaCom '12 and '13, and other conferences. His research focuses on optical transport networks, packet transport networks, and more.

BISWANATH MUKHERJEE [F] is a Distinguished Professor at University of California, Davis, where he was Chairman of Computer Science during 1997–2000. He received his B.Tech. degree from the Indian Institute of Technology, Kharagpur (1980) and his Ph.D. from University of Washington, Seattle (1987). He was General Co-Chair of the IEEE/OSA Optical Fiber Communications (OFC) Conference 2011, Technical Program Co-Chair of OFC '09, and Technical Program Chair of IEEE INFOCOM '96. He is Editor of Springer's Optical Networks Book Series. He has served on eight journal editorial boards, most notably *IEEE/ACM Transactions on Networking* and *IEEE Network*. In addition, he has guest edited Special Issues of *Proceedings of the IEEE*, *IEEE/OSA Journal of Lightwave Technology*, *IEEE Journal on Selected Areas in Communications*, and *IEEE Communications Magazine*. He has supervised 61 Ph.D. theses to completion and currently mentors 18 advisees, mainly Ph.D. students. He was the winner of the 2004 Distinguished Graduate Mentoring Award and the 2009 College of Engineering Outstanding Senior Faculty Award at the University of California, Davis. He is co-winner of Optical Networking Symposium Best Paper Awards at IEEE GLOBECOM 2007 and 2008. He is author of the graduate-level textbook *Optical WDM Networks* (Springer, 2006). He served a five-year term on Board of Directors of IPLocks, a Silicon Valley startup company. He has served on the Technical Advisory Boards of several startup companies, including Teknovus (acquired by Broadcom).

Transmission Media for an SDM-Based Optical Communication System

Kazuhide Nakajima, Pierre Sillard, David Richardson, Ming-Jun Li, René-Jean Essiambre, and Shoichiro Matsuo

ABSTRACT

Space-division multiplexing technology is an attractive candidate for overcoming a potential future "capacity crunch" in optical networks based on conventional single-mode fiber. Transmission media for SDM optical communication systems can be designed based on both single- and multi-core concepts. This article investigates the potential of various SDM transmission media.

INTRODUCTION

Space-division multiplexing (SDM) is expected to play a critical role in future optical communication systems. SDM can potentially be applied to a wide range of applications, and it can be supported by various transmission media. If we are to overcome a potential "capacity crunch" in existing single-mode fiber (SMF), one can first consider increasing the number of modes guided in the same core region and to use these to define multiple spatial data transmission channels. However, the overall transmission capacity is not simply proportional to the number of guided modes because the transmission characteristics are strongly dependent on the detailed mode coupling characteristics. Moreover, the differential mode group delay (DMGD) is not negligible since it causes a difference in detection timing. Alternatively, one can imagine increasing the number of cores and corresponding core density to improve the spatial utilization efficiency. Figure 1 shows the geometrical relationship between the relative effective core density and the cladding diameter D assuming hexagonally packed cores. The effective core density is defined as the ratio of the total effective area A_{eff} normalized with respect to the A_{eff} = 80 μm^2 and D = 125 μm associated with conventional SMF. Here, the minimum cladding thickness t was assumed to be 35 μm as required to achieve the required level confinement loss for the outermost cores. Black, red, and blue circles show the relationship when the value of the core pitch Λ was assumed to be 60, 70, and 80 μm, respectively. Figure 1 shows that a relative effective core density of more than five can be expected by reducing Λ and/or increasing the D value. However, the smaller Λ enhances the coupling between neighboring cores, and large values of D degrade the mechanical reliability of the fiber.

This article discusses the potential and subject of various SDM transmission media. We review single-core-based SDM transmission media, such as high-density optical fiber cable and few-mode fiber (FMF). Hollow core photonic bandgap fiber (HC-PBGF) is also considered as a potential candidate providing additional desirable transmission characteristics including low nonlinearity, low latency, and potential low loss. We cover multi-core-based transmission media, such as multi-core fiber (MCF) with and without core coupling. This section also describes progress in FM-MCF.

POTENTIAL OF SINGLE-CORE-BASED TRANSMISSION MEDIA

SINGLE-CORE HIGH-DENSITY OPTICAL FIBER CABLE

Progress on transmission technologies has improved both transmission capacity and cable density. In an electrical communication network, the cable density of a 400-pair copper cable was 0.23 mm^{-2}. In an optical network, a slot type 400 optical fiber cable achieved a cable density of 1.4 mm^{-2}, which is six times higher than that of copper cable. The latest 400 optical fiber cable realized a cable density of 4.1 mm^{-2} by employing a rollable fiber ribbon. In a conventional fiber ribbon, several optical fibers are completely fixed along their length to form a uniform sheet. This feature is particularly important to splice the multiple optical fibers effectively. In rollable fiber ribbon, the neighboring optical fibers are bonded partially along their length, which enables dense packing similar to that in copper cable while maintaining the operability obtained with conventional fiber ribbon. Here, it should be noted that the maximum core density of a 400 optical fiber cable is limited to around 5.3 mm^{-2} geometrically when we assume hexagonal-

Kazuhide Nakajima is with NTT Corporation.

Pierre Sillard is with Prysmian Group.

David Richardson is with the University of Southampton

Ming-Jun Li is with Corning Inc.

René-Jean Essiambre is with Alcatel Lucent.

Shoichiro Matsuo is with Fujikura Ltd.

ly packed 250 μm optical fiber (250 μm corresponds to the coating diameter of standard SMF) with a 2 mm thick cable sheath. Therefore, it can be said that the latest optical fiber cable employing rollable fiber ribbon has already achieved a cable density close to the geometrical limit.

Traditionally, these cables have been developed to make it possible to construct a transmission line more effectively and economically. If we are to overcome the geometrical limit while maintaining a single-core structure, we have to reduce the cladding diameter D to less than 125 μm (D = 125 μm corresponds to standard SMF) and bundle the fiber densely. The D value in a single-core fiber is mainly restricted by the confinement loss characteristic. Roughly speaking, we need a D value of 70–80 μm in order to maintain a typical mode-field diameter (MFD) and avoid transmission loss increase after the cabling process. It can be confirmed from Fig. 1 that we can expect to improve the space utilization efficiency of single-core fiber more than threefold if we can utilize a 70 μm cladding diameter.

However, it should be noted that the effective core density in a bundled thin-cladding fiber is reduced because the influence of coating thickness is neglected in Fig. 1. Although a multi-element fiber [1] can be considered one way of realizing high-density bundled fibers, we need to undertake further study of the optimum cladding and coating thickness. Moreover, the handling and/or operating properties of a single-core ultra-high-density optical fiber cable should be studied carefully because a backbone network has to maintain stable transmission performance during its lifetime once installed.

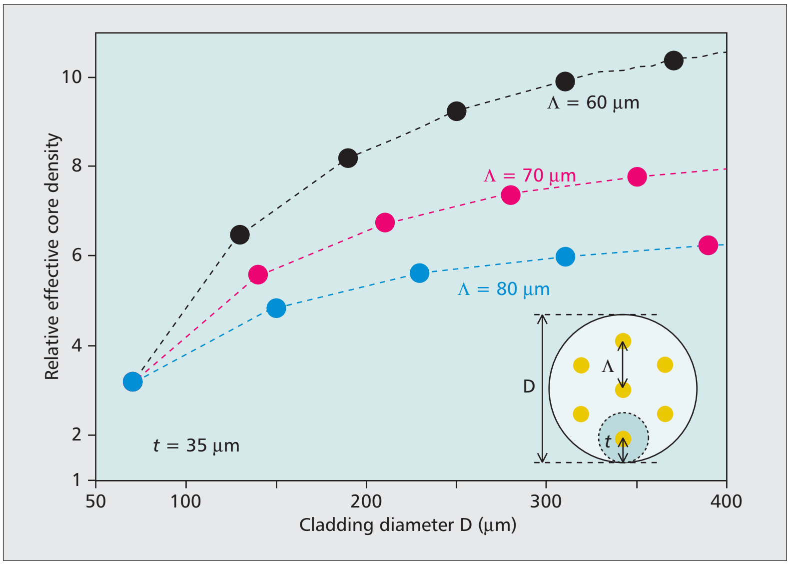


Figure 1. Geometrical relationship between relative effective core density and cladding diameter when t = 35 μm. Black, red, and blue plots show the results when Λ was assumed to be 60, 70, and 80 μm, respectively.

Few-Mode Fiber

FMFs adapted to SDM are of particular interest because of their ability to multiply system capacity by the (high) number of modes they can support. FMFs can be classified into two categories [2]. In the first weakly coupled category, modal crosstalk is minimized so that each LP mode is separately detected without using complex multiple-input multiple-output (MIMO) signal processing. Few FMFs belong to this category because LP modes spatially overlap, which is not favorable for crosstalk minimization. Most FMFs belong to the second low-differential-mode-group-delay (low-DMGD) category. Here, the DMGDs are minimized so that all modes can be detected simultaneously at reception, and MIMO can efficiently compensate for crosstalk.

Step-index profiles (red inset, Fig. 2) are well adapted to the first category [2]. When optimizing such FMFs, there is a trade-off to find between the differences of effective indices between LP modes, Δn_{eff}, which have to be as high as possible to limit crosstalk, and the effective areas, A_{eff}, which have to be as large as possible to limit nonlinearity. For a given number of LP modes, higher Min$|\Delta n_{eff}|$ is obtained by increasing the core index, yielding a smaller core radius and eventually smaller Min$|A_{eff}|$ of all LP modes. Min$|A_{eff}|$ also has to be decreased when the number of LP modes increases in order to keep Min$|\Delta n_{eff}|$ constant (red circles, Fig. 2). If 7-LP-mode fibers have acceptable characteristics (Min$|A_{eff}|$ > 90 μm^2 for Min$|\Delta n_{eff}|$ > 1 × 10^3 at 1550 nm), 9-, 10-, and 11-LP-mode fibers already reach critical levels (Min$|A_{eff}|$ ≤ 75 μm^2 for Min$|\Delta n_{eff}|$ = 1 × 10^{-3} at 1550 nm). Attenuation is another important characteristic. It increases with the number of LP modes, which is mainly due to the increase of Rayleigh scattering related to higher core indices and larger core radii required to support more LP modes. The average attenuation is estimated at 0.25 dB/km at 1550 nm for 7-LP-mode fibers with germanium-doped-core structures, which is near the highest acceptable value (9-, 10-, and 11-LP-mode fibers are ≥ 0.30 dB/km). The differences of attenuations between LP modes remain small (≤ 0.010 dB/km), however, due to similar confinements ensured by the step-index structure, which notably limits mode-dependent-loss impairments [3].

Trench-assisted graded-index profiles (blue inset, Fig. 2) are very well suited to minimize the DMGDs of FMFs of the second category. Low Max|DMGD|, ranging from ~0 ps/km for 2-LP-mode fibers to 73 ps/km for 12-LP-mode fibers at 1550 nm, can be obtained (blue squares, Fig. 2). DMGDs, however, are known to be highly sensitive to process variability that prevents reaching these optimized low values [2, 4] required to limit MIMO complexity. One way to compensate for this sensitivity is to concatenate fibers with LP modes with DMGDs with opposite signs. We have simulated the best possible concatenations of fibers representative of actual manufacturing using Gaussian distributions of profiles with average values corresponding to the optimized designs and with standard deviations that match manufacturing tolerances. The resulting DMGD-compensated links, which use more than 95 percent of these fibers, have Max|DMGD| delimited by the blue area of Fig. 2. DMGD-compensated 9-LP-mode links, with Max|DMGD| between 20 and 45 ps/km, could already allow to achieve thousands of kilometers

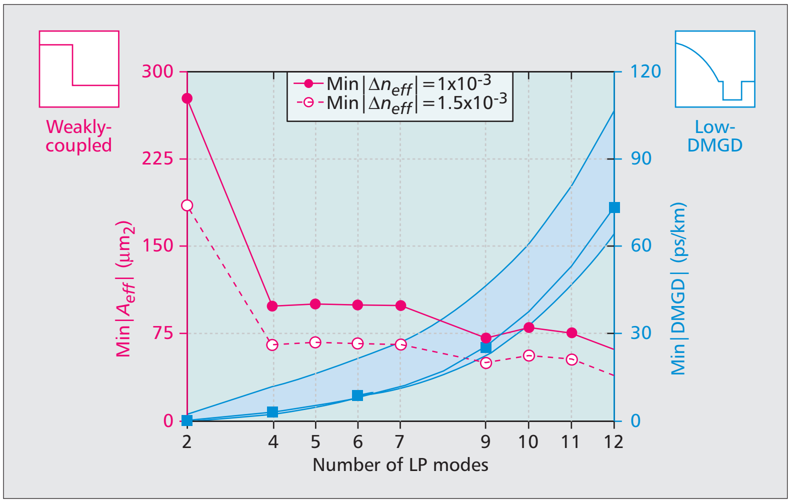


Figure 2. Min$|A_{eff}|$ at 1550 nm of step-index weakly coupled FMFs for Min$|\Delta n_{eff}|$ = 1.0 and 1.5 × 10^{-3} (red circles) and Max|DMGD| at 1550 nm of trench-assisted graded-index low-DMGD FMFs (blue squares for optimized designs, blue area delimiting the lowest and highest values of concatenations taking process variability into account) as a function of the number of LP modes. Lines are guides for the eye.

with 30 × 30 MIMO with tens of nanoseconds memory, while DMGD-compensated 12-LP-mode links require tighter process controls to reduce their DMGD values for practical use. Note that for this category, Min$|A_{eff}|$ are ≥ 90 μm^2 at 1550 nm, which ensures small nonlinearity. For a given number of LP modes, the attenuations are smaller than those of the first category, mainly because there is no constraint on Min$|\Delta n_{eff}|$ that imposes high core indices. The average attenuation is then estimated at 0.24 dB/km at 1550 nm for 12-LP-mode fibers. But the differences of attenuations, due to different mode confinements, are much higher and reach 0.035 dB/km. The 9-LP-mode fiber is at acceptable levels with an average attenuation of 0.22 dB/km and differences of attenuations ≤ 0.02 dB/km.

FMFs offer promising properties for SDM, but in order to give their full potential, the number of their LP modes has to be increased. For weakly coupled FMFs, the decrease of A_{eff} and the increase of attenuations are the main limiting factors of this increase, while for low-DMGD FMFs, these are the high sensitivity of DMGDs and the high differences of attenuations between LP modes. We have seen that if weakly coupled 7-LP-mode fibers and low-DMGD 9-LP-mode fibers already appear to be feasible, FMFs with higher numbers of LP modes require deeper investigation to overcome these limitations.

Photonic Bandgap Fiber

Hollow core photonic bandgap fibers (HC-PBGFs) (Fig. 3) provide a potential alternative medium for ultra-high-capacity transmission and offer a number of unique optical properties. These derive primarily from the fact that the optical signals propagate in air rather than in glass, and include the effective elimination of nonlinear effects between signal channels, low latency transmission due to the effective vacuum light speed signal propagation, and the potential for lower loss due to the reduced light matter interaction [5]. On this latter point it is to be appreciated that the spectral location of the minimum of loss in these fibers occurs at wavelengths around 2 µm. This results from the combination of surface roughness scattering at the inner core ring of the fiber, which dominates at short wavelengths, and the reduced multiphonon absorption due to the relatively small fractional overlap of the optical mode with the glass, which dominates at longer wavelengths. This is convenient since it means that transmission in HC-PBGFs, which can have bandwidths of several hundred nanometers depending on the detailed geometry, is potentially compatible with the use of the thulium doped fiber amplifier (TDFA), which is capable of providing gain over a > 300 nm spectral bandwidth in this wavelength range.

Substantial work on HC-PBGFs was undertaken in the early 2000s, and although encouraging results on loss reduction were obtained, with losses as low as 1.2–1.7 dB/km at 1565 nm reported by researchers at the University of Bath in 2005, this work effectively terminated with the downturn in the telecom market that occurred during this period. However, interest in the approach resurfaced around 2010 due to concerns over a potential future "capacity crunch," the ever increasing interest in low latency transmission, and the potential to exploit the multiple core guided modes associated with low loss PBGF fiber designs to define multiple independent spatial information channels in the context of SDM.

Since then significant progress has been made on further developing the technology. In particular, improvements have been made in controlling surface modes within HC-PBGFs (thereby extending the bandwidth available for data transmission at the current ~2 dB/km level from ~20 nm to >> 100 nm) and developing a more detailed understanding of the fundamental origins of loss in such fibers (especially in terms of surface scattering and the impact of characteristic distortions that occur during fabrication). Steady progress towards reducing loss to the < dB/km level in broadband structures is being made and experimentally validated models indicate that < 0.2 dB/km loss levels should ultimately be possible providing that the large-core geometries needed to realize low loss can be reliably achieved in practice. In addition, improved means of exciting, controlling, and exploiting the multiple guided modes within such fibers have been developed along with techniques for splicing HC-PBGFs both to themselves and to solid fibers.

These improvements have led to a number of high-capacity transmission experiments that have shown the viability of low latency transmission of DWDM signals [6]. Data rates as high as 30 Tb/s and distances in excess of > 1 km have so far been achieved for single-mode C-band transmission. Moreover, mode-division multiplexing at 73.7 Tb/s using 3-mode and 96-channel C-band dense wavelength-division multiplexing (DWDM) has also now been demonstrated [7]. The relatively high DMGD values of current fibers (~ ps/m level) and the intrinsically higher loss values as the mode order increases present significant challenges to the use of HC-PBGFs in long-haul SDM transmission, and approaches to mitigate these issues are the subject of ongoing research.

The viability of TDFA-amplified HC-PBGF data transmission at 2 μm has also now been shown for both single-wavelength and coarse WDM channels, with the majority of the components needed to implement DWDM transmission now emerging. Experiments illustrating the anticipated reduced nonlinearity and increased radiation hardness benefits of HC-PBGFs have also been undertaken, which, as well as being significant in terms of data transmission, underline the value of the technology in other application spaces, such as laser power delivery and sensing in harsh environments.

Work on producing longer lengths of HC-PBGF per draw is also progressing well, enabled by the development of new non-destructive tools that allow the quality of both preforms and fibers to be assessed and monitored along their length. All the evidence to date indicates that scaling the draw process to well beyond the current few-kilometer fiber yield per-draw level should ultimately be possible.

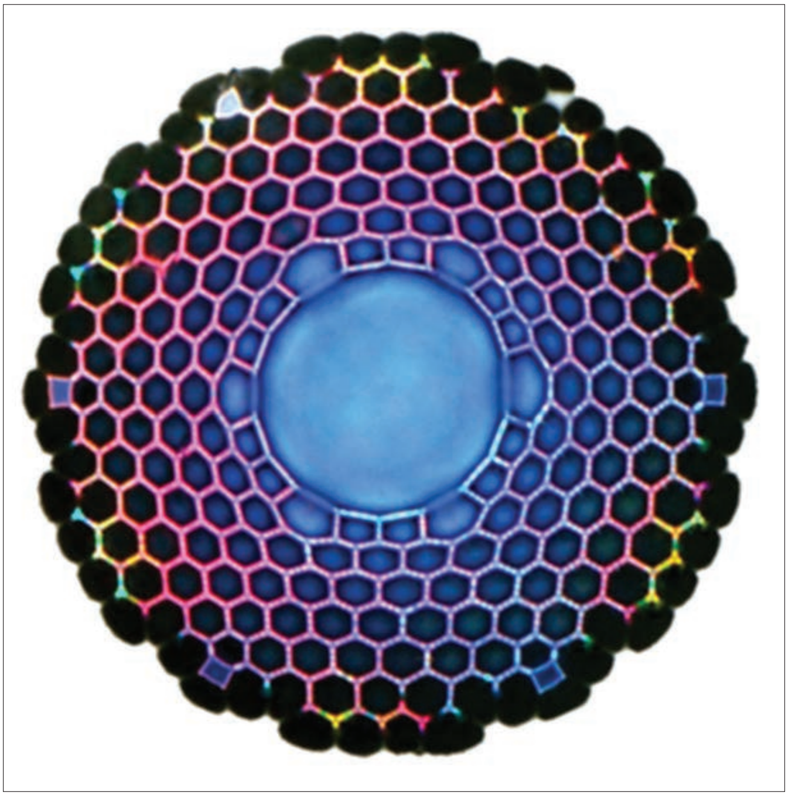


Figure 3. Microscope image of a typical large core HC-PBGF under white light illumination conditions. (Figure courtesy of S.R. Sandoghchi, Optoelectronics Research Centre, University of Southampton.)

Potential of Multi-Core-Based Transmission Media

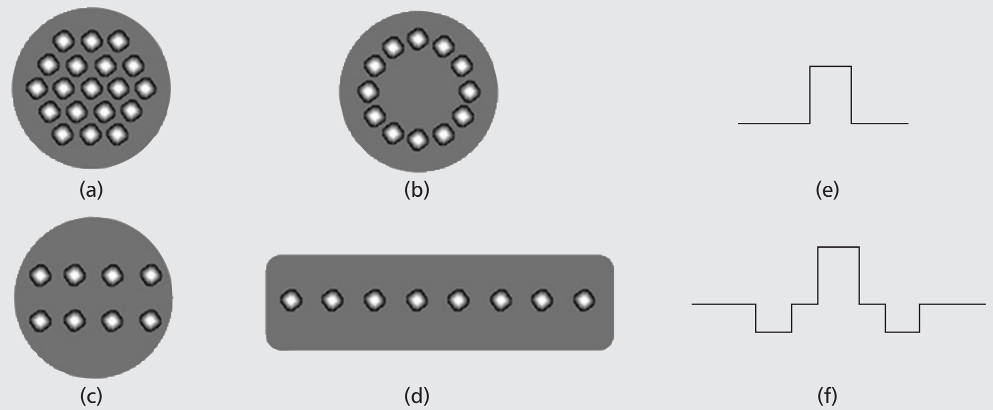
Multi-Core Fiber with Negligible Coupling

Multi-core fiber (MCF) is a potential fiber for SDM. Different MCF structures have been proposed, as shown in Fig. 4. Figure 4a is a hexagonal design that has a higher packing density. However, the crosstalk is worse because each core in the center of a hexagon has six neighbors. To avoid this problem, one can use a one-ring design as shown in Fig. 4b. Figure 4c is a linear array design with $n \times m$ cores, which can match semiconductor transceiver arrays. The coupling of a transceiver array to a multicore fiber can be done by direct butt coupling, through a lens array or an array of diffractive gratings. The designs of Figs. 4a–c have a round fiber cladding in which the number of cores is limited by the cladding diameter due to mechanical reliability considerations. To overcome this limit, we can use a ribbon MCF design as shown Fig. 4d, which offers the advantage of core scalability in one dimension and mechanical flexibility in the other direction. Because a ribbon fiber always bends around the axis parallel to the long side, the mechanical reliability is guaranteed by keeping the small dimension below the bending limit.

One of the most important aspects of MCF design is the crosstalk. For an ideal two-core fiber, from the coupled mode theory, the power crosstalk can be calculated by simply considering the power coupling coefficient κ and the mismatch $\Delta\beta$ in propagation constant. This simple model provides good guidelines for designing MCFs. The most important factor for low crosstalk is to reduce κ by reducing the overlap of the electrical fields of the modes. Another factor is the mismatch $\Delta\beta$ between the two cores, which reduces the maximum power that can be transferred from one core to another. Therefore, a heterogeneous core design can have lower crosstalk than a homogeneous core design.

Based on the coupled mode theory, the power in an ideal fiber oscillates sinusoidally with propagation distance. However, the measured crosstalk of fabricated homogeneous MCFs does not oscillate, but accumulates linearly along the fiber length. Such discrepancy is due to inhomogeneity along the fiber due to random perturbations and effects of fiber deployment conditions such as bending and twist. For a homogeneous MCF with random perturbations, constant phase can be maintained only in short sections of fiber. For a long length of fiber, it can be shown that the crosstalk is proportional to the fiber length L and the average correlation length Δl [8]. Reference [8] also derives an analytical crosstalk for an MCF under small bends by considering the propagation constant, bending radius, and core spacing.

To design an MCF with low crosstalk, it is important to reduce the overlap between the electrical fields of the two cores. Step and trench assisted profiles as shown in Figs. 4e and 4f have been studied. For an effective area of 80 μm^2 with a step index profile design, the core spacing needs to be greater than 45 μm to ensure a crosstalk of less than –30 dB after 100 km propagation. With a trench profile design, the core spacing can be reduced to about 37 μm for the same effective area.



Fiber				System						
No of cores	Core spacing (µm)	Fiber diameter (µm)	Effective area (µm²)	Span (km)	Data rate (Gbit/s)	No of channels	Channel spacing (GHz)	Reach length (km)	Total capacity (Tbit/s)	Spectral efficiency (bit/Hz)
12	37	225	81	52	456	222	50	52	1010	91.4
12 SM 2 FM	45	216	80	3	92, 103, 213	385	25	3	1050	109
7	56	200	99	55	128	40	50	6160	28.8	14.4
7	56	196	120	45.5	100	201	25	7326	140.7	28.0
19	35	200	72	10.1	86	100	100	10.1	305	30.5

(g)

Figure 4. Multi-core fiber designs (a–d), refractive index profiles (e–f), and the most recent SDM transmission experiments using MCF (g).

To accommodate more cores in an MCF, the cladding diameter can be increased. However, the fiber diameter is limited by the mechanical reliability requirements. The maximum fiber diameter should be smaller than 230 µm to ensure an acceptable failure rate under 50 mm bend radius. With this fiber diameter, the number of cores that can be put into a fiber is about 19 assuming the worst tolerable 100 km crosstalk.

Low loss and low crosstalk MCFs have been demonstrated and used in transmission experiments. Table in Fig. 4 summarizes most recent experiments using MCFs, which represents the state of the art in MCF transmission. A total transmission capacity of over 1 Pbit/s has been demonstrated with a spectral efficiency around 100 b/Hz. Long distance transmission over 7000 km has been achieved using a multicore EDFA.

MCFs are also attractive for high density short reach parallel optical data links. MCFs with linear geometry arrangements have been proposed for use with Silicon Photonics linear array transceivers. MCFs with 1 × 4 and 2 × 4 linear arrays have been made. A step index profile design with core separation of 47 µm results in crosstalk below –45 dB for a fiber length of 200 m, which is suitable for short reach applications. A ribbon MCF with 8 cores in a rectangular fiber cladding has been demonstrated with low crosstalk [9].

Coupled-Core Multi-Core Fiber

In contrast to a MCF designed to minimize linear coupling, a coupled-core MCF (CC-MCF) is designed to produce strong linear coupling between cores. When the signals exiting each core are not jointly processed, linear coupling generates crosstalk. In contrast, when MIMO processing between the cores is performed at the receiver, the data can be "untangled" even in the presence of strong coupling and an arbitrary large amount of crosstalk can be compensated.

There are several reasons why a CC-MCF can be an attractive fiber for long-haul SDM transmission [10]. First, such fiber allows for a higher density of cores than a MCF designed for low coupling. This can be visualized in Fig. 5a that shows high-density hexagonal-core structures based on 9-µm diameter cores and a core-to-core spacing of 18 µm. The numbers of cores represented are 91 and 19 for the 220- and 125-µm fiber diameters, respectively. A second advantage of CC-MCFs is the low DMGD they exhibit [11, 12]. In CC-MCF, DMGD between individual cores is greatly reduced since the signals propagate nearly as much in each core due to linear coupling, reducing the average DMGD to nearly zero, with some statistical spread of DMGD remaining. The width of the DMGD spread decreases with increasing coupling strength between cores. Similar effect occurs with mode-dependent loss (MDL) that also averages to zero due to linear coupling and exhibits a residual spread of MDL. Even though FMF can offer a higher spatial density of modes, it is more difficult to achieve a low DMGD between all modes in such fibers since the coupling between mode groups is insufficient to produce a good averaging of DMGD. A third advantage of CC-MCFs is the possibility of lowering the impact of fiber nonlinearity using these fibers. One notes that CC-MCFs can have a larger

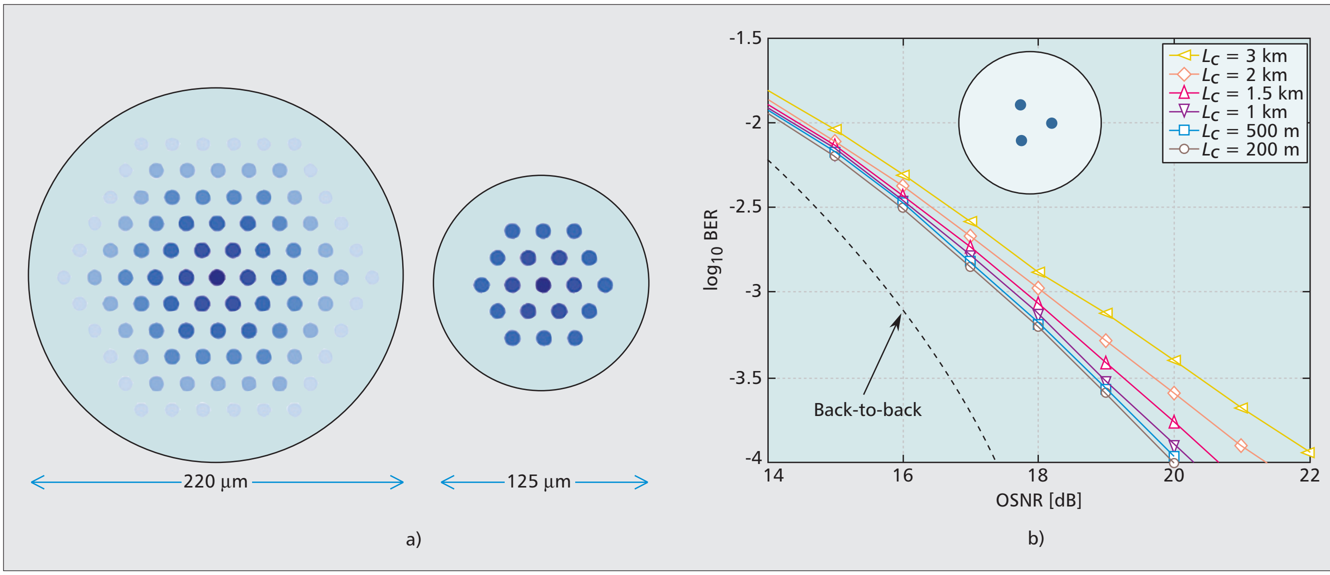


Figure 5. a) High-density hexagonal-core MCFs with 91 and 19 cores in a 220 and 125 μm diameter fiber, respectively. b) Bit error-rate (BER) as a function of optical signal-to-noise ratio (OSNR) after 1000-km transmission over a 3-core MCF having different coupling lengths L_c. A single channel is launched in each of the three cores with 7 dBm average power (from [13]).

mode field diameter, i.e. a larger effective area A_{eff}, per core than a MCF designed to suppress linear coupling between cores. The strength of linear coupling can be described by the coupling length L_c, which is the length over which the signal transfers from one core to another. A shorter L_c indicates a stronger coupling. Figure 5b shows the BER as a function of OSNR for various values of L_c after propagation over 1000 km in a 3-core CC-MCF [13]. One notes that, as L_c decreases, the BER improves suggesting that strong coupling reduces nonlinear distortions. A fourth advantage of CC-MCF, shared with most MCFs, is that the propagation and hardware at the transmitter and detector are similar for all modes. This homogeneity minimizes custom components, helping reduce the cost of commercial systems. Even though CC-MCFs have been considered for long-distance transmission more than two decades ago, modern use of these fibers to increase capacity started in 2011. The longest fiber transmission achieved using SDM MIMO processing is 4,200 km. The main current limitation for laboratory demonstrations of ultrahigh spectral efficiency using CC-MCFs is the large quantity of equipment necessary for MIMO processing.

Few-Mode Multi-Core Fiber

An FM-MCF is an uncoupled multi-core fiber with few-mode cores. In 2012, several organizations proposed FM-MCFs almost coincidentally: four-core, seven-core and 19-core FM-MCFs with all solid structures and a seven-core FM-MCF with a hole-assisted structure. All the FM-MCFs that were presented this year have two LP mode cores. One should be aware of many issues that are related to both MCF and FMF for designing an FM-MCF: inter-core crosstalk, inter-mode crosstalk, differential mode group delay (DMGD), core count, mode count, cladding diameter and core layout. The realization of high spatial multiplicity with an FM-MCF more than makes up for the issues. Two remarkable FM-MCFs have been reported recently. One is a two-LP mode, twelve-core fiber with a cladding diameter of 230 μm [14] and the other is a two-LP mode six-core fiber with a cladding diameter of 125 μm [15]. When we consider two-LP modes, LP_{01} and LP_{11}, MIMO technology enables us to use three spatial modes because the LP_{11} mode contains two degenerated modes, LP_{11a} and LP_{11b}. Accordingly, the 12-core FM-MCF realizes a spatial multiplicity of 36 (= 12 cores × 3 modes), which is the record multiplicity of SDM and the six-core FM-MCF realizes a multiplicity of 18, which is the largest multiplicity of 125-μm cladding fiber.

Figure 6 summarizes the characteristics of an FM-MCF that achieved the spatial multiplicity of 36 [14]. Figure 6a shows a cross sectional view of the FM-MCF. The cores, which had a multi-step index profile with an index trench as shown in Fig. 6b, were arranged on a square lattice structure. A multi-step index profile helps suppress DMGD. Figure 6c illustrates measured DMGD characteristics of the fabricated 12-core FM-MCF. The DMGDs were less than |600| ps/km over the C+L band for all cores owing to a multi-step index profile.

An index trench is useful for reducing inter-core crosstalk. For further reduction of inter-core crosstalk, the 12-core FM-MCF uses a heterogeneous core structure in which modes of adjacent cores have different propagation constants and the same A_{eff}. The structure is quite advantageous for reducing inter-core crosstalk. The set of heterogeneous cores, however, is restricted to two regarding cutoff wavelength of the third mode and bending loss of the LP_{11} mode. We cannot use the heterogeneous structure for a hexagonal closed pack structure, which is the most popular structure for MCFs, because three kinds of cores are needed to realize the heterogeneous structure with the hexagonal closed pack structure. The square lattice struc-

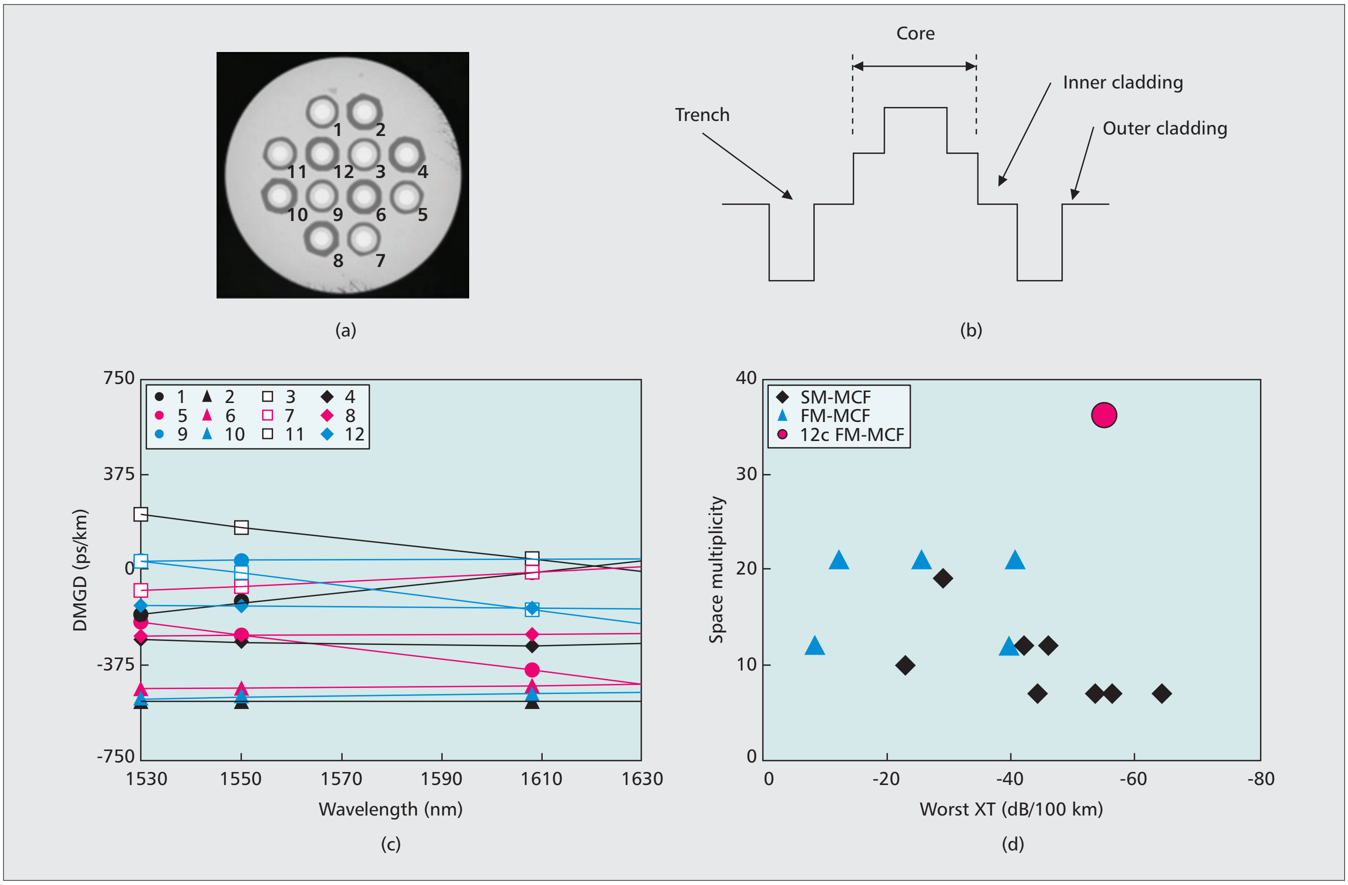


Figure 6. Characteristics of a fabricated FM-MCF.

ture, however, enables the heterogeneous core arrangement with only two kinds of cores. Because of the trench structure and heterogeneous core structure, the 12-core FM-MCF realizes 100 km crosstalk between LP_{11} modes lower than –50 dB at 1550 nm.

Figure 6d shows the spatial multiplicity (mode count × core count) of single-mode MCFs and FM-MCFs as a function of inter-core crosstalk. The 12-core FM-MCFs realize the highest spatial multiplicity of 36 and very low crosstalk simultaneously. Dense SDM (DSDM) transmission over the 12-core FM-MCF has been demonstrated. In the experiment, 20 WDM polarization-division multiplexed 32-quadrature amplitude modulated (QAM) signals were transmitted over 12-core × 3-mode fiber, and an aggregate spectral efficiency of 247.9 b/s/Hz was achieved.

The result demonstrated the potential of FM-MCFs. However, there are some issues to address: the optimum combination of core count and mode count, splicing, and amplifier and fabrication technology. We believe that the issues will be solved in the near future.

Conclusion

This article reviews the potential of various transmission media for use in SDM-based optical communication systems. SDM transmission media using multiple modes or cores can provide a more than tenfold transmission capacity increase compared to existing SMF. However, a tenfold improvement will not provide adequate room for growth if we are to construct an SDM-based backbone network with a 20–30-year lifetime. Although an FM-MCF can potentially bring us further transmission capacity by combining the benefits of increased mode and core count, this will require the simultaneous optimization of several key parameters, including crosstalk, transmission loss, A_{eff}, DMGD, and so on. HC-PBGF and CC-MCF potentially provide novel spatial channels with attractive features including low nonlinearity/latency and low mode dependent loss, respectively. However, if we are to use these spatial channels effectively, it will be essential to conduct a harmonized study with other SDM transmission technologies. It is to be appreciated that the usefulness of SDM technologies is not limited purely to overcoming the "capacity crunch" but also provides the potential for cost-per-bit reduction; arguably this, more than anything else, is the primary driver for research on SDM transmission media. It is important to consider the real overall value of SDM technologies in order to create commercially meaningful application of SDM transmission media. We hope that intensive work on SDM technologies will lead to a new era of optical communications, and that additional uses of SDM technology, for example, in sensing and laser applications, will also emerge in due course.

References

[1] S. Jain *et al.*, "Multi-Element Fiber for Space-Division Multiplexed Transmission," *Proc. ECOC '13*, Mo.4.A.2, 2013.

[2] P. Sillard, M. Bigot-Astruc, and D. Molin, "Few-Mode Fibers for Mode-Division-Multiplexed Systems," *J. Lightwave Tech.*, vol. 32, no. 16, 2014, pp. 2824–29.
[3] P. J. Winzer and G. Foschini, "MIMO Capacities and Outage Probabilities in Spatially Multiplexed Optical Transport Systems," *Opt. Express.*, vol. 19, no. 19, 2011, pp. 16,680–96.
[4] L. Grüner-Nielsen *et al.*, "Few Mode Transmission Fiber with Low DGD, Low Mode Coupling, and Low Loss," *J. Lightwave Tech.*, vol. 30, no. 23, 2012, pp. 3693–98.
[5] F. Poletti, M. N. Petrovich, and D. J. Richardson, "Hollow-Core Photonic Bandgap Fibers: Technology and Applications," *Nanophotonics*, vol. 2, no. 5–6, 2013, pp. 315–40.
[6] F. Poletti *et al.*, "Towards High-Capacity Fibre-Optic Communications at the Speed of Light in Vacuum," *Nature Photonics*, vol. 7, no. 4, Apr. 2013, pp. 279–84.
[7] V. A. J. M. Sleiffer *et al.*, "High Capacity Mode-Division Multiplexed Optical Transmission in a Novel 37-Cell Hollow-Core Photonic Bandgap Fiber," *J. Lightwave Tech.*, vol. 32, no. 4, Feb. 15 2014, pp. 854–63.
[8] M. Koshiba *et al.*, "Analytical Expression of Average Power-Coupling Coefficients for Estimating Intercore Crosstalk in Multicore Fibers," *IEEE Photon. J.*, vol. 4, no. 5, 2012, pp. 1987–95.
[9] O. N. Egorova *et al.*, "Multicore Fiber with Rectangular Cross-Section," *Opt. Lett.*, vol. 39, no. 7, 2014, pp. 2168–70.
[10] R.-J. Essiambre *et al.*, "Space-Division Multiplexing in Multimode and Multicore Fibers for High-Capacity Optical Communication," *Photon. J.*, vol. 5, no. 2, 2013, paper 0701307.
[11] R. Ryf *et al.*, "Impulse Response Analysis of Coupled-Core 3-Core Fibers," *Proc. ECOC '12*, 2012, paper Mo.1.F.4.
[12] S. O. Arik and J. M. Kahn, "Coupled-Core Multi-Core Fibers for Spatial Multiplexing," *Photon. Tech. Lett.*, vol. 25, no. 21, 2013, pp. 2054–57.
[13] S. Mumtaz, G. P. Agrawal, and R.-J. Essiambre, "Spatially Multiplexed Transmission in Multicore Fibers: Role of Core Coupling on System Performance," *Opt. Soc. Amer. Frontiers in Optics*, 2012, paper FW1D.2.
[14] Y. Sasaki *et al.*, "Few-Mode Multicore Fibre with 36 Spatial Modes (Three Modes (LP01, LP11a, LP11b) × 12 Cores)," *Proc. ECOC'14*, Th.1.4.1, 2014.
[15] T. Sakamoto *et al.*, "Moderately Coupled 125-μm Cladding 2 LP-Mode 6-Core Fiber for Realizing Low MIMO-DSP and High Spatial Density," *Proc. ECOC '14*, Tu.4.1.3, 2014.

Biographies

KAZUHIDE NAKAJIMA [M] received his M.S. and Ph.D. degrees in electrical engineering from Nihon University, Chiba, Japan, in 1994 and 2005, respectively. In 1994, he joined NTT Access Network Systems Laboratories, Tokai, Ibaraki, Japan, where he engaged in research on optical fiber design and related measurement techniques. He is currently a senior research engineer, supervisor (Distinguished Technical Member) of NTT Access Network Service Systems Laboratories, Tsukuba, Ibaraki, Japan. He has been acting as a Rapporteur of Q5, WP2, SG15 of ITU-T since 2009. He is a member of the Optical Society of America (OSA), and the Institute of Electronics, Information and Communication Engineers (IEICE) of Japan.

PIERRE SILLARD [M] received his engineering diploma from the Ecole Nationale Supérieure des Télécommunications, Télécom ParisTech, in 1994, and his Ph.D. degree in optics from the University of Paris VI in 1998, in collaboration with Thales Research & Technology, on the subject of nonlinear interactions in laser resonators. He has been working in the field of optical fibers and optical networks since 1999. He has been involved in the development and deployment of new fibers for FTTH, long-haul, and local area networks. His current research interests include low-loss and large-effective-area single-mode fibers, few-mode and multi-core fibers, and multi-mode fibers. He has published more than 150 papers and has been granted more than 60 patents. In 2004, he received the TR35 innovator award from *MIT Technology Review*. He is a member of the OSA, and serves as a reviewer and committee member for several journals and conferences.

DAVID RICHARDSON [F] joined the Optoelectronics Research Centre (ORC) at Southampton University as a research fellow in May 1989 and was awarded a Royal Society University Fellowship in 1991 in recognition of his pioneering work on short pulse fiber lasers. He has been deputy director of the ORC with responsibility for fiber and laser related research since 2000. His current research interests include, among others, optical fiber communications, microstructured optical fibers, and high-power fiber lasers. He has published more than 1000 conference and journal papers, has produced more than 20 patents, and was one of the co-founders of SPI Lasers Ltd, an ORC spin-off venture acquired by the Trumpf Group in 2008. He is an active member of the international photonics community and a Fellow of the OSA and IET. He was elected a Fellow of the Royal Academy of Engineering in 2009 and is currently a Royal Society Wolfson Research Merit Award holder.

MING-JUN LI [F] joined Corning Inc. in 1991 and is currently a Corporate Fellow. He has been a key inventor and contributor to many telecom and specialty fibers including bend-insensitive ClearCurve® optical fiber for FTTH, which received the R&D 100 Award and was featured as one of *Time*'s Best Inventions of the Year; LEAF® fiber, which won four industrial awards and was deployed with more than 30 million km of fiber in long haul networks, ultra-low PMD fiber and low loss fiber for high data rate transmission, low SBS fiber for analog signal transmission, high-bandwidth multimode fiber for data centers, various specialty fibers for use in connectors, fiber lasers, sensors, and endoscopes, and new multicore fibers and few mode fibers for future space division multiplexing. Recently, he has also been working on new glass measurement techniques for measuring chemically strengthened glass and laminate glass, and creating optical functionalities on glass substrates. His research efforts have resulted in the generation of a significant number of new products and revenue. He received the 1988 French National Prize on Guidedwave Optics for his work on Cerenkov second harmonic generation, the 2005 Stookey Award for exploratory research at Corning Inc., and the 2008 Northeast Regional Industrial Innovation Award from American Chemical Society. He is a Fellow of the OSA. He has served as Associate Editor and Coordinating Committee Member, and currently is Deputy Editor for the *Journal of Lightwave Technology*. He has also served as a Guest Editor for several special journal issues and as committee chair or member for many international conferences. He holds 116 U.S. patents, has published 3 book chapters, and authored and coauthored over 220 technical papers in journals and conferences. He received a B.Sc. degree in applied physics from Beijing Institute of Technology, China, in 1983, an M.Sc. degree in optics and signal processing from University of Franche-Comté, Besancon, France, in 1985, and a Ph.D. degree in physics from the University of Nice, France, in 1988.

RENÉ-JEAN ESSIAMBRE [F] received his Ph.D. from Université Laval in Canada and studied at the University of Rochester before joining Lucent Technologies (now Alcatel-Lucent) in 1997. His interests include optical fiber nonlinearity, space-division multiplexing, information theory, advanced modulation formats, and the design of optically routed networks. He has served on or chaired many conference committees including OFC, ECOC, CLEO, and the Photonics Society (formerly LEOS). He was program and general co-chair of CLEO Science & Innovation in 2012 and 2014, respectively. He is a recipient of the 2005 Engineering Excellence Award from OSA. He is a fellow of the OSA, and a Distinguished Member of Technical Staff (DMTS) at Bell Labs, Alcatel-Lucent.

SHOICHIRO MATSUO [M] received B.E and M.E degrees in electrical engineering from Kyushu University, Fukuoka, Japan, in 1988 and 1990, and his Ph.D. degree in production and information science from Utsunomiya University, Tochigi, Japan, in 2008. He has been with Fujikura Ltd., Chiba, Japan, since 1990, working on research and development of transmission fibers for long-haul and FTTH networks, and rare-earth-doped fibers, photonic band-gap fibers, and manufacturing technology of these optical fibers. Now, he is a general manager of the Optics and Electronics Laboratory, Fujikura Ltd. He is an author or coauthor of more than 100 journals and conference papers, and is an inventor of more than 90 granted patents globally. He is a member of the OSA, the Japanese Society of Applied Physics (JSAP), and the IEICE.

Physical Layer Transmission and Switching Solutions in Support of Spectrally and Spatially Flexible Optical Networks

Roland Ryf, S. Chandrasekhar, Sebastian Randel, David T. Neilson, Nicolas K. Fontaine, and Mark Feuer

ABSTRACT

Fiber optic communication networks in the era of large data centers are required to provide large capacity, often orders of magnitude larger than the capacity of a single-mode fiber. The transport capacity must also be rapidly reconfigurable to adapt to the applications and services provided by the data centers. This article explores new transponders capable of providing larger capacities and making optimal use of the optical transmission spectrum, new optical switches capable of flexible routing of wavelength and multiple spatial channels, and new amplifiers supporting multiple spatial channels and broad optical bandwidth.

INTRODUCTION

Optical fiber communication technology is essential to provide the bandwidth and long distance connectivity required by the data centric communication infrastructure. Whereas in the past optical networks were often perceived as slowly changing aggregators of fixed endpoints, the build-out of massive data centers has dramatically changed the requirements imposed on them. Today's data centers demand a strong synergy between the data processing capability inside the data center and the data transport provided by the external optical network. For example, required bandwidth may exceed by orders of magnitude the capacity accessible in a wavelength-division multiplexed (WDM) single-mode fiber system, driving a need for optical components supporting multiple standard single-mode fibers (SMFs) or parallel spatial channels, also referred to as space-division multiplexing (SDM). Since capacity requirements can drastically change over time, following traffic patterns or scheduled activities, rapid reconfiguration time of a few seconds or less is highly desirable. To sustain the exploding capacity demands of data centers, a new generation of optical devices is required.

In particular, new transponders capable of providing larger capacities while making optimal use of the optical transmission spectrum are needed. Despite the impressive progress in high-speed digital signal processing (DSP) and opto-electronic conversion, meeting the capacity requirements will make it necessary to bundle multiple channels within a single transponder, thus forming so-called *super-channels*. Super-channels can be formed in the optical spectral domain by combining closely spaced wavelength channels, but also in the spatial domain by bundling multiple spatial channels[1]. From the transponder perspective, spectral and spatial super-channels exhibit similar architectures, but the spatial super-channel offers potential for cost saving by reducing the number of required lasers. However, it is in optical switching and optical amplification devices where the spatial super-channel concept has the greatest impact.

Optical switching in combined WDM/SDM networks presents a formidable technological challenge, and the optimum solution strongly depends on the number of WDM and SDM channels and the required switching granularity. For example, in a situation where the number of SDM channels significantly exceeds the number of WDM channels, it is conceivable to perform switching using a wavelength transparent switch (i.e., switching the full WDM spectrum at once). Switches capable of transparent switching have been demonstrated in the past for switch fabrics with more than 1000 × 1000 ports. In the short term, however, an upgrade path for single-fiber-based transmission systems must be provided, and the initial number of SDM channels is expected to be small compared to the 100 WDM channels present in a typical transmission system today. Wavelength selective switches (WSSs) capable of switching between N input ports and M output ports are therefore required. Implementing a fully

Roland Ryf, S. Chandrasekhar, Sebastian Randel, David T. Neilson, and Nicolas K. Fontaine are with Bell Laboratories, Alcatel Lucent.

Mark Feuer is with the College of Staten Island.

 0163-6804/15/$25.00 © 2015 IEEE

flexible $N \times M$ WSS that can switch any wavelength from any input spatial channel to any output spatial channel is technically challenging, although early demonstration with five input/output ports has been reported[2]. The switching capacity can be dramatically increased if the switching capability is restricted to a subset of switching configurations. One subset of particular interest is the switching of spatial super-channels as a whole. In this configuration there is no switching capability within the super-channel: all spatial sub-channels at a given wavelength are routed the same way. At the device level, this translates into an architecture referred to as *joint switching* [3], where a common switching element (micro-electro-mechanical system, MEMS, mirror or liquid crystal on Silicon LCoS spatial light modulator) is used to steer multiple light beams at the same time.

Using joint switching, it is possible to extend currently available $1 \times N$ WSSs to support super-channels with S spatial channels, where S is independent of the switching elements, and can be scaled up to at least 7, limited only by the lens design of the WSS. Another key feature of WSSs for high-capacity systems is the ability to provide flexible bandwidth allocation for the switching channel. WSS based on LCoS can provide such flexible bandwidth allocation limited only by the spectral resolution of the WSS optical design, thanks to the large number of switching elements (pixels). Additionally, a large number of pixels are advantageous to increase the total supported bandwidth and also support a large number of switching ports: $1 \times N$ WSSs with port number N larger than 90 are theoretically possible.

The concept of spatial super-channels also plays an important role in optical amplification, where new concepts like cladding pumped multicore amplifiers [4] or few-mode fiber amplifiers [5–7] are currently under investigation as a way to provide efficient amplification of multiple spatial channels.

In the first part of the article we review new trends in optical transponders, present the concept of a super-channel, and discuss how it can be applied to increase the transponder capacity. The following three sections discuss optical switching applied to spatial super-channels, and present a new kind of WSS that supports multiple input and multiple output ports as a new building block for optical networks. Furthermore, we review the scalability of LCoS technology in terms of bandwidth management and number of ports. The final two sections are dedicated to high-capacity optical amplification achieved either through multiple spatial channels or by increasing the bandwidth of the optical amplifier.

Spectrally Flexible High-Capacity Transponders

There has been great progress in the last five to eight years toward the realization of software-defined optical transponders (SDOTs) that address the needs of next generation high-capacity optical systems as well as enabling novel networking architectures. This progress is attributed to the development of high speed digital-to-analog converters (DACs) and analog-to-digital converters (ADCs), and the associated digital signal processing (DSP) engines[8]. Thanks to these electronic devices, the preferred mode of long-reach transmission is now to encode information in the full electric field of light, using coherent intradyne detection to gain access to the complete phase and amplitude of the received signal. Such SDOTs are beginning to reshape the optical networking landscape, leading toward dynamic, flexible, and agile transport networks.

Four-channel DACs coupled to dual-polarization vector modulators are being used to generate advanced multi-level modulation formats such as quadrature phase shift keying (QPSK), quadrature and amplitude modulation (QAM), orthogonal frequency-division multiplexing (OFDM), and, more recently, 4D modulation. These formats support spectral efficiencies from 2 b/s/Hz to well over 10 b/s/Hz. The sampling speed of the DACs has reached 75 GSa/s, enabling generation of signals with symbol rates as high as 70 GBaud. In addition, since the DACs have effective number of bits (ENOB) of over 6 bits, digital spectral shaping-cum-equalization is commonly used to mitigate component bandwidth limitations, incorporate electronic dispersion precompensation and fiber non-linearity pre-compensation, and limit the spectral extent by suitable root-raised-cosine (RRC) filtering. All these attributes have allowed generation of spectrally efficient compact modulation formats that can be juxtaposed closely in frequency to achieve high capacity. The information rate carried on a single optical carrier can be varied by changing either the modulation format, the symbol rate, or a combination of both, all typically without any hardware changes.

Four-channel ADCs used as part of the polarization diversity intradyne coherent receiver digitize the received electric field of the signal, and the DSP following it recovers the full information through a series of steps [8]. There has been more progress on the ADC front than on the DAC front, particularly in the instrumentation field. Sampling speeds on digital sampling oscilloscopes range from 50 GSa/s to 160 GSa/s on each of four analog channels, with analog bandwidths approaching 65 GHz. The ADCs used in SDOTs are known to operate up to 70 GSa/s. The DSP engines in the receiver allow for compensation of accumulated fiber chromatic dispersion, partial mitigation of intersymbol interference (ISI), carrier frequency and phase recovery, and bit error ratio (BER) counting. An important characteristic of the coherent receiver is the fact that the beat signal between the local oscillator laser (OLO) and the received signal is limited to the bandwidth of the receiver, which is typically close to the symbol rate of the signal being detected. As a result, multiple signals with different wavelengths can be incident on the receiver, and the OLO can be tuned to select only the channel of interest without the need for any optical filter to select the received channel. These characteristics have opened the doors to what is now called *colorless* detection, and is another key enabler for the next generation optical networks.

Four-channel DACs coupled to dual-polarization vector modulators are being used to generate advanced multi-level modulation formats such as quadrature phase shift keying, quadrature and amplitude modulation, orthogonal frequency-division multiplexing, and, more recently, four-dimensional modulation.

With increasing integration densities, in both the electrical and the photonic domains, the analog interfaces start becoming a bottleneck and crosstalk between spatial channels will start to occur. In the electrical domain, such crosstalk is likely to occur between high-bandwidth parallel analog interfaces.

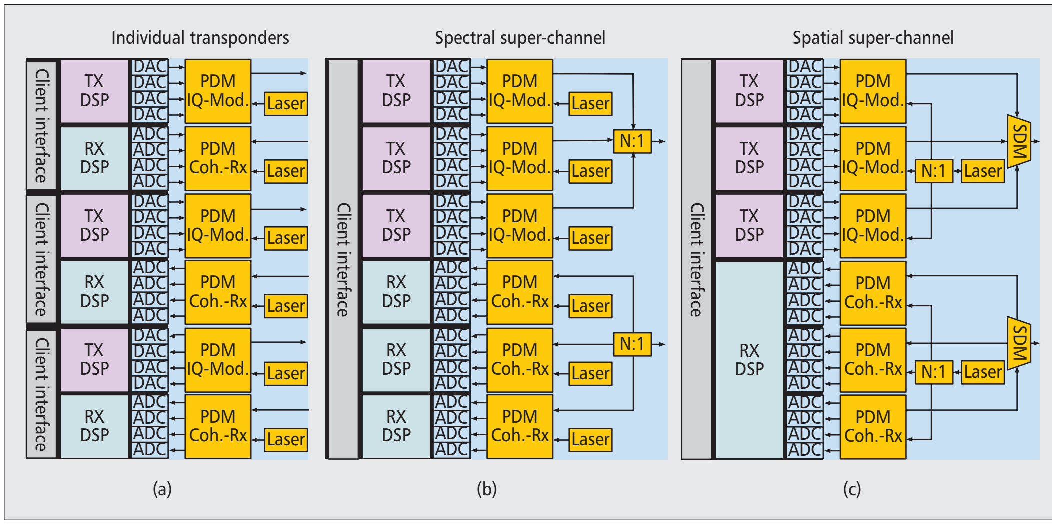


Figure 1. Examples of super-channel architectures with three subchannels: a) individual transponders; b) spectral super-channel transponder; c) spatial super-channel transponder.

The ability of the SDOT to generate and receive a flexible format/symbol rate signal with spectral shaping has led to the evolution of flex-grid optical networking, where the available optical spectrum is partitioned non-uniformly in multiples of a *spectral slice*, typically 12.5 GHz wide[9]. This approach allows better utilization of the optical spectrum with minimal wasted bandwidth between modulated channels. Such flex-grid networking requires reconfigurable optical add/drop multiplexers (ROADMs) that are constructed from WSSs which allow flexible passband allocation. Nevertheless, concatenation of several such flex-grid WSSs imparts similar system penalties due to spectral truncation as fixed grid WSSs. The ability to spectrally "pre-emphasize" the signal spectrum at the transmitter using the DACs to account for such truncation allows some recovery of penalties.

Transponders Supporting Spectral and Spatial Super-Channels

In order to further increase the capacity of transponders in a cost-effective way, spectral super-channels formed by combining multiple closely spaced frequency channels have been introduced. In Fig. 1a, we show a system architecture with individual transponders for each spectral channel. While this concept provides a high degree of flexibility and fine granularity, it does not scale well in terms of cost and power consumption. Continuous electronic and photonic integration is required to further increase the transponder efficiency on a per bit measure. Significant advantages for spectral super-channels are achieved by integrating multiple components in the electronic but also in the optical domain (Fig. 1b). As the capacity requirements scale in excess of the single-mode fiber limit, parallel spatial channels are required, and the concept of super-channels can be extended to the spatial domain.

In Figs. 1b and 1c, we show examples of spectral and spatial super-channel transponders with three sub-channels. The DSP engines are now integrated and include three processing cores, following a similar trend to that observed in the case of central processing units (CPUs) over a number of years. In addition, super-channel transponders include multiple electro-optical components such as polarization diverse inphase/quadrature modulators (PDM IQ-Mods) and polarization diverse coherent receivers (PDM Coh. RX). Progress in photonic integration technologies will allow combining these components into a single package. In the case of the spectral super-channel, individual lasers are required for each subchannel at the transmitter and the receiver. The transmit and receive signals can be multiplexed and demultiplexed in a colorless fashion using power combiners/splitters (N:1). In the case of the spatial super-channel, only one laser at the transmitter and one at the receiver are required and low-loss space-division (SDM) multiplexers and demultiplexers can be used.

With increasing integration densities in both the electrical and photonic domains, the analog interfaces start becoming a bottleneck, and crosstalk between spatial channels will start to occur. In the electrical domain, such crosstalk is likely to occur between high-bandwidth parallel analog interfaces. Note that in the case of three subchannels, already 12 ADCs and 12 DACs must be interfaced to the optics, all from a single chip. In the optical domain, crosstalk between spatial modes can occur, for example, in fibers, where the light propagating in different modes gets mutually coupled. By adding multiple-input multiple-output (MIMO) equalization capabilities to the receiver DSP of super-channel transponders, this crosstalk can be compensated or at least mitigated to a large extent [1]. The additional resources required to implement the MIMO equalizer matrix grow linearly with the number of subchannels, but are generally only a modest increase compared to the necessary resources for the DSP functions in a coherent receiver, which are typically dominated by chromatic dispersion compensation and forward error correction [10].

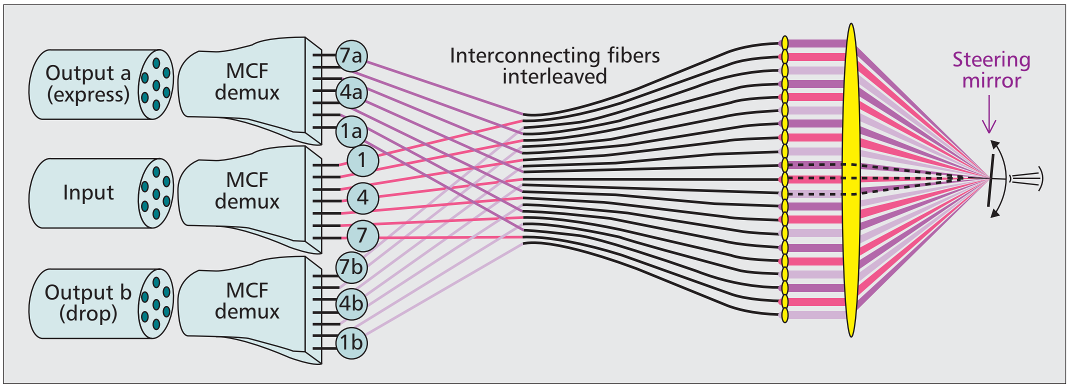


Figure 2. A single mirror can simultaneously steer multiple beams that impinge on it at different angles of incidence. In a WSS, there are many such mirrors, one for each wavelength. In this illustration, spatial modes from multi-core fibers are converted to single-mode fiber signals in the component labeled MCF demux (after [3]).

For fiber optic networks operating at 10–100 Gb/s per channel, ROADMs based on 1 × N WSSs have brought substantial benefits in network efficiency and flexibility, so extending these WSSs to support multiple spatial channels is a promising direction.

Wavelength-Selective Switches Supporting Spatial Super-Channels

We have seen that transponders incorporating spectral or spatial super-channels can enable terabit-class long distance optical links. To ensure that such terabit links will be supported by flexible networks with transparent optical routing, the industry is working to develop ROADMs that support spatial modes as smoothly as they do wavelength channels. For fiber optic networks operating at 10–100 Gb/s per channel, ROADMs based on 1 × N WSSs have brought substantial benefits in network efficiency and flexibility, so extending these WSSs to support multiple spatial channels is a promising direction.

Undoubtedly, the greatest network flexibility would be achieved by WSSs that could route each wavelength and each spatial mode independently, but such independence introduces complexity and cost that may not be justified. If networks with terabit-class interfaces rely on spatial super-channels (i.e., each end-to-end demand is transmitted across a group of spatial modes in the same fiber at the same frequency), all of the subchannels in a given super-channel must be routed together to reach the same endpoint, so there is no added network value in an optical switch that can route them independently in that scenario. Therefore, the choice of super-channel format goes hand in hand with a successful WSS design.

For example, consider a 1 Tb/s spatial super-channel comprising seven channels of ~140 Gb/s, occupying seven spatial modes at a single frequency. Spatial super-channel switching can be realized by using seven conventional 1 × N WSSs. Alternatively, Fig. 2 shows super-channel switching by exploiting ancillary optical paths that exist in a single standard 1 × 20 WSS in order to convert it into a septuple 1 × 2 WSSs capable of switching seven subchannels simultaneously[3]. The concept is referred to as *joint switching* and relies on the fact that multiple beams at different angles of incidence may be switched by a single steering element formed by an LCoS phase modulator or a MEMS mirror as shown in Fig. 2, where the 20 + 1 single-mode fiber ports are reconfigured to form 2 + 1 groups of seven-fiber ports. Scaling up this approach to reach a seven-subchannel 1 × 8 WSS (1 × 8 WSSs are used in most of today's 10–100 Gb/s networks) would require a WSS with a total of 7 × (1 + 8) = 63 single-mode fiber ports, comparable to the largest WSS demonstrated in research to date. Nevertheless the tilt requirements for the switching element, would not be any different for a septuple 1 × 8 WSS than for a conventional 1 × 8 WSS, using the interleaved arrangement of Fig. 2. Crosstalk among spatial subchannels can occur within the multiple 1 × N WSSs, but can be mitigated by appropriate subchannel layouts. The joint switching architecture can be applied to any type of SDM fiber (e.g., few-mode fiber, multicore fiber, fiber ribbon, or multiple single-mode fibers) by incorporating the appropriate spatial multiplexer. For example, laser-inscribed 3D waveguides can be used to transition between multicore fiber and the multiple single-mode inputs of the WSS.

A system that routes spatial super-channels can be tolerant to crosstalk among spatial modes (a.k.a. mode mixing), since the outputs from all spatial modes are available at the receiver for DSP. This could be important for networks utilizing few-mode fibers, which are likely to have strong mode mixing. On the other hand, if the network is based on multi-core fiber with low spatial-mode crosstalk, it might be beneficial to use spectral super-channels, ensuring that a bundle of adjacent wavelengths reach the receiver to facilitate compensation of fiber nonlinearity. Finally, note that the ROADMs in a network make up a large part of its total cost when only a few channels are lit, but the transponders dominate the network's cost once it is fully loaded. Thus, a super-channel format using more expensive ROADMs might be justified if the format enables lower-cost transponders.

New $N \times M$ Wavelength Crossconnect Switches

Currently, the key component of reconfigurable optical mesh networks is the 1 × N WSS. A 1 × N WSS directs any set of wavelengths at an input

Liquid crystal on silicon phased array is one of the most attractive switching technologies for building wavelength-selective switches today, since it allows flexible wavelength allocation, and reuses a technology platform developed for displays.

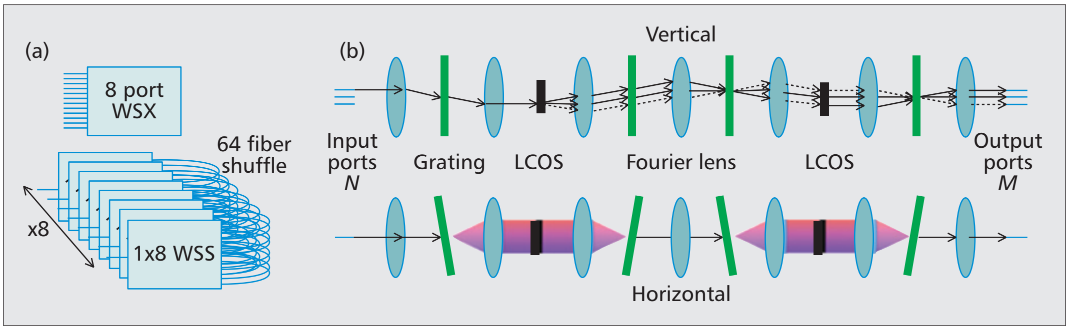


Figure 3. a) Functionality of a wavelength selective cross connect built using many WSSs and fiber shuffle; b) schematic of the optical arrangement for an $N \times M$ wavelength crossconnect [2].

to any of the N outputs and manages the routing of the wavelength channels within the network. To have colorless, directionless, and contentionless (CDC) switching between N inputs and N outputs, these WSSs are placed into more complex switching nodes. A CDC $N \times N$ WSS or wavelength selective crossconnect (WSX) allows the same wavelength to be routed between any set of input and output ports even if that wavelength is already used by another pair of ports. A contentionless switching node based on the route and select architecture is constructed by interconnecting the outputs of N $1 \times N$ WSSs such that each WSS is connected to all other WSSs so that any wavelength or group of wavelengths entering a port can be routed to all output ports. The output ports could be connected to other transmission fibers or transponders.

These nodes are particularly suited for LCoS technology, which typically shows a higher level of crosstalk compared to MEMS-based WSSs. In fact, because all signals pass through two WSSs, the total crosstalk rejection in dB is doubled, which relaxes the crosstalk requirement of a single WSS. Large-port-count nodes require high-port-count LCoS-based switches (20+ ports), which typically have crosstalk levels around 25 dB. After double passing, a crosstalk rejection of 50 dB is obtained, which is sufficiently large for cascaded optical elements in coherent transmission systems. However, the same effect will also apply to the passband, where the signal will now be subject to twice the passband narrowing effect, and therefore optical devices with high spectral resolution are desirable.

The tremendous switching capacity of LCoS technology is mostly unused in $1 \times N$ WSSs. Tapping the unused capacity enables construction of $N \times N$ switching nodes in a single package with nearly identical optics, and similar complexity and form factor as a single $1 \times N$ WSS. In the example of Fig. 3a, a single 8 $N \times 8$ WSX can replace eight WSSs, resulting in huge cost savings. Additionally, compared to switching nodes built from a $1 \times N$ WSS, the $N \times N$ WSX is contained in a single free-space section where the optics can be optimized independently for wavelength and steering. As a result, the $N \times N$ WSX has reduced crosstalk, suffers less passband narrowing, and can have nearly the same loss as a single WSS (only one extra bounce off an LCoS device contributes as added loss).

LCoS technology is favorable for building WSXs because the beam-steering element can be defined in software as a phase hologram after the system is aligned. Compared to fixed MEMS mirrors, the hologram boundaries can be curved, gapless, have any aspect ratio, and have subpixel resolution (averaging of many pixels). A WSX contains two 2D arrays of pixels facing each other. Wavelengths are dispersed on one axis using diffraction gratings, and the ports are imaged onto the LCoS in the other direction. In between the two arrays, wavelengths are imaged (4-f system), and the ports are Fourier transformed (2-f). The mirrors are then programmed in the software to match the beam layouts for every wavelength and port. This corresponds to 2000 mirrors for a 20×20 WSX with 100 wavelength channels. Other strengths of LCoS technology in a WSX is the ability of multicast functionality (beam-splitting) and inband spectral shaping (optical equalization).

Scalability of Wavelength-Selective Switches Based on Liquid Crystal-on-Silicon

Liquid crystal on silicon phased array is one of the most attractive switching technologies for building WSSs today, since it allows flexible wavelength allocation and reuses a technology platform developed for displays. The ability to program the hologram on the LCoS phase modulator array allows other attenuation, beam splitting, and shaping functions to be implemented.

Arrays with larger numbers of pixels promise higher switching functionality, but simply using arrays with smaller pixels (pixel widths as small as 4 μm are common for high-resolution LCoS backplanes) may not be sufficient, because phase modulator LCoS panels for operation at telecom wavelengths require a typical cells thickness in a range from 4–15 μm, which is much larger than the cell thickness required for visible display applications. When the pixels' width to thickness approaches a 1:1 aspect ratio, the pixels do not provide full modulation depth, and the holograms are not properly resolved. Failure to resolve the holograms can impact switching efficiency and crosstalk and wavelength pass band shapes.

In the wavelength direction, the spectral resolution is typically determined by the resolution of the optical system and not the pixel granularity. Increasing the number of pixels in this axis can be advantageous to extend the range of wavelengths (e.g., to support multiple optical transmission bands as a seamless device). Additionally, when combined with an increased resolution of the optical system, the switching granularity (number of switchable wavelength channels) can also be improved if desired.

Increasing pixel count will have a dramatic impact on the number of ports. The number of spatial switching states and hence ports that can be supported by an LCoS array depends on the number of pixels illuminated and the acceptable crosstalk between the ports. There are two main mechanisms of crosstalk. The first is given by the partial overlap of the steered beam into the neighbor ports, and can be reduced by increasing the beam steering angle difference between ports. The second mechanism of crosstalk is generated by imperfection of the beam steering hologram, which creates undesired alternate diffraction orders. If we choose the beam steering angle difference between ports to give 43 dB crosstalk suppression, and choose a beam size on the LCoS where the $1/e^2$ power fills 2/3 of the pixels to avoid clipping effects, the number of switch states is given by

$$K = \frac{2.P}{3.Q},$$

where P is the number of resolvable pixels in the steering direction and Q is the minimum number of pixels per 2π phase. For a 1920 × 1080 LCoS array using the 1080-pixel direction for switching with $Q = 8$ pixels per 2π, this gives 90 switch states, while for a 4k (4096 × 2400 pixels) LCoS array using the 4096-pixel direction for switching with $Q = 8$ gives 340 switch states. Typically, not all of these switch states can be used because of crosstalk from other diffraction orders. There is a zero order reflection from the LCoS, which typically makes some ports unusable because of reflections from the LCoS backplane. Depending on hologram design, certain harmonics of the diffracted beam may require elimination of ports or irregular port spacing, which results in a reduction in the number of ports.

The switch states can be divided among multiple parallel switches, so an LCoS with 90 switch states could provide a pair of up to 45 port switches. The 340 switch states of a 4K array might be used to create eight such 45-port switches.

For $N \times M$ wavelength selective crossconnects, there are two switching planes, so the geometric crosstalk is the result of two spatial steering steps and therefore additive. If we keep the same net crosstalk (43 dB), we get the square root of two times as many switch states. The 4k array would have 480 switch states, allowing 22 switches with 22 ports per array, so using two LCoS arrays, a 22 × 22 wavelength cross-connect is theoretically possible.

The availability of $1 \times N$ and $N \times M$ WSSs with such high port counts is expected to have a dramatic impact on SDM and WDM meshed optical networks.

Optical Amplifiers Supporting Multiple Spatial Channels

Optical fiber amplifiers have been instrumental for the success of WDM systems. A single amplifier is capable of amplifying a wide frequency range extending over multiple terahertz. This large bandwidth can support 100 or more wavelength channels. The Erbium-doped fiber amplifier (EDFA), in particular, has been extremely successful, because it offers amplification in a frequency window that coincides with the low loss region of the standard single-mode fiber. Modern EDFAs present low noise figures, sometimes only a few tenths of a dB above the theoretical limit of 3 dB, can operate over a wide gain range, and can provide output power in excess of 1 W.

In order to further increase the capacity of fiber optic links above the capacity limit of a single SMF, amplifiers capable of efficiently amplifying multiple spatial channels are required. As with transponders, integration is considered to be a main driver for cost reduction. A typical EDFA consists of the control electronics that monitor the power levels in the amplifier input and output, and drive the pump lasers. Pump light is supplied to the Erbium-doped fiber using discreet WDM combiners, and other components like optical isolators or optical taps to monitor the power level are also present. For each component there is a potential for integration. Some integration steps, like combining the driver and control electronics, are straightforward, whereas for the optical components, the optimum solution is less clear. Some components like monitoring taps or WDM combiners can be realized using planar waveguide circuit technology, which becomes more compelling for multichannel amplifiers, but some other components, like optical isolators, are typically realized in free space or sandwiched between a pair of fiber optic collimators.

For the Erbium-doped fiber itself, numerous options are currently under study, the most obvious being the use of multiple single-core doped fibers (e.g., by using reduced cladding diameter [11] fibers to reduce the size of the amplifier fiber). Also of interest is the use of multi-core fibers, where multiple Erbium-doped cores are present in a single cladding [12]. The advantage of the multi-core amplifier is the possibility to share bulk optical components, for example, optical isolators or a dichroic beam splitter used to separate and combine the pump light from the signal. In the core-pumped multi-core EDFA the pump light is coupled directly to each individual core, and the multichannel amplifier can then be operated similar to multiple conventional EDFAs. In an alternative approach, cladding pumping, the multi-core EDFA is pumped using a multimode pump laser that is coupled to the cladding [4]. The solution is of interest because multimode pump lasers at 980 nm are low-cost and have higher electrical-to-optical efficiency than single-mode pumps. Also, a single pump that can be easily coupled into the multicore fiber is then used to

Optical fiber amplifiers have been instrumental for the success of WDM multiplexed systems. A single amplifier is capable of amplifying a wide frequency range extending over multiple terahertz. This large bandwidth can support 100 or more wavelength channels.

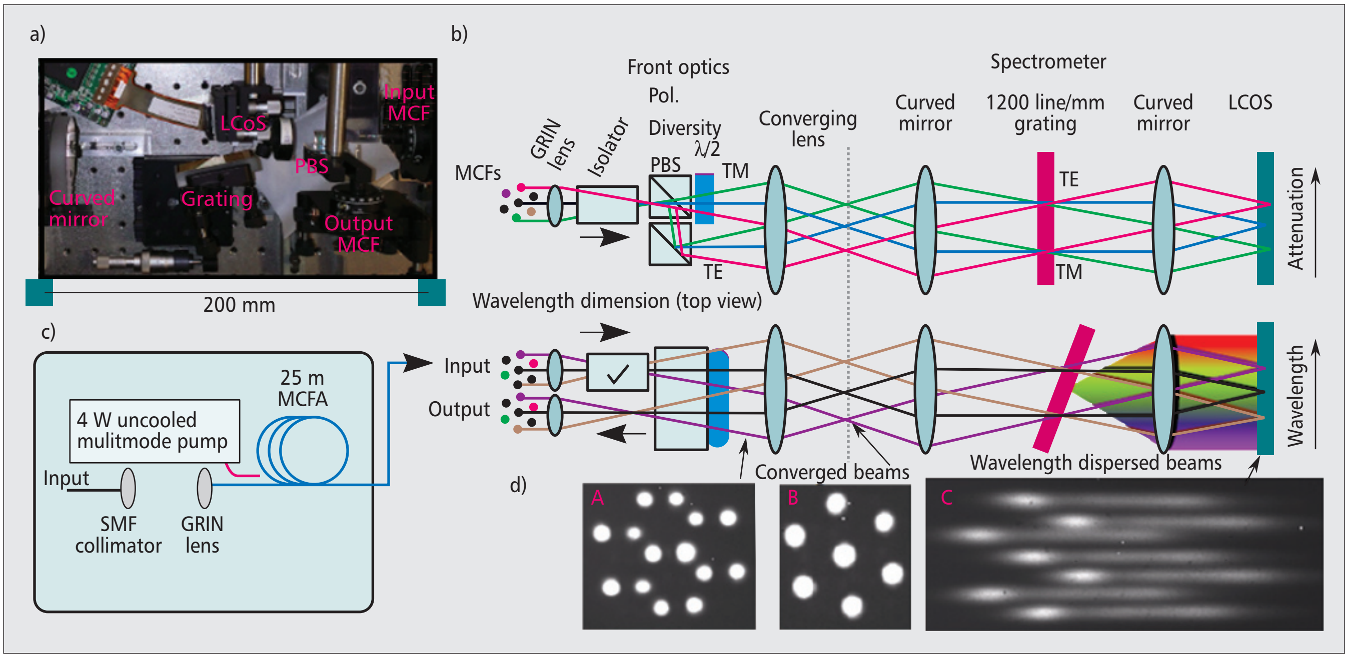


Figure 4. a) Picture of a table-top version of a cladding pumped multicore Erbium doped amplifier with integrated core-wise gridless dynamic gain equalizer; b) schematic optical layout of the core-wise dynamic equalizer (PBS denotes polarizing beam splitter, and λ/2 denotes a birefringent plate with λ/2 retardance); c) cladding pumped multicore fiber amplifier (MCFA); d) images of the beams at different locations when illuminated by C-band amplified spontaneous emission [13].

amplify multiple channels simultaneously, drastically reducing the number of components in the amplifier. The drawback of the approach is that the gain in the individual cores can no longer be controlled. In general, for spatial super-channels all individual signals will have similar power levels and wavelength components, and require the same overall amplification gain. If desired, additional controlling elements like an array of variable attenuators or dynamic gain equalizers can be integrated [13]. Figure 4 shows an example where a single LCoS device is used to individually control the spectrum of a cladding pumped multi-core amplifier. Note also that multi-core amplifiers can be used in combination with multiple single-mode transmission fiber while maintaining the advantages of the multi-core amplifier.

A more radical approach to implement an amplifier that supports multiple spatial channels is based on Erbium-doped few-mode fiber (ED-FMF) [6, 7]. The amplifier looks similar to a conventional EDFA, but multiple signal modes can be amplified at the same time. Because the signal modes have a strong intensity overlap in the fiber, the modes practically share the gain medium. The difficulty of this approach is to minimize the gain variation between modes. For this purpose, fibers with optimized spatial doping profiles have been developed, and additionally, by injecting the power from the pump laser into specific modes, additional control of the modal gain properties can be achieved. Typically, the modes in the amplifier will undergo mode coupling, which will require joint DSP at the receiver to undo the coupling; therefore, this type of amplifier cannot be used as a general-purpose amplifier.

Optical Amplifier Supporting Multiple Spectral Bands

Single-mode fibers present a wide transmission window spanning over 500 nm and therefore offer great potential to increase the capacity by increasing the wavelength range supported by optical amplifiers. Currently, most fiber optical communication systems for long-distance communication are based on a single wavelength band, the most popular being the C-band (1530–1565 nm wavelength range) or the L-band (1565–1625 nm). Both bands are supported by Erbium-doped fibers, where amplification in the L-band is achieved by increasing the length of the doped fiber. Optical amplification in the S-band (1460–1530 nm) has been demonstrated using Erbium-doped fibers based on a depressed cladding design and also in Thulium-doped fibers. Furthermore, Raman amplification has been used often in combination with Erbium-doped fiber amplifiers to extend the wavelength range of amplified communication systems up to 80 nm wavelength range. More recently, Thulium-doped fibers have been used to demonstrate broadband amplification around 2000 nm wavelength, where an almost 200-nm-wide wavelength range was achieved [15].

Multiple amplifiers supporting different bands can be exploited by using low-loss WDM combiner/splitters, available as low-pass or high-pass filters. After the signal is split in bands, components like amplifiers, transponders, and switches specific to the particular band are used. Components that seamlessly support multiple bands without requiring splitting the signal in bands are highly desirable but technically challenging, particularly for optical amplification.

Multiband systems are of interest in fiber-poor environments to avoid the cost of deploying new additional fibers. In fiber-rich environments, it is often more cost effective to use multiple C-band systems, because components are readily available and cost reduced by the economy of scale, and no band splitter/combiners are required. Note also that a typical fiber optic cable can contain hundreds of fibers.

In fiber-rich environments, it is often more cost effective to use multiple C-band systems, because components are readily available and cost reduced by the economy of scale, and no band splitter/combiners are required. Note also that a typical fiber-optic cable can contain hundreds of fibers.

Conclusion

In this article we have reported the latest device-level developments and research activities to increase the transmission and switching capacity in optical networks based on combined wavelength and space multiplexing. The concept of the spatial super-channel, where multiple spatial channels are bundled and treated as a single entity, is expected to be a key driver to develop new high-capacity low-cost-per-bit devices like spectrally flexible multichannel transponders, multi-channel wavelength selective switches, and multi-channel optical amplifiers, as required to interconnect the next generation of data centers.

References

[1] P. J. Winzer, "Spatial Multiplexing: The Next Frontier in Network Capacity Scaling," *ECOC '13*, tutorial paper.

[2] N. K. Fontaine, R. Ryf, and D. T. Neilson, "$N \times M$ Wavelength Selective Crossconnect with Flexible Passbands," *OFC/NFOEC '12*, Postdeadline Paper PDP5B.

[3] L. E. Nelson *et al.*, "Spatial Super-Channel Routing in a Two-Span ROADM System for Space Division Multiplexing," *J. Lightwave Technology*, vol. 32, no. 4, 2014, pp. 783–89.

[4] K. Abedin *et al.*, "Cladding-Pumped Erbium-Doped Multicore Fiber Amplifier," *Opt. Express*, vol. 20, no. 18, 2012, pp. 20,191–200.

[5] Y. Jung *et al.*, "First Demonstration and Detailed Characterization of a Multimode Amplifier for Space Division Multiplexed Transmission Systems," *Opt. Express*, vol. 19, no. 26, 2011, pp. B952–B957.

[6] E. Ip *et al.*, "146λx6x19-Gbaud Wavelength- and Mode-Division Multiplexed Transmission over 10x50-km Spans of Few-Mode Fiber with a Gain-Equalized Few-Mode EDFA," *OFC/NFOEC '13*, Postdeadline Paper PDP5A.2.

[7] G. Le Cocq *et al.*, "Modeling and Characterization of a Few-Mode EDFA Supporting Four Mode Groups for Mode Division Multiplexing," *Opt. Express*, vol. 20, no. 24, 2012, pp. 27,051–61.

[8] S. J. Savory, "Digital Coherent Optical Receivers: Algorithms and Subsystems," *IEEE J. Selected Topics in Quantum Electronics*, vol. 16, no. 5, 2010, pp. 1164–79.

[9] X. Liu, S. Chandrasekhar, and P. J. Winzer, "Digital Signal Processing Techniques Enabling Multi-Tb/s Super-channel Transmission," *IEEE Sig. Proc. Mag.*, vol. 31, no. 2, 2014, pp.16–24.

[10] S. Randel *et al.*, 'Complexity Analysis of Adaptive Frequency-Domain Equalization for MIMO-SDM Transmission," *ECOC '13*, pp. 801–03.

[11] M. Yamada *et al.*, "Optical Fiber Amplifier Employing a Bundle of Reduced Cladding Erbium-Doped Fibers," *IEEE Photonics Technology Letters*, vol. 24, no. 21, 2012, pp. 1910–13.

[12] Y. Tsuchida *et al.*, "Multicore Erbium Doped Fiber Amplifier with 5 THz Bandwidth, High-Gain and Low NF," *Photonics Society Summer Topical Meeting Series*, 2013, pp. 105–06.

[13] N. K. Fontaine *et al.*, "Programmable Gain Equalizer for Multi-Core Fiber Amplifiers," *OFC '14*, Postdealine Paper TH5C.

[14] H. Louchet *et al.*, "Capacity Upgrade of Legacy Fiber Plant Using Space Division Multiplexing and Multimode/Multicore EDFAs," *ECOC '11*, Paper We.10.P1.74.

[15] Z. Li *et al.*, "Diode-Pumped Wideband Thulium-Doped Fiber Amplifiers for Optical Communications in the 1800–2050 nm Window," *Opt. Express*, vol. 21, 2013, pp. 26,450–55.

Biographies

Roland Ryf (roland.ryf@alcatel-lucent.com) is a Distinguished Member of Technical Staff at Bell Laboratories, Alcatel-Lucent, Holmdel, New Jersey. He received his diploma and Ph.D. in physics from the Swiss Federal Institute of Technology (ETH) Zürich, working on nonlinear optics and optical parallel processing. Since joining Bell Labs in May 2000 he has worked on large port-count optical cross-connect switches, high-resolution optical wavelength filters, wavelength-selective switches, and numerous first experimental demonstrations of long distance space-division multiplexed transmission over multimode fibers and coupled-core multicore fibers. He has authored/coauthored over 150 journal and conference publications, and holds over 30 patents.

S. Chandrasekhar [F] received his Ph.D. degree from the University of Bombay, India, in 1985, and joined Bell Labs, Alcatel-Lucent in 1986. He has worked on high-speed optoelectronic integrated circuits (OEICs) and WDM optical networking. He is a Distinguished Member of Technical Staff at Bell Labs, a Fellow of the Optical Society of America, and was an Associate Editor of *Photonics Technology Letters*. He received the IEEE 2000 Engineering Achievement Award and the OSA 2004 Engineering Excellence Award.

Sebastian Randel [SM] received his Dr.-Ing. degree in electrical engineering from TU Berlin, Germany, in 2005 for his work on high-speed OTDM transmission systems. From 2005 to 2010 he was a research scientist at Siemens AG, Munich, Germany, working on optical networks with applications in access, automotive, and industrial automation. In 2010 he joined Bell Labs, Alcatel-Lucent, where he is engaged in research activities on high-capacity optical transport networks. He was a founding chairman of the DKE Working Group Polymer Optical Fiber. He is an editor of *IEEE Transactions on Communications*.

David T. Neilson [F] is a technical manager at Bell Labs, Alcatel-Lucent. He holds B.Sc.(Hons) and Ph.D. degrees in physics from Heriot-Watt University, Scotland. He has been with Bell Laboratories since 1998, where he has conducted research on optical switches and compensators, highly integrated InP based optoelectronic components and subsystems, the role of optical interconnects and switching for high-capacity optical switches and routers, and the energy consumption of networks. Prior to joining Bell Labs he was a visiting scientist at NEC Research Institute, Princeton, New Jersey, researching optical interconnects for high-performance computing. He has over 150 publications and 20 issued patents in the field of optical interconnects, switching and optoelectronic devices. Dr. Neilson is a Senior Member of the Optical Society of America (OSA).

Nicolas K. Fontaine obtained his Ph.D. in 2010 at the University of California, Davis in the Next Generation Network Systems Laboratory [http://sierra.ece.ucdavis.edu] in electrical engineering. In his dissertation he studied how to generate and measure the amplitude and phase of broadband optical waveforms in many narrowband spectral slices. Since June 2011, he has been a member of the technical staff at Bell Laboratories, Crawford Hill, New Jersey, in the advanced photonics division. At Bell Labs, he develops devices for space-division multiplexing in multi-core and few mode fibers, builds wavelength crossconnects and filtering devices, and investigates spectral slice coherent receivers for THz bandwidth waveform measurement.

Mark D. Feuer is currently an associate professor at the College of Staten Island, City University of New York, where his research focuses on transforming society through advanced optical and communications technologies. His industrial research experience at AT&T Labs, JDS Uniphase, and Bell Labs spanned photonic devices, systems, and metrology, including invention of the reflective semiconductor optical amplifier (RSOA) and pioneering work on reconfigurable photonic networks.

Switching Solutions for WDM-SDM Optical Networks

Dan M. Marom and Miri Blau

The authors are with Hebrew University.

Abstract

Over the last few decades, network traffic has consistently grown at an exponential rate and was efficiently satisfied using WDM and more efficient coding schemes requiring coherent detection. There is no indication that the network traffic growth trend will cease anytime soon, and we are nearing the day when the capacity of the ubiquitous single-mode fiber will be fully exploited. Space-domain multiplexing (SDM) for high-capacity transmission is the promising solution with the scaling potential to meet future capacity demands. However, there is still a large technological gap between current WDM optical communication system designs and SDM network implementations. In this article we lay the foundation of switching node designs for future WDM-SDM optical networks.

Introduction to WDM-SDM Optical Networks

Ubiquitous information access available to us today in multifaceted forms (laptops, tablets, smart phones, and wearable devices such as watches and glasses) has enabled us to enjoy unfettered data exchange between individuals, businesses, the cloud, and is coming soon to an appliance near you. The pervasive infrastructure supporting our information-based society is founded on optical networks traversing the globe, whether through undersea links across the oceans or in terrestrial networks spanning continents, countries, and cities, and terminating at points of business and private premises. The technology behind these optical networks has evolved considerably over the past 40 years, with each technology cycle offering an orders of magnitude capacity growth at a typically incremental cost basis. As such, we have witnessed sustained exponential network capacity growth for decades, and have reaped its benefits. However, at present we are seeing warning signs on the distant horizon that question how we can continue to evolve the infrastructure and offer continued exponential growth opportunities on an economically viable basis [1].

Optical networks are predominantly based on the omnipresent single-mode optical fiber, which, alongside conventional Erbium-doped fiber amplifiers offers essentially 5 THz of useable bandwidth for information transmission. The available bandwidth can be efficiently exploited using wavelength-division multiplexing (WDM), allowing multiple frequency-separated channels to co-transmit, and using modulation formats that encode information over amplitude, phase, and polarization at each wavelength channel. However, when trying to encode more bits of information on every communicated symbol for even more aggressive modulation formats, a transmission distance limitation emerges due to nonlinear impairments of the fiber medium itself. Hence, higher capacities tend to lose transmission range. This implies that future higher-capacity optical transport networks might need to be constructed using either progressively aggressive modulation formats over limited ranges and then cascading networks to enable extended reach, or operating using today's modulation formats that have extended reach but lower spectral efficiency and increasing capacity by deploying additional optical networks in parallel. Barring any other development, either scenario implies that supporting future exponential capacity growth rates will be accompanied by exponentially rising costs as networks are reaching their capacity-distance limits due to the fundamental physical attributes of single-mode fibers and Erbium-based optical amplifiers. One plausible avenue of research is investigating alternative fiber dopants for enabling optical amplification in additional frequency bands to provide some relief of the capacity constraint [2], but this solution will not provide more than an order of magnitude capacity increase. The alternate research avenue is investigating different optical fiber designs that offer additional guided modes for optical communication. These new fibers are heralding a new era of space-division multiplexing (SDM), which can more significantly scale in spatial mode counts and hence capacity [3].

Optical fibers in support of SDM transmission may come in many forms (Fig. 1). Single-mode fiber (SMF) is designed to allow light guiding of a single spatial mode in the core region by tailoring the refractive index profile

 0163-6804/15/$25.00

and core dimensions. Multiple cores can be placed within a single fiber cladding, forming a multi-core fiber (MCF), with each core now supporting a single spatial mode. Hence, MCF may offer a capacity multiplier equal to the core count. Alternatively, the core dimensions or refractive index contrast can be modified to support additional optically guided spatial modes. These few-mode fibers (FMFs) may offer a capacity multiplier equal to the mode count. A somewhat more exotic fiber design is that of the annular core fiber (ACF), which supports multiple spatial modes confined to the annular core region [4]. The annular structure is designed to support a single radial mode and multiple azimuthal modes. One of the key differentiating metrics between SDM supporting fibers is whether the modes remain *uncoupled* in transmission or potentially may be *coupled* due to manufacturing imperfections as well as environmental effects such as bends, stress, and temperature gradients. Coupled transmission implies that the modes intermix, but the information is maintained within the set of modes. However, coupled transmission does require that these mixed modes must remain together in order to unravel the mixing at the receiver using digital signal processing. FMF and ACF are inherently prone to mixing as the modes spatially overlap and hence are categorized as a coupled SDM transmission medium. Even though in a simple MCF the cores are distinct, they may still couple if the cores are closely packed through evanescent fields; conversely, the MCF can be specifically designed to suppress coupling to very low values and remain essentially uncoupled. One flavor of SDM purports to use an array of existing SMF (or several SMF co-packaged [5], Fig. 1a); obviously, this SDM solution can be categorized as uncoupled because the fibers are separated. The SDM fiber can also be designed to support multiple coupled spatial mode subgroups, but have no coupling between the groups. Two such examples are an FMF-MCF hybrid, where there are uncoupled multiple higher index cores and each core supports several spatial modes [6], and an MCF design with uneven spacing allowing only the closely packed cores to couple [7, 8].

Transitioning to an optical network based on high-capacity SDM fiber links is still far off and entails resolving many issues related to its implementation, such as identifying the best SDM fiber options, the optical amplification means, and efficient space multiplexing and demultiplexing methods. This article focuses on the all-optical switching operations occurring at network nodes, where information-bearing channels have to be routed toward their destinations within an optical mesh topology. Such switching can occur for the individual wavelength and space channels, or possibly performed jointly over all modes or all wavelengths, and other mixed solutions in an effort to simplify the implementation. These options are weighed in terms of their physical realization requirements, as well as routing flexibility and scaling potential. Integration is key for the economic viability of the SDM solution, and the cost of implementation of the WDM-SDM network switching nodes will be a decisive factor

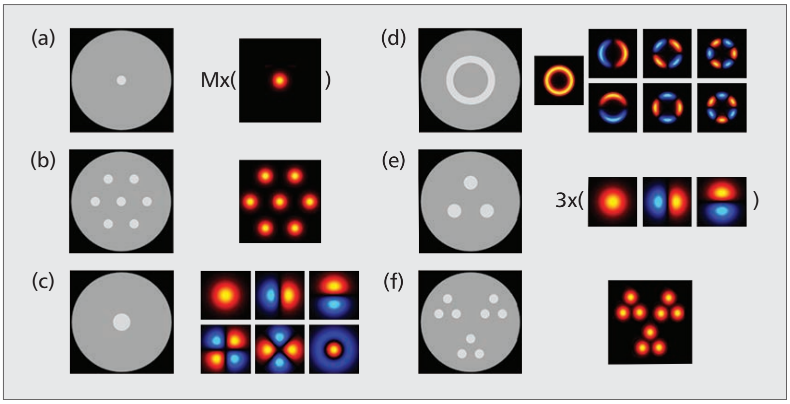


Figure 1. Different types of fibers in support of SDM, showing their geometrical form and propagating spatial mode distributions: a) single-mode fiber array; b) multi-core fiber; c) few-mode fiber; d) annular-core fiber; e) multi-core supporting few modes; f) multi-core arranged in coupled subgroups.

to be reckoned with. While this article lays the foundation for switching node designs of future WDM-SDM optical networks, their complete assessment and merit identification will likely require intense research effort by the optical networking community to identify the trade-offs and costs in advance of deployment and uptake of such networks.

WDM-SDM Switching Classification

The adoption of new SDM-supporting fibers in the optical network potentially increases the capacity per fiber by a factor of M, the number of guided spatial modes. Considering that each fiber mode still spectrally spans the optical communication band, WDM can be applied, carrying N wavelength channels per spatial mode. Hence, the WDM-SDM fiber capacity can be defined by a two-dimensional spatial-spectral array, with wavelengths ($\lambda 1, \ldots, \lambda N$) and spatial modes ($\sigma 1, \ldots, \sigma M$) defining its columns and rows (Fig. 2). Before addressing the switching options available in a WDM-SDM optical network, it is first instrumental to consider the current state of switching in optical networks and emerging trends.

Today's optical networks employ wavelength-selective switches (WSSs) at network nodes to perform the switching, and are exclusively based on SMF interfaces [9]. A conventional WSS accepts a single fiber carrying WDM traffic and switches each wavelength channel to one of K possible output fiber ports. Note that each output fiber port can have any number of wavelength channels switched to it, as the switching is independently performed for each wavelength in the input port. The switch can also operate in reverse, receiving K input fibers with WDM traffic, and selecting on a wavelength channel basis which input fiber will be interconnected to the single output fiber port. Two identical wavelengths cannot be physically switched to the single output fiber in a WSS; such a network routing assignment is prohibited, and referred to as wavelength contention. In an optical mesh

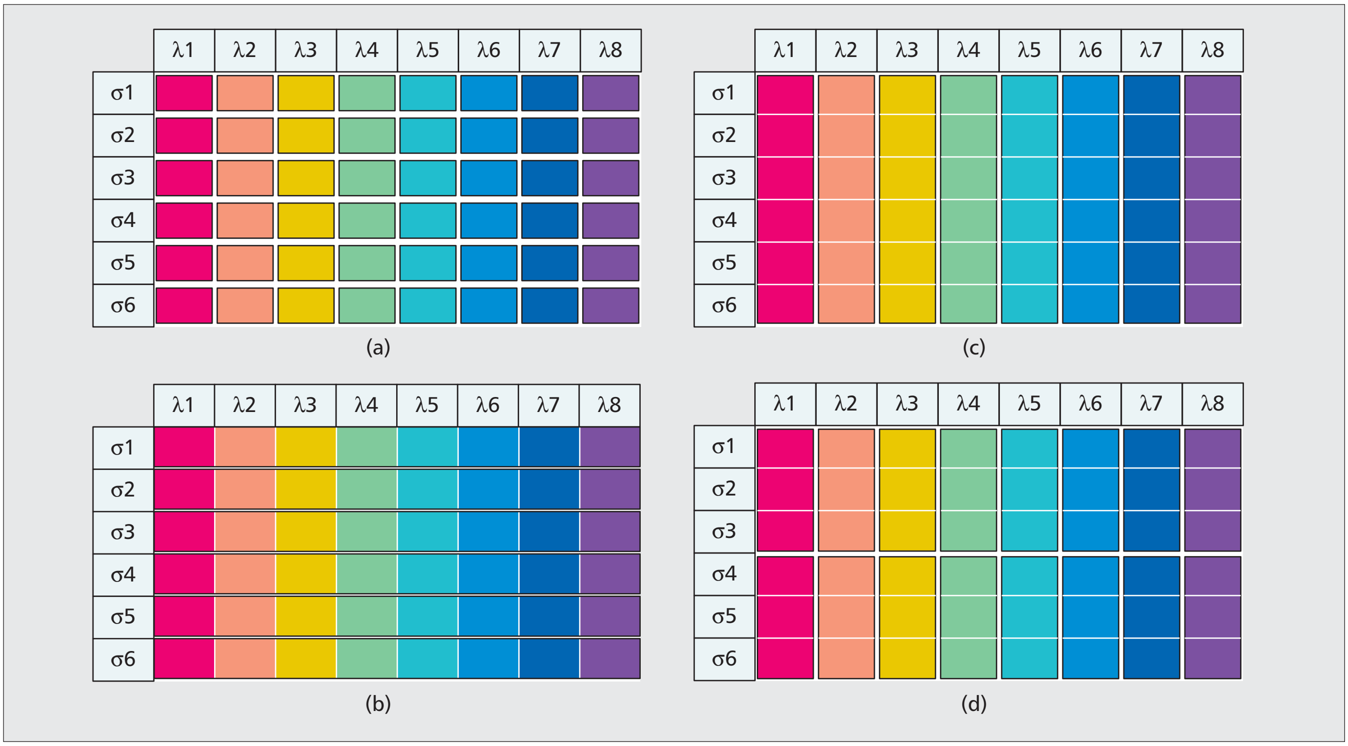


Figure 2. Parsing the fiber's SDM and WDM channels for switching, where space modes (σ1, …, σ*M*) and wavelength channels (λ1, …, λ*N*) are fully utilized: a) space-wavelength granularity (each mode/wavelength channel can be independently switched); b) space granularity (switching performed on a mode basis across all wavelengths); c) wavelength granularity (switching performed on a wavelength basis across all modes); d) fractional space-full wavelength granularity (switching performed on a wavelength basis and spatial mode sub-groups).

network node, a broadcast-and-select or route-and-select architecture can perform the switching function, with the route function performed with a 1 × *K* WSS, and the select function with a *K* × 1 WSS [10]. The wavelength channels of a WSS are defined by the spectral switching elements, and typically reside on a 100 GHz (supporting about 50 channels) or 50 GHz grid (supporting nearly 100 channels). The fixed WSS channel plan imparts a certain inefficiency, as guard bands are introduced between adjacent channels that reduce the bandwidth available for information encoding. Today, as data volumes grow and SMF capacity is becoming a limiting factor, flexible-grid WSSs are being introduced. In a flexible-grid WSS the spectral selection does not adhere to a fixed grid and can be provisioned as required; large contiguous spectral channel bandwidths can be defined [11]. This capability is being utilized to create spectral super-channels, which consist of closely packed subchannels that are routed through the network as one entity with few or no guard bands between the subchannels. For example, a 200 GHz wide spectral super-channel can support 1 Tb/s capacity by combining several subchannels on separate carriers. Hence, employing super-channels for transport, the number of WDM channels tends to decrease while their assigned bandwidth increases. In SDM research experiments we are witnessing a reverse trend. Initial SDM experiments were conducted with few spatial modes (e.g., three) and are gradually increasing. Hence, the WDM-SDM capacity array might have, in certain scenarios, a column count of ~100 × 50 GHz wide wavelength channels or maybe only ~25 × 200 GHz spectral super-channels (or even mixed assignments); the spatial mode row count may be only a few (e.g., 6–10) modes initially, but can certainly scale to tens to 100 in the future. These expected values are important to consider when weighing the different WDM-SDM switching technologies.

While there are many SDM fiber alternatives one can consider for implementing future WDM-SDM optical networks, for mesh node switching purposes we categorize them into three general archetypes:

Uncoupled spatial modes: Spatial channels remain distinct in fiber propagation and all ancillary network equipment, as would be experienced in uncoupled MCF (Fig. 1b) or an SMF bundle (Fig. 1a). Therefore, individual spatial modes can be switched from one SDM fiber link to another, or add/drop operations applied for any spatial mode/wavelength channel combination.

Coupled spatial modes: Spatial channels mix throughout fiber transmission, as occurs in FMF (Figs. 1c, 1d) and coupled MCF (Fig. 1b), forming spatial superchannels in which multiple data streams are transported as groups of subchannels occupying the same wavelength in independent modes/cores. As a result of the channel mixing, multiple-input multiple-output (MIMO) processing is required in order to unravel the mixed information, which occurs after coherent detection of all the spatial modes at the SDM receiver. Since the information is mixed across all spatial modes, modes cannot be separated for switching to other destinations, or information loss will be experienced. A complete MIMO receiver has to be employed in order to separate the SDM data, an operation performed at the channel destination and not desirable at every mesh network node as network transparency will be lost.

Coupled spatial subgroups: Spatial modes may mix only within subgroups of the total spatial mode count. The subgroups are defined by the SDM fiber design, and the spatial modes belonging to a subgroup must not be separated in switching operation (Figs. 1e, 1f). Subdividing the spatial modes into subgroups of smaller size eases the switching limitations with respect to full mode coupling, while not reaching the full flexibility of uncoupled modes.

At the optical network nodes, the WDM-SDM traffic on each inbound fiber link has to be either redirected to outbound fiber links as part of the network information flow, or dropped to receivers for local consumption at the node's geographical location. Additional data is typically reintroduced, or added, in its place, originating at clients associated with the node. The SDM fiber categories dictate the permissible switching operations that can transpire at the optical network nodes. We identify four alternative switching scenarios that strongly correlate to the SDM fiber categories, forming multi-dimensional switching nodes of different granularities (Fig. 2).

Independent spatial mode/wavelength channel switching (space-wavelength granularity): The WDM-SDM fiber capacity can be switched independently for every spatial mode and wavelength combination (Fig. 2a). This forms the finest switching capacity granularity, leading to the greatest flexibility at the cost of increased realization complexity. Employing independent mode/wavelength switching requires uncoupled SDM fiber.

Spatial mode switching across all wavelength channels (space granularity): The WDM-SDM fiber capacity is switched at the spatial mode level, independent of wavelengths (Fig. 2b). Hence, the entire communication band per mode is jointly switched. At low spatial mode counts, space switching granularity is coarse (all WDM channels) but simple to realize. Employing independent mode switching also requires uncoupled SDM fiber.

Wavelength switching across all spatial modes (wavelength granularity): The WDM-SDM fiber capacity is switched at the wavelength level across all spatial modes, forming spatial superchannels that are routed through the network as one entity (Fig. 2c). Since the spatial modes are not separately routed, the network topology is similar to today's SMF networks, but benefitting from the SDM capacity multiplier. Employing wavelength switching across all modes is obligatory for coupled SDM fiber, but can also be applied to uncoupled SDM fibers.

Wavelength switching across spatial mode subgroups (fractional space-full wavelength granularity): The WDM-SDM fiber capacity is switched at the wavelength level across smaller spatial mode subgroups (Fig. 2a). The switching operation is still in support of spatial superchannels within each subgroup, but applied independently to the subgroup elements. Employing the fractional space and full wavelength granularity capacity switching supports coupled subgroups of SDM fiber, but can also be applied to uncoupled SDM fibers.

Routing in the WDM-SDM optical network is constrained by the employed switched capacity granularity, as the network provisioning algorithms must assign each information flow request onto a route that can be supported by the switching nodes and is contention-free. An additional degree of freedom implicit in the switching capacity granularity involving the wavelength space (options A, C, and D), is the ability to flexibly define the switched spectrum. The independent space-wavelength granularity (option A) is the smallest capacity block size and offers the greatest routing flexibility, as even single wavelength and spatial mode requests can be accommodated. The alternative solutions (options B–D) utilize larger switching capacity granularities by addressing all wavelengths (B), spatial modes (C), or spatial mode subgroups (D) as one entity. Such switching solutions may become inefficient when addressing small capacity requests, but the reduced hardware required to realize these degenerate switching solutions may be favorable implementation-wise. If using coupled SDM fiber, jointly switching all spatial modes (option C) is mandatory. The capacity granularity can be reduced by provisioning narrower spectral bands by the switching hardware, which may require some customization. For example, if the switching hardware can support minimal bandwidth provisioning of 35 GHz, a six spatial mode SDM fiber can offer enough capacity to support 1 Tb/s as the minimal switched granularity, which is a reasonable starting point for future WDM-SDM optical networks.

WDM-SDM Switching Solutions

After having defined the switched capacity granularity at the WDM-SDM optical network nodes, we turn our attention to their implementation details. The switching node must complete two functions: routing traffic from an input SDM fiber to an output SDM fiber and performing channel add/drop to be terminated at optical transceivers. Each solution entails its unique switching hardware, and various levels of complexity are associated with each granularity level. However, some elements are recurring, and we briefly explain their operation. The WSS has been introduced earlier, and its flexible-grid implementation is assumed here. The WSS utilizes SMF at its input/output ports (although WSSs with direct FMF interfaces have been demonstrated [9]), and must be properly interfaced to the SDM fiber solutions. For MCF, a breakout device separates the M cores to M individual SMFs. For FMF, a mode demultiplexer converts the M modes to M individual SMFs [10, 11]. This operation does not necessarily require the modes to be mapped to individual output fibers; a unitary mode-mixing operation may be associated with the demultiplexer, which can be subsequently undone in MIMO processing at the receiver.

The switching node must complete two functions: routing traffic from an input SDM fiber to an output SDM fiber and performing channel add/drop to be terminated at optical transceivers. Each solution entails its unique switching hardware, and various levels of complexity are associated with each granularity level.

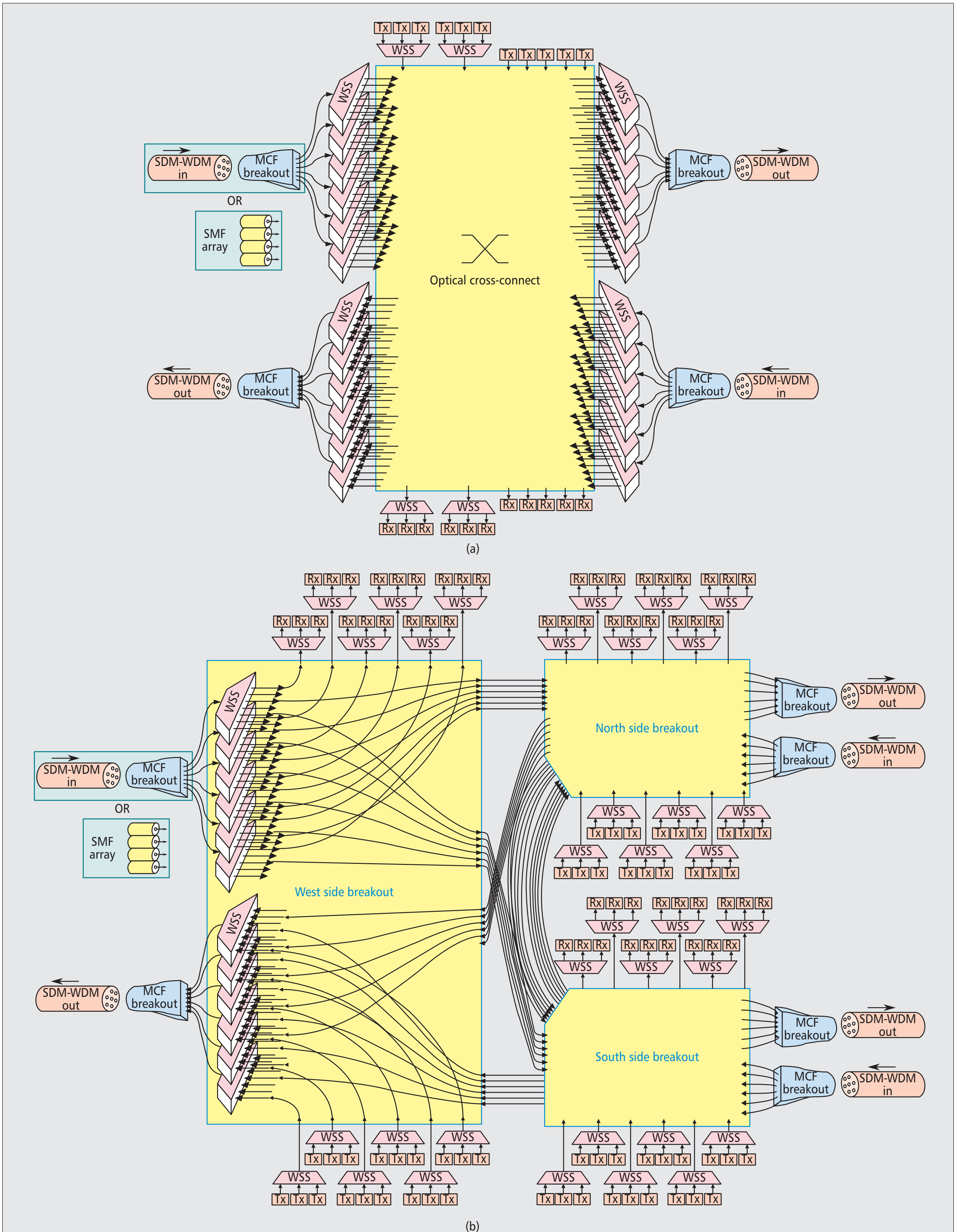


Figure 3. Implementations for independent spatial mode/wavelength channel switching: a) optical cross-connect for full connectivity (only two fiber links shown); b) route-and-select per each spatial mode. North and south side breakouts are identical to the west side breakout.

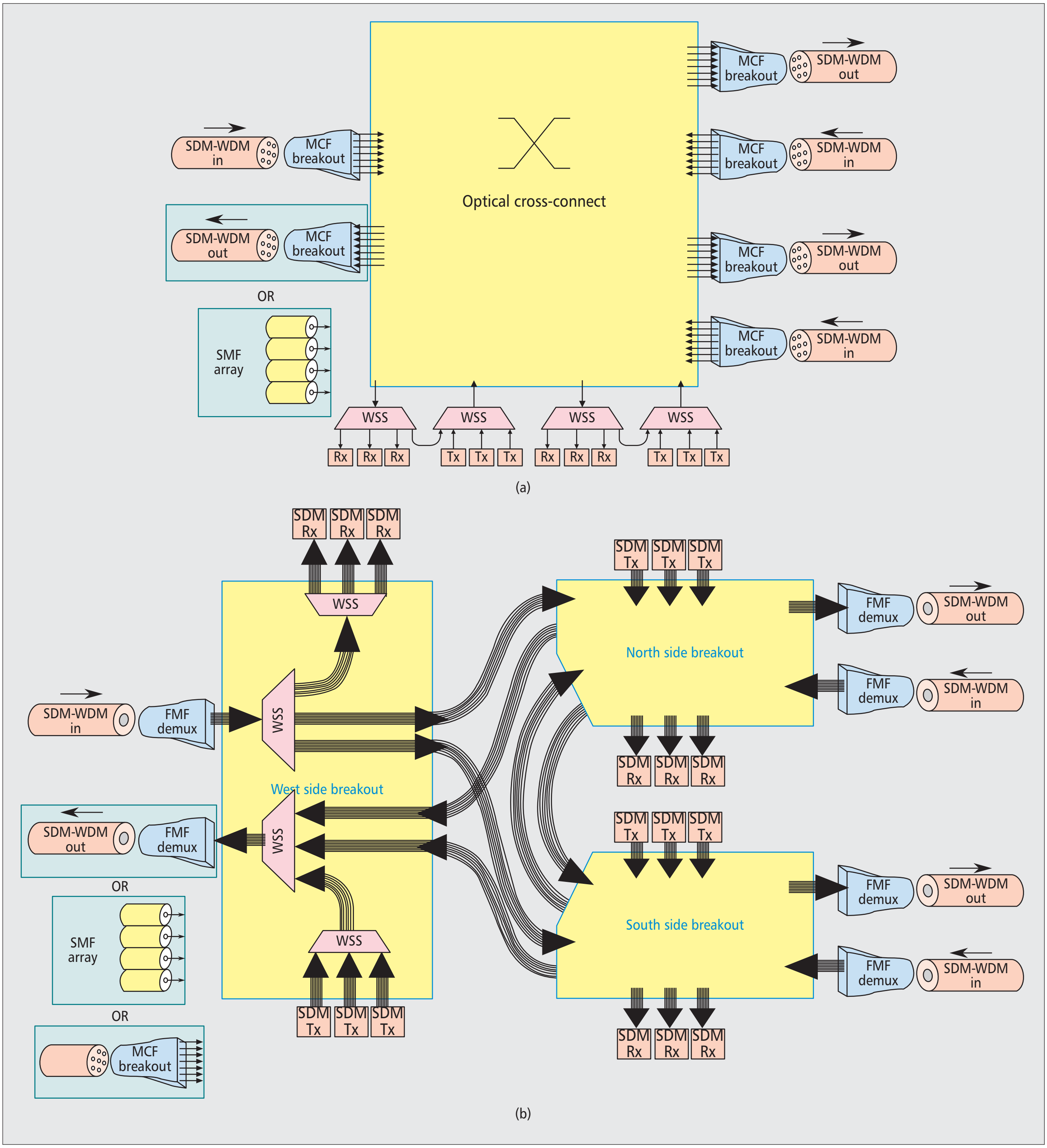


Figure 4. Switching node designs for implementing: a) space granularity, routing the entire communication band per spatial mode; b) wavelength granularity, routing all modes per wavelength.

Two possible implementations of independent spatial mode/wavelength channel switching (space-wavelength granularity) offering different levels of flexibility are shown in Fig. 3. The first implementation makes use of a large optical cross-connect (OXC). To interface between the uncoupled SDM fiber solution and the OXC, each independent spatial mode is preprocessed with a $1 \times K$ WSS. The WSS subdivides the WDM channels on each spatial mode according to destination, whether to an output SDM fiber (on a particular spatial mode) or to a drop port. On the output side a $K \times 1$ WSS multiplexes or post-processes the channels onto a spatial mode of the SDM fiber. Single drop channels can be terminated directly at conventional receivers, and multiple drop channels can be further separated with another WSS. This architecture provides full routing flexibility thanks to the OXC, especially wavelength contention by enabling

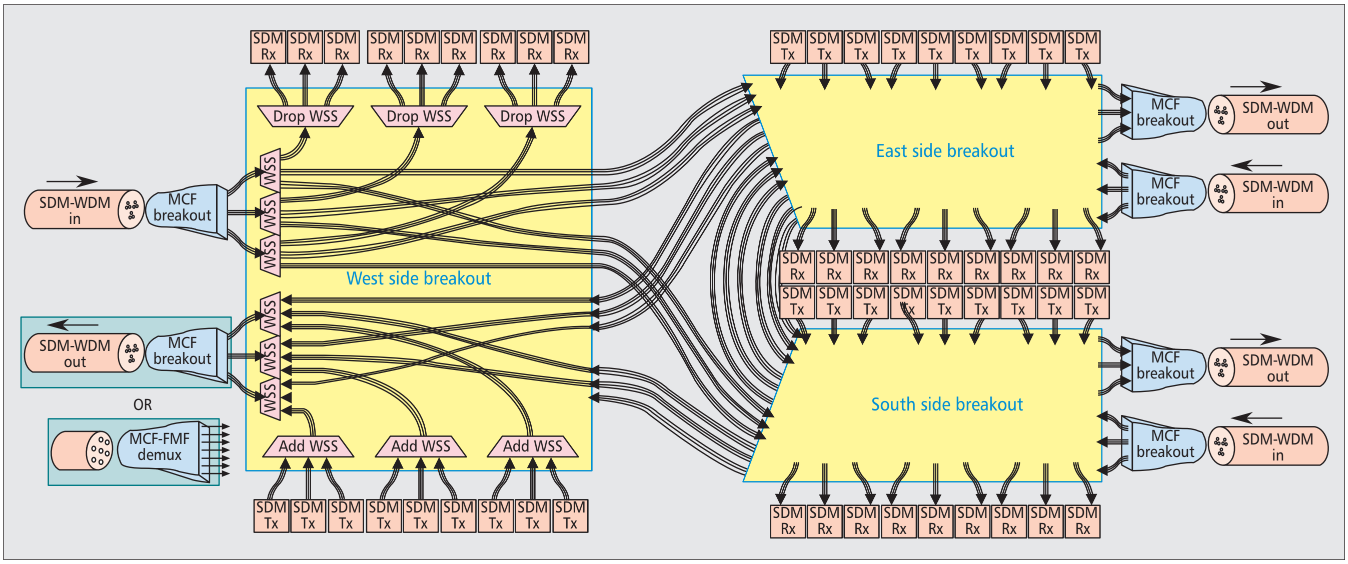


Figure 5. Switching node design for implementing hybrid fractional space-full wavelength switching granularity, routing spatial super-channels spanning spatial subgroups.

SDM "lane changes" or routing of a wavelength channel from one spatial channel to another instead of the much harder wavelength conversion solutions proposed for SMF networks. Additionally, the transceiver elements are accessible to all fiber directions (and are further colorless and contentionless), resolving the directional limitation in route-and-select solutions. Having WSS pre- and post-process WDM channels enables flexible bandwidth allocation, and also conserves OXC fiber ports as WDM channels destined to the same direction and able to be routed jointly.

The alternative independent spatial mode/wavelength switching solution eliminates the OXC, which is a costly element and a single point of failure threat. Here, route-and-select switching is performed for each spatial mode independently. Each independent spatial mode is subdivided by a $1 \times K$ WSS and routed to output fiber destinations (mapped to the same mode on the output fiber, eliminating SDM lane change operation), or to a set of receivers associated with the spatial mode (i.e., directional in both the fiber and mode senses). Hence, eliminating the OXC results in routing constraints.

The biggest disadvantage of the independent space-wavelength granular switching node designs is the amount of hardware required to implement them. Essentially, this entails scaling the WSS count M-fold, the same as the capacity increase. Hence, the mesh node cost scales linearly with the capacity gain, which is counter to the value proposition of SDM. We are seeking a sublinear cost increase with capacity to maintain network economics by way of better device and sub-system integration. With independent space-wavelength granular switching this requirement is not met.

Eliminating the WDM switching elements and realizing a space-granular capacity switching design significantly reduces the node hardware and cost (Fig. 4a). An OXC receives all spatial modes along its input ports and switches each mode to an output SDM fiber, thereby completing the routing of the entire optical communication band (all WDM channels carried on a spatial mode) to the output destinations. For the add/drop operation, the OXC can switch the dropped spatial mode to a port where a WSS separates the channels to be detected, and the remaining channels are fed through to a second WSS that combines these channels with additional channels added back into the OXC for output fiber assignment. The WSS count in the node depends on the number of spatial modes allowed to be dropped, with a minimum of one spatial mode per fiber. However, as this number increases, which is required to offer reasonable routing flexibility in a multi-node network, the WSS count can again become prohibitively high.

Alternatively, the wavelength-granular capacity switching design utilizes a recently introduced WSS modification specifically designed for routing of spatial super-channels [12, 13]. The WSS is based on a conventional (SMF-based) high-port-count WSS, and is made to operate with all the spatial modes of the input SDM fiber feeding a first subset of the WSS ports. The internal wavelength switching mechanism of the WSS (based on beam steering) steers the set of input ports onto a second subset of the WSS ports. When the ports are arranged in a linear equispaced array, the fiber ports are imaged from the first subset onto the second, switching in parallel all the fibers and hence the entire spatial mode set. This joint-switching WSS can be used to construct the conventional route-and-select architecture of SMF networks, with the M-fold parallelism applied across all modes with a single switching module. The routed spatial super-channels traverse a first $M \times (1 \times K)$ for destination selection and a second $M \times (K \times 1)$ for combining the wavelength channels to the output SDM fiber. The dropped spatial super-channels are interfaced to SDM transceiver elements where MIMO processing is performed for information extraction in the case where the modes

	Space-wavelength granularity	Space granularity	Wavelength granularity	Fractional space-full wavelength granularity
Minimum switching granularity	Bandwidth of a single WDM channel present at a single spatial mode.	Bandwidth of entire optical communication band carried on a single spatial mode.	Bandwidth of a single WDM channel spanning over all spatial modes.	Bandwidth of a single WDM channel over a subset of spatial modes.
Realization	*With OXC*: High-port count OXC and at least 2*M* conventional WSSs per I/O fiber link. *Without OXC*: 2*M* conventional WSSs per I/O fiber link. 4*M* if WSSs placed on add/drop.	Moderate port count OXC, and 2 WSSs per mode selected for WDM channel add/drop.	4 joint switching WSSs per I/O fiber link in route-and-select topology applied to all spatial modes in parallel.	4 × *M*/*P* joint switching WSS modules per I/O fiber link.
Flexibility	*With OXC*: Each mode/WDM channel independent provisioned and routed. Supports SDM lane change. Single point of failure. *Without OXC*: Each mode/WDM channel independently provisioned and routed. Spatial mode maintained. Prone to wavelength contention.	The complete optical communication band is routed across network. Coarse granularity. If WDM channels need to be extracted from many modes, WSS count quickly escalates.	Each spatial super-channel provisioned across all modes. Susceptible to wavelength contention. Add/drop bound to direction.	Compromise solution using small SDM groups. More efficient when provisioning low-capacity demands.
Scaling	*With OXC*: Can quickly escalate to very large port counts. Switching node cost linearly scales with capacity, no price benefit to SDM.	Conventional OXC can support foreseeable mode and fiber counts. OXC is a single point of failure. Pricing favorable but with greater add/drop require more WSS modules.	Cost roughly independent of SDM count. Inefficient for low-capacity connections due to minimum BW provisioned across SDM. Large SDM Rx/Tx are an integration and DSP challenge.	Cost scales as group count. Groups can be turned on as capacity grows, offering a pay-as-you-go alternative. Maintaining small group sizes facilitates MIMO processing at Rx.

Table 1. Comparison of WDM-SDM switching alternatives.

are mixed due to mode coupling from the SDM fiber. The scalability of this wavelength-granular solution to high mode counts (tens of modes) is presently undetermined, as joint switching WSSs have limited fiber port counts.

The final variant is a space and wavelength hybrid design, which offers fractional space and full wavelength granular capacity switching (Fig. 5). This switching scenario is matched to an SDM fiber solution offering M spatial modes, where the modes can be divided to M/P independent subgroups, where modes are coupled within the individual subgroups having P modes. Each mode group must be switched jointly due to the inherent intra-coupling, but the groups can be switched independent of one another. The switching solution is the basic route-and-select topology applied to each subgroup using modified WSSs that support the joint-switching concept for groups of size P. The solution is replicated M/P times, matching the subgroup count. This hybrid solution offers finer granularity than switching all modes (wavelength granularity) at a price of increased switching hardware; however, the cost is a fraction ($1/P$) of the capacity gain. While the fractional space-full wavelength solution requires specific forms of SDM fiber, it can also be applied to uncoupled SDM fiber. This switching solution can more effectively address SDM fibers with very high spatial mode counts, provided they are designed according to coupling within grouped modes.

Concluding Remarks on WDM-SDM Optical Networking

SDM transmission is a promising solution to the capacity limitation of SMF, but addressing the physical SDM elements of novel fiber types, supporting optical amplifiers, and mode multiplexers, without careful attention to optical networking implications misses an important element of the entire value proposition. In this article we highlight some of the implications of designing a WDM-SDM optical mesh network, concentrating on the switching node designs by which information flows need to be provisioned. We identify four categories of capacity granularities to be provisioned, applied across the space and wavelength domains. Each category can be realized with different optical switching gear at the network node, affecting the realization complexity and cost, flexibility, and scalability. These findings are summarized in Table 1.

It is premature at this early stage to deduce if there is an optimal solution to the WDM-SDM optical network. This has to be assessed for spe-

It is premature at this early stage to deduce if there is an optimal solution to the WDM-SDM optical network. This has to be assessed for specific network layouts and traffic patterns. Different networking applications are likely to have divergent conclusions.

cific network layouts and traffic patterns. Different networking applications are likely to have divergent conclusions. Assessment of the complete optical network must take into account the physical layer attributes, the expected information flow scales and churn, how efficiently they can be met given the minimum capacity granularity that is routed by the network, blocking probabilities due to contention for the provisioning of information flows, and cost of implementation, amongst other. WDM-SDM mesh node switching solutions, the focus of this article and a key factor in network operation and cost, can be assessed with a routing power metric, borrowed over from reconfigurable wavelength add-drop node designs [14, 15]. The routing power metric can provide a measure of the number of connection states for the switching node, and will decrease for larger switching granularity, but this loss of flexibility is favorably accompanied by lower implementation costs. As such, a complete analysis involves the contributions from different skillsets and it will likely require the concerted effort of many researchers in the field to analyze the performance level and benefits offered by WDM-SDM optical networks.

Acknowledgments

The research leading to these results has received funding from the European Community's Seventh Framework Program (FP7/2007-2013) under grant agreement no. 619732 (INSPACE).

References

[1] A. D. Ellis, J. Zhao, and D. Cotter, "Approaching the Non-Linear Shannon Limit," *J. Lightwave Tech.*, vol. 28, 2010, pp. 423–33.

[2] M. A. Melkumov *et al.*, "Laser Diode Pumped Bismuth-Doped Optical Fibre Amplifier for 1430 nm Band," *Optics Letters*, vol. 36, 2011, pp. 2408–10.

[3] D. J. Richardson, J. M. Fini, and L. E. Nelson, "Space-Division Multiplexing in Optical Fibres," *Nat'.. Photonics*, vol. 7, no. 5, 2013, pp. 354–62.

[4] N. Bozinovic *et al.*, "Terabit-Scale Orbital Angular Momentum Mode Division Multiplexing in Fibers," *Science*, vol. 340, 2013, pp. 1545–48.

[5] S. Jain *et al.*, "Multi-Element Fiber Technology for Space-Division Multiplexing Applications," *Opt. Express*, vol. 22, 2014, pp. 3787–96.

[6] Y. Sasaki *et al.*, "Large-Effective-Area Uncoupled Few-Mode Multi-Core Fiber," *Opt. Express*, vol. 20, no. 26, 2012, pp. B77–84.

[7] Y. Kokubun, M. Koshiba, and T. Morioka, "Coupled System Multi-Core Fiber, Coupling Mode Multiplexer and Demultiplexer, System for Tranmission Using Multi-Core Fiber and Method for Transmission Using Multi-Core Fiber," US Patent Application 20110188855 A1, Aug. 4, 2011.

[8] R. Ryf *et al.*, " Space-Division Multiplexed Transmission Over 3 × 3 Coupled-Core Multicore Fiber," *OFC*, *OSA Tech. Digest* (online), 2014, paper Tu2J.4.

[9] D. M. Marom *et al.*, "Wavelength-Selective 1?K Switches using Free-Space Optics and MEMS Micromirrors: Theory, Design, and Implementation," *IEEE J. Lightwave Tech.*, vol. 23, 2005, pp. 1620–30.

[10] S. Gringeri *et al.*, "Flexible Architectures for Optical Transport Nodes and Networks," *IEEE Commun. Mag.*, vol. 48, no. 7, 2010, pp. 40–50.

[11] S. Frisken, S. B. Poole, and G. W. Baxter, "Wavelength-Selective Reconfiguration in Transparent Agile Optical Networks," *Proc. IEEE*, vol. 100, May 2012, pp. 1056–64.

[12] R. Ryf *et al.*, "Wavelength-Selective Switch for Few-Mode Fibers Transmission," *ECOC 2013* PD1.C.2.

[13] J. A. Carpenter *et al.*, "1 × 11 Few-Mode Fiber Wavelength Selective Switch using Photonic Lanterns" *OFC*, *OSA Tech. Digest* (online), 2014, paper Th4A.2.

[14] N. K. Fontaine *et al.*, "Few-Mode Fiber Wavelength Selective Switch with Spatial-Diversity and Reduced-Steering Angle," *OFC*, *OSA Tech. Digest* (online), 2014, paper Th4A.7.

[15] L.E. Nelson *et al.*, "Spatial Superchannel Routing in A Two-Span ROADM System for Space Division Multiplexing," *J. Lightwave Tech.*, vol. 32, no. 4, 2014, pp. 783–89.

Biographies

DAN M. MAROM [SM] obtained his B.Sc. degree in mechanical engineering and M.Sc. degree in electrical engineering, both from Tel-Aviv University, Israel, in 1989 and 1995, respectively, and was awarded a Ph.D. in electrical engineering from the University of California, San Diego in 2000. From 2000 until 2005, he was a member of technical staff at the Advanced Photonics Research Department of Bell Laboratories, Lucent Technologies, where he invented and headed the research and development effort on MEMS-based wavelength-selective switching solutions for optical networks. Since 2005, he has been with the Applied Physics Department, Hebrew University, Israel, where he is now an associate professor leading a research group pursuing his research interests in creating photonic devices and subsystems for switching and manipulating optical signals, in guided-wave and free-space optics solutions using light modulating devices, nonlinear optics, and compound materials. He has published more than 150 conference and journal papers, and produced nearly 30 granted patents. He was awarded the 2014–15 IEEE Photonics Society Distinguished Lecturer Award.

MIRI BLAU (miri.blau@mail.huji.ac.il) is presently a Ph.D. candidate student in the Department of Applied Physics at Hebrew University under the supervision of Prof. Marom. She received her B.Sc. degree in electro optics engineering from Jerusalem College of Technology, Israel, in 2010. Her research deals with space division multiplexing for overcoming the capacity crunch accompanied by lower implementation costs than single-mode-based optical communications.

Spectrally and Spatially Flexible Optical Network Planning and Operations

Dimitrios Klonidis, Filippo Cugini, Ori Gerstel, Masahiko Jinno, Victor Lopez, Eleni Palkopoulou, Motoyoshi Sekiya, Domenico Siracusa, Gilles Thouénon, and Christophe Betoule

Dimitrios Klonidis is with Athens Information Technology.

Filippo Cugini is with CNIT.

Ori Gerstel is with Sedona Systems.

Masahiko Jinno is with Kagawa University.

Victor Lopez is with Telefónica.

Eleni Palkopoulou is with Cisco.

Motoyoshi Sekiya is with Fujitsu Laboratories of America.

Domenico Siracusa is with CREATE-NET.

Gilles Thouénon and Christophe Betoule are with Orange Labs.

ABSTRACT

The advent of spectrally flexible (a.k.a. elastic) optical networking is widely identified as the next generation optical network solution that permits varying bandwidth demands to be dynamically assigned over flexible spectral containers, targeting optimum use of the available network resources. Additionally, the adoption of the space dimension is identified as a promising solution for the capacity expansion of future networks, while novel spatial-spectral switching solutions show that the flexible networking concept can be further expanded over both the spatial and spectral dimensions. This article provides an overview of the latest developments and possible approaches with respect to flexible optical networking and the emerging benefits that spatially flexible networking approaches can offer. The focus is on the network planning and resource optimization functions, the main network operations related to fragmentation and IP/optical layer integration, and the control plane solutions.

INTRODUCTION

The scale of network capacity — but also of energy consumption and general economic viability — is pressuring the industry to identify innovative technologies and network architectures in order to overcome the expected fiber capacity limitations. This issue is exacerbated by the new traffic profile, which is mainly dominated by rich content video and cloud services, imposing high traffic churn (i.e., high peak-to-average traffic ratio). In turn, this results in large bandwidth and capacity variations over time, considerably increasing the required capacity overprovisioning far beyond the average network capacity.

The concept of flexible (a.k.a. elastic) optical networking refers to the ability of a network to dynamically adjust its resources, such as the optical bandwidth and modulation format, according to the bandwidth requirements and transmission characteristics of each connection. The key benefit of this approach is that the network resources are always adapted to traffic demands and optimized resource utilization is maintained, thus avoiding the need for overprovisioning.

Spectrally flexible optical networking [1] combines bandwidth-efficient and software defined transmission schemes with advanced network planning and resource optimization methodologies, enabled by spectrally adaptive switching schemes at the nodes. An extension of the concept in the space domain [2, 3] is envisioned, in which the "spatial resources" can be flexibly assigned to different traffic demands, increasing the utilized degrees of flexibility, and the network planning and optimization capabilities of the network. Such spatial resources may refer to fiber cores or modes in multi-core fibers (MCFs) or multi-mode fibers (MMFs), or even single-mode fiber (SMF) bundles, in a more short-term and practical approach. Obviously, as additional degrees of freedom are introduced (i.e., wavelength, signal bandwidth, and space), the networking complexity increases, and efficient control plane solutions are required.

This article first defines the different channel allocation options over the available fiber dimensions (wavelength, bandwidth, space) enabled by the latest relevant technology advances and research efforts. Next, the possible flexible networking approaches considering both spectral and spatial network flexibility are commented on and discussed, identifying their benefits, limitations, and synergies. The last part focuses on the key network design, planning, operation, and control issues, discussing the latest advances in flexible optical networking, while also highlighting the important research directions and required innovations for the introduction of flexibility in the space dimension.

CHANNEL ALLOCATION SCHEMES AND ENABLING TECHNOLOGIES

Spectrally efficient formats [1] have evolved toward spectrally overlapped orthogonal frequency-division multiplexing (OFDM) schemes and time overlapped Nyquist wave-

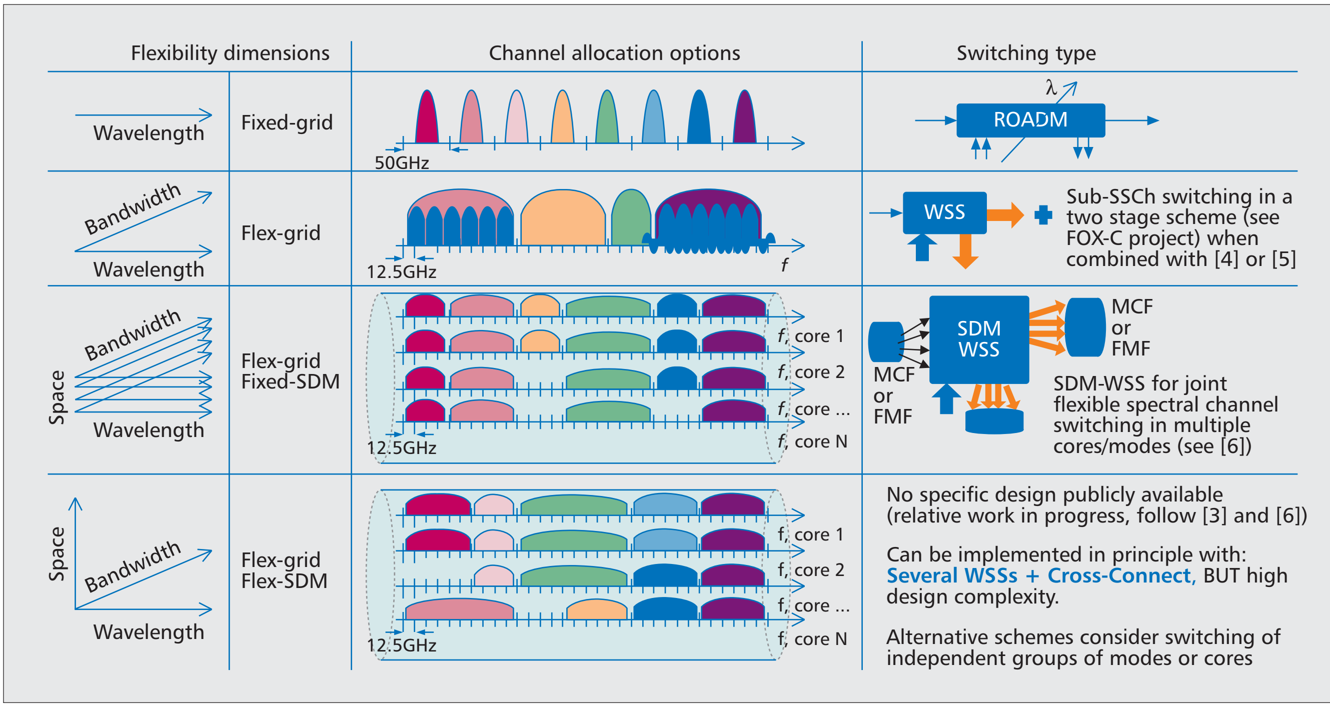


Figure 1. Different channel allocation options according to their degrees of freedom and the related enabling switching solutions. (Note: in the channel allocation options, each color indicates an allocated channel/SCh).

length-division multiplexing (NWDM) schemes, each being implemented in either the optical or electronic domain. The key characteristic of all these schemes is the bit rate and bandwidth adaptability of the transceivers [1], achieved by varying the number of subcarriers and the modulation levels (i.e., the format). Such bandwidth variable transceivers (BVTs) enable the definition of the spectral super-channel (S-SCh) as a networking entity with variable spectral content able to transport a flexible capacity demand over an end-to-end path in the network. By expanding this concept in the space dimension, one may consider multiplexing several S-SChs over a number of cores or modes in MCFs or MMFs, respectively, or even over SMF bundles in a multi-fiber cable. In turn, this defines the spatial-spectral super-channel (S^2-SCh) networking entity in which the channel allocation flexibility spans over both the spectrum and space dimensions.

While the BVT defines the basic optical layer transport entity in flexible networks, the routing of the flexible demands in the network depends purely on the type of bandwidth-adaptive switches and their capabilities. Wavelength-selective switches (WSSs) are the key network elements with elastic bandwidth characteristics that can be dynamically adapted to the S-SCh bandwidth. Although commercially available solutions today offer a spectral resolution close to 8 GHz, the latest developments show that the construction of WSSs at sub-1 GHz resolution is possible [4]. This allows the reduction of spectral guard bands between neighboring S-SCh, increasing the spectral efficiency of the links. Quite recently, an important breakthrough in all-optical switching was presented in [5], demonstrating the capability to add/drop and erase the contents of an S-SCh even in the case of spectrally overlapped OFDM formats. Such solutions are the enablers of all-optical grooming in intermediate nodes, providing additional degrees of freedom for network resource optimization, which remain to be explored. With respect to spatially flexible networking, a significant WSS redesign presented in [6] expanded the spatial dimension of a WSS element, providing switching of spectral contents jointly from all cores in MCFs. Researchers are currently exploring the extension of this switch design in order to enable the spatial switching of spectral contents from a variable number of cores in MCFs, resulting in truly flexible S^2-SCh allocation over both dimensions.

The technology advances, in terms of both BVTs and switching elements, enable the definition of different channel allocation options and in turn different networking approaches (discussed in the next section). The channel allocation options are summarized in Fig. 1 according to the maturity of the relevant technology. First, the traditional fixed-grid WDM approach is shown with tunability in the wavelength dimension only. Next, the flex-grid approach is presented, which spans over both the wavelength and bandwidth dimensions, allowing the allocation of spectrally elastic S-SChs with different data rates and formats. The third case in Fig. 1 considers the use of the space dimension only for multiplexing purposes, increasing the overall capacity per fiber. In this case spectrally flexible S-SChs expand over all or some of the spatial fiber resources (i.e., modes or cores) and are jointly switched in the nodes. The real use of all three dimen-

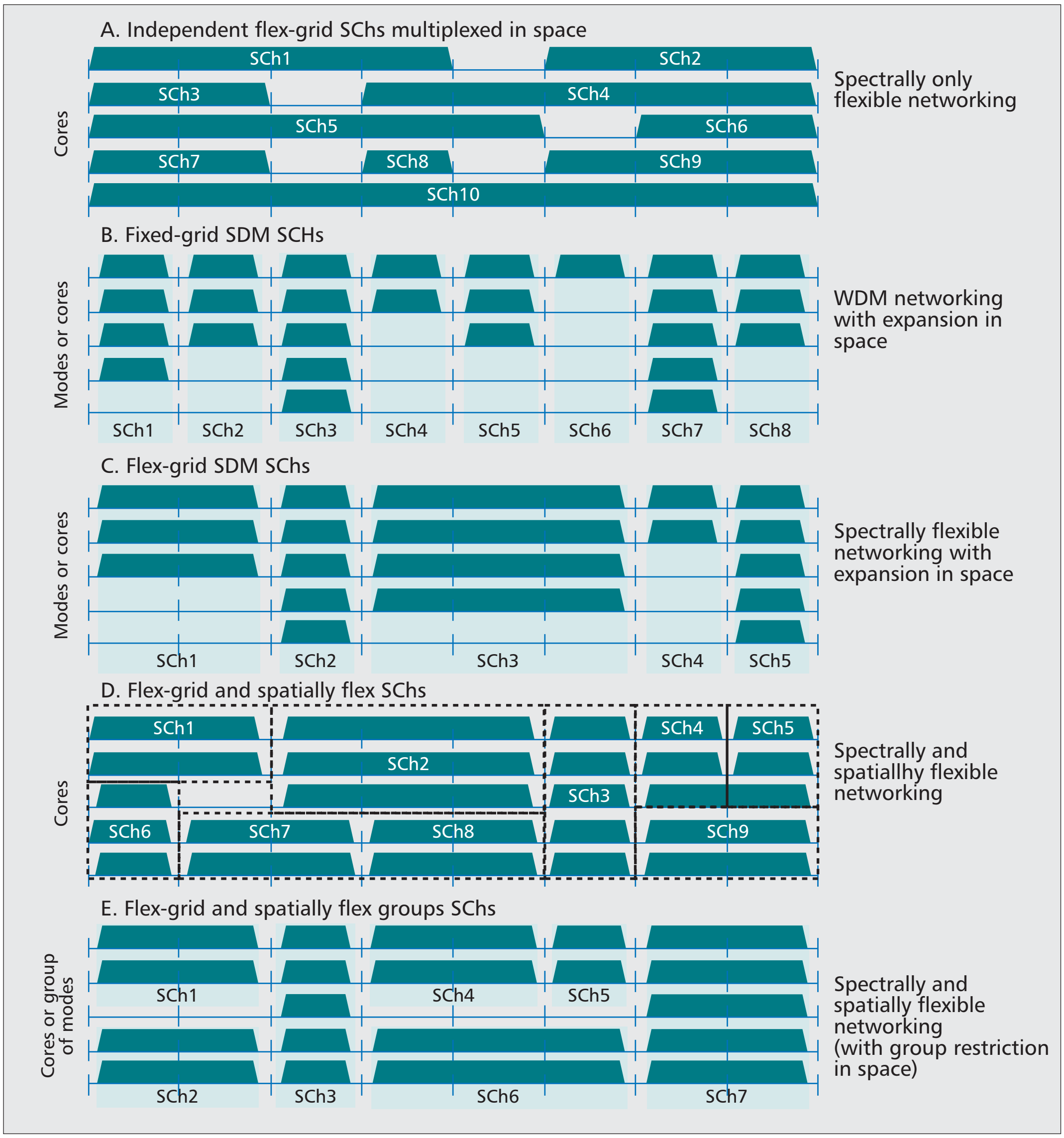


Figure 2. The different flexible optical networking options offered with the use of the space dimension in addition to the spectrum dimension.

The channel allocation flexibility increases significantly with the introduction of the space dimension. However, the key question from the networking point of view is: What added value could the space dimension bring to the flexible optical networking concept?

sions (wavelength, bandwidth, and space) is presented in the last allocation scheme in which the spectral and spatial fiber resources are used independently to allocate the traffic demands over S^2-SChs. For example, an S^2-SCh may be defined over a number of fiber cores, while in the remaining cores another S^2-SCh may be allocated, even with different spectral contents.

Despite the evident multi-dimensional flexibility offered by the latter scheme, the available switching technology solutions are currently in their design phase, and significant research effort is still required for the first proof of concept. Beyond the potential capabilities that the technology innovations could bring, what is of equal importance is to investigate the new networking capabilities that emerge, particularly with respect to the increased flexibility the space dimension could offer. The possible added value of spatial flexibility to the spectrally flexible networking concept is discussed in the following section.

Spatial, Spectral, or Combined Flexible Optical Networking?

The use of the space dimension in fibers is introduced primarily in order to offer the ability of space-division multiplexing (SDM) in addition to wavelength division multiplexing (WDM) to increase the overall transmission capacity in a cost-effective manner [3] (i.e., by integrating to a certain extent multiple transmission systems in parallel). Also, as discussed above, the channel allocation flexibility increases significantly with the introduction of the space dimension. However, the key question from the networking point of view is: What added value could the space dimension bring to the flexible optical networking concept?

The use of the space dimension in addition to the spectrum dimension could potentially lead to five different flexible optical networking options, which are summarized in Fig. 2 and explained in the following paragraphs.

The pure SDM approach (as defined in the

The additional use of the space dimension is expected to significantly increase the complexity of the network planning and resource optimization algorithms compared to the spectrally flexible networking approach, but can potentially lead to significant benefits.

physical layer) means that at the end of each transmission link a spatial demultiplexer is placed prior to a spectral demultiplexer to process and extract the spatially multiplexed data. The spectrally flexible networking concept studied for common SMF-based systems can be equally applied to SDM systems, assuming that a spatial demultiplexing process applies at the end of the links (Fig. 2, case A). In this case, the space dimension is related to the multiplexing stage of the physical layer and has no role in networking layers. Therefore, in principle there is no difference with the spectrally flexible networking concept, which spans over multiple parallel links. Evidently, this scheme represents the natural evolution of today's transmission systems toward higher-capacity systems, considering the use of links with bundles of SMF. However, it is also applicable to MCF systems with uncoupled cores in order to avoid cross-talk between cores carrying independently assigned end-to-end S-SChs.

An alternative could be to use elasticity in the space dimension instead of the spectrum dimension (Fig. 2, case B). Channels can be allocated over a fixed spectral grid (as in the case of WDM systems) and may have the ability to flexibly expand over some or all of the modes/cores in the fiber. For example, a 100 Gb/s dual polarized quadrature phase shift keying (DP-QPSK) demand can be allocated inside a 50 GHz grid and occupy one core in an MCF, while a 400 Gb/s demand is allocated over four cores inside the same grid. Networking can then rely on common WDM allocation and routing schemes. Also, the transmitter and receiver designs are simplified due to the use of a common transmission and local oscillator laser. In this case, the spatial elasticity makes the spectral elasticity requirements obsolete. Moreover, a pure spatially flexible optical network maintains the network planning and operation simplicity of fixed-grid WDM networks, avoiding issues like spectral fragmentation and complex resource optimization and control, as in the case of spectrally flexible networks. However, this networking solution is not able to make real use of the capacity increase benefits of SDM. Limited resource utilization is expected due to the required spectral gaps between adjacent fixed-grid channels and mainly to the unutilized spatial resources from low capacity demand. The scheme is applicable to any type of MMF or MCF with coupled or uncoupled cores.

A more efficient scheme than the pure spatially flexible networking concept can be proposed by allowing spectrally flexible contents to be allocated over the space dimension (Fig. 2, case C). This approach leads to the flex-grid SDM scheme and is primarily enabled by the novel joint spatial switching elements proposed in [6]. From the networking point of view the network planning and operation complexity follows that of spectrally flexible networks with the main difference being the allocation options for the SChs. Similar to a spectrally flexible approach, the scheme also allows the adaptation of the signal format (i.e., bandwidth) to the transmission reach target, which can be an important issue for SDM transmission due to cross-channel interference in MMFs and MCFs with coupled cores. The flex-grid SDM solution is expected to have better resource utilization than the pure spatially flexible (i.e., fixed-grid SDM) solution for large capacity demands due to efficient use of the spectral resources; however, optimum use of the spatial resources still remains an issue.

Evidently, the most flexible networking approach is to combine elasticity in both the spectrum and space dimensions (Fig. 2, case D). In this case, the available end-to-end resources could be identified over a pool of spatial and spectral slots in the links that form the end-to-end path, assuming that innovative multi-dimensional nodes are in place, able to perform switching in both dimensions independently. The additional use of the space dimension is expected to significantly increase the complexity of the network planning and resource optimization algorithms compared to the spectrally flexible networking approach, but can potentially lead to significant benefits. Spectral fragmentation issues, discussed in the next section, can be resolved via switching of the fragmented spectral slots in the space dimension, while new approaches for network resource virtualization and content-aware IP mapping over the optical layer can also be adopted with the spatial allocation of virtualized network segments. The scheme requires the use of MCFs with uncoupled cores to achieve independent switching of cores in the space dimension.

Finally, an alternative scheme for combined spectral and spatial networking could consider the allocation of SChs on independent groups of cores or modes rather than individual ones (Fig. 2, case E). This case can be seen as a combination of cases C and D presented above. The spatial groups can be treated as a single switching entity whether or not they are fully utilized, as in case D, while spectral elasticity is maintained within them, as in case C. This approach could potentially relax the technology related design complexity of the node (i.e., by reducing the number of required spatial ports), but at the expense of some resource underutilization compared to the previous case. The scheme is applicable to MCFs with uncoupled cores and possibly to MCFs with coupled cores assuming that the spatial groups are properly allocated over distant cores with minimum interference between them. Moreover, this scheme could be especially attractive for the case of the newly developed MCF with few-mode cores (FM-MCF) [7], where each of the few-mode cores is a spatial group that can be handled independent of the rest of the fiber cores.

It is noted that the combined spectral and spatial flexible networking is a new research topic, and significant research effort is required in order to identify and quantify its full potential and complexity limitations. Moreover, a detailed evaluation of the different flexibility schemes presented above is currently missing from the literature. In the following two sections the ongoing progress in important fields of flexible networking and network control is reviewed, and the required extensions with respect to spatial and spectral flexibility are identified.

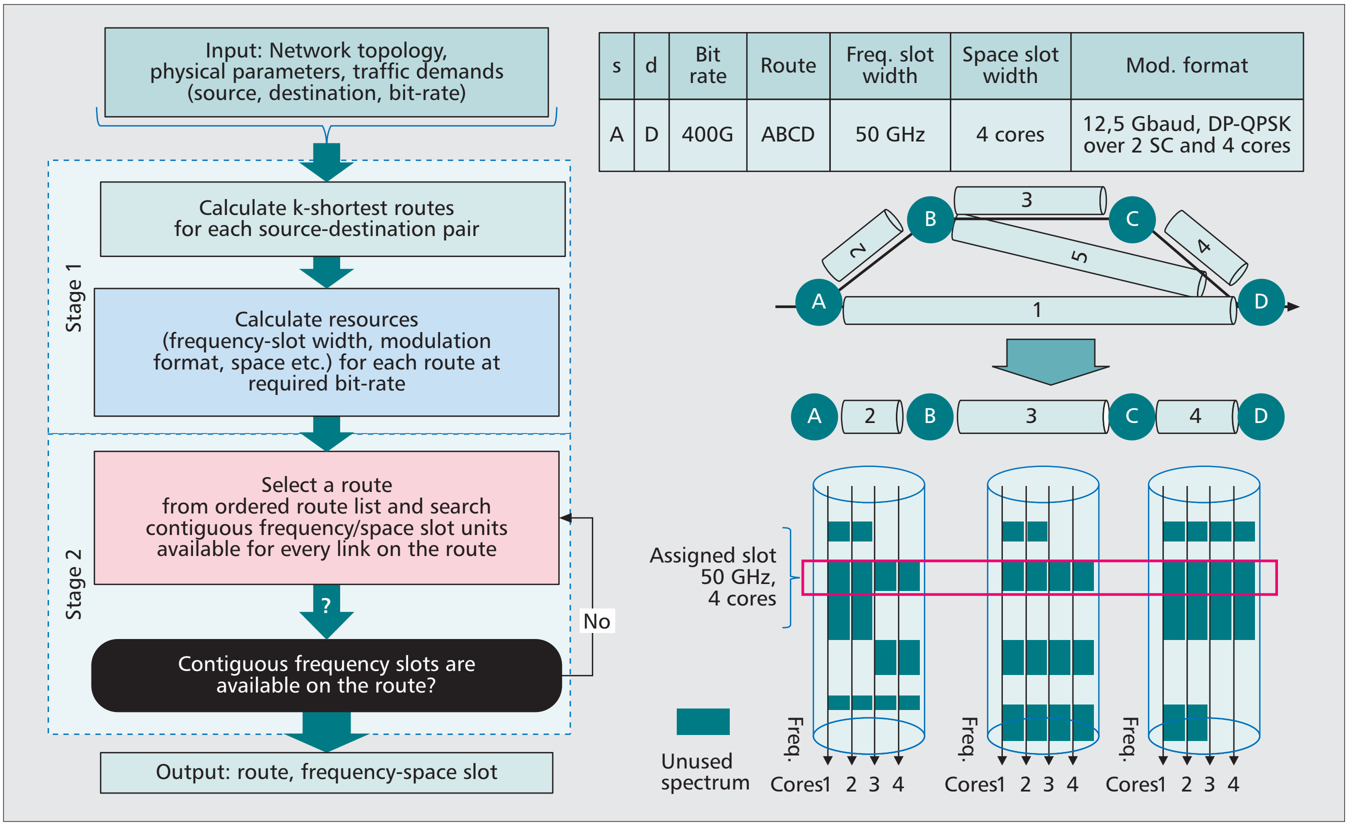


Figure 3. An example of the process and the main calculation steps for identifying and assigning a capacity demand in a flexible network.

Network Planning and Operation Challenges

Network Planning and Resource Optimization

Network planning enables the assignment of network resources (transponders, switching elements, spectrum, fiber cores/modes, etc.) to connections in order to satisfy a defined optimization objective while adhering to a given set of constraints. In flexible optical networks, network planning is considered to be a dynamic online process that adapts consistently to the changes of the connection demands obeying specific optimization criteria, such as minimization of the required spectral resources, and capital and/or operational expenditures, and maximization of energy efficiency. The capabilities offered by the network equipment impose a set of constraints on the feasible assignments. For example, the granularity to which the spectrum of an optical channel can be increased is restricted by the capabilities of the optical switches deployed along its path.

Traditional WDM optical networks require routing and wavelength assignment (RWA) algorithms that ensure the wavelength continuity of connections along the optical path and the single wavelength assignment on each link. With the introduction of spectrally flexible optical networking, it was discovered that conventional RWA algorithms are no longer applicable, since a number of contiguous spectrum slots (instead of single wavelengths) must be assigned. Additionally, the continuity of these spectrum slots should be guaranteed in a similar manner as wavelength continuity constraints are imposed. This led to the development of routing and spectrum allocation (R-SA) algorithms, which evolved to routing, modulation level, and spectrum allocation (R-ML-SA) algorithms, covering all the degrees of freedom allowed by spectrally flexible optical networks and their relevant new constraints. The advent of spatial flexibility introduces an additional degree of freedom, the space dimension, thus leading to the definition of routing, space, modulation level, and spectrum allocation (R-S-ML-SA) algorithms. Moreover, new constraints related to the types of fibers and the spatial interference among copropagating S^2-SChs must be included.

All the network planning and resource optimization algorithmic approaches generally follow two stages, as shown in the example of Fig. 3. In the first stage, a route list with necessary resources (frequency slot width, modulation format, number of cores, etc.) is created. Given the network topology and physical parameters, an ordered list of *k*-shortest routes for each source-destination pair is first created. Then the required resources (spectral and/or spatial) are calculated for each route at the required bit rate, taking into account linear and nonlinear impairment factors or, more simply, based on expected and pre-calculated transmission reach values. The second stage allocates contiguous resources to the route. When a connection request arrives, a route is selected from the ordered route list created in the previous stage in sequence. The

The advent of spatial flexibility introduces an additional degree of freedom, the space dimension, thus leading to the definition of routing, space, modulation level, and spectrum allocation (R-S-ML-SA) algorithms. Moreover, new constraints related to the types of fibers and the spatial interference among copropagating S^2-SChs must be included.

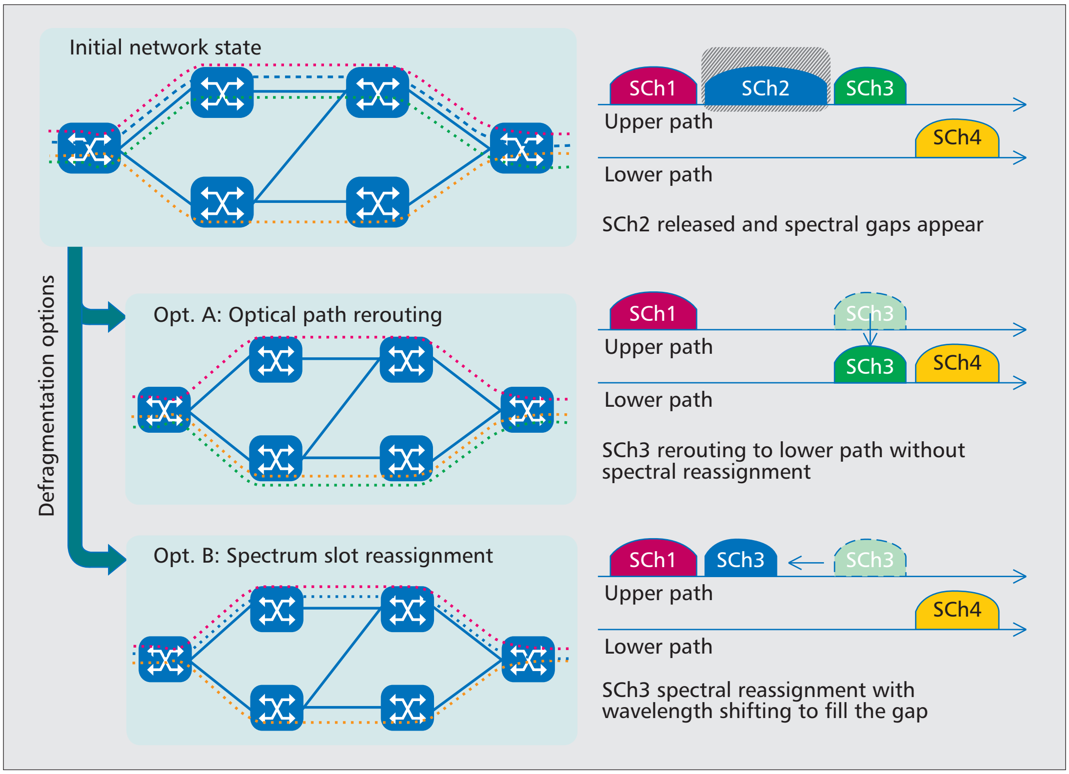


Figure 4. Network example describing the two defragmentation options of optical path rerouting and spectrum slot reassignment when one SCh (SCh2-dashed line) is released.

necessary resources are found from the list. Then a search is performed for the available contiguous resources in every link on the route, and the fewest available contiguous resources are selected. If no available contiguous resources are found on the route, an alternate route is selected from the route list.

The RWA and RMLSA algorithmic approaches have been extensively examined over the last years [8]. However, the network planning approaches that also consider the space dimension are quite limited. The resource allocation for flex-grid SDM networks with MCF was addressed for a dynamic scenario in [9]. Also, recently a joint spectral and spatial network planning strategy for networks employing MCFs was presented [10] taking into account inter-core cross-talk impairments.

Fragmentation

Spectrum fragmentation is an intrinsic issue in spectrally flexible optical networks where multi-rate and multi-modulation-format signals are allocated various number of spectrum slots [11]. As signals, occupying different number of slots, are established and released over time, spectrum gaps (fragmented slots) inevitably appear, making it difficult to allocate a certain number of contiguous slots to new signals. For example, a released 400 Gb/s 16-quadrature amplitude modulation (QAM) connection typically leaves seven empty slots of 12.5 GHz. If four of these slots are re-occupied by a 100 Gb/s DP-QPSK signal, it results in three fragmented slots.

The fragmentation problem is addressed by either trying to alleviate fragmentation or, more effectively, performing a defragmentation process. The alleviation of fragmentation is generally based on enhanced RMLSAs, which define the optimum spectral position of new requests taking into account:

- The already established SChs
- The number of unutilized slots between SChs supporting a possible SCh spectral expansion
- The selection criteria of the network path that will accommodate a demand with the smallest probability of fragmented spectrum

On the other hand, the defragmentation process improves the spectrum utilization by taking active measures on the established traffic (e.g., rerouting and/or reallocating of spectrum resources), with the goal of reducing the unusable spectrum gaps. There are two options to perform defragmentation:

- Reroute optical paths with or without spectrum slot reassignment.
- Reassign spectrum slots maintaining the same optical path routing.

These are described in the example of Fig. 4. In the first option, the optical path is switched to an alternative route to create vacant slots with more contiguous spectrum on some congested links. This approach entails service disruption due to switching of paths. In the second option, spectrum reassignment can be performed by continuously shifting the wavelength and then tracking it using coherent receivers, thus avoiding service disruption and traffic delay since the route and path length is not changed. However, their implementation requires the use of advanced laser tuning and digital signal process-

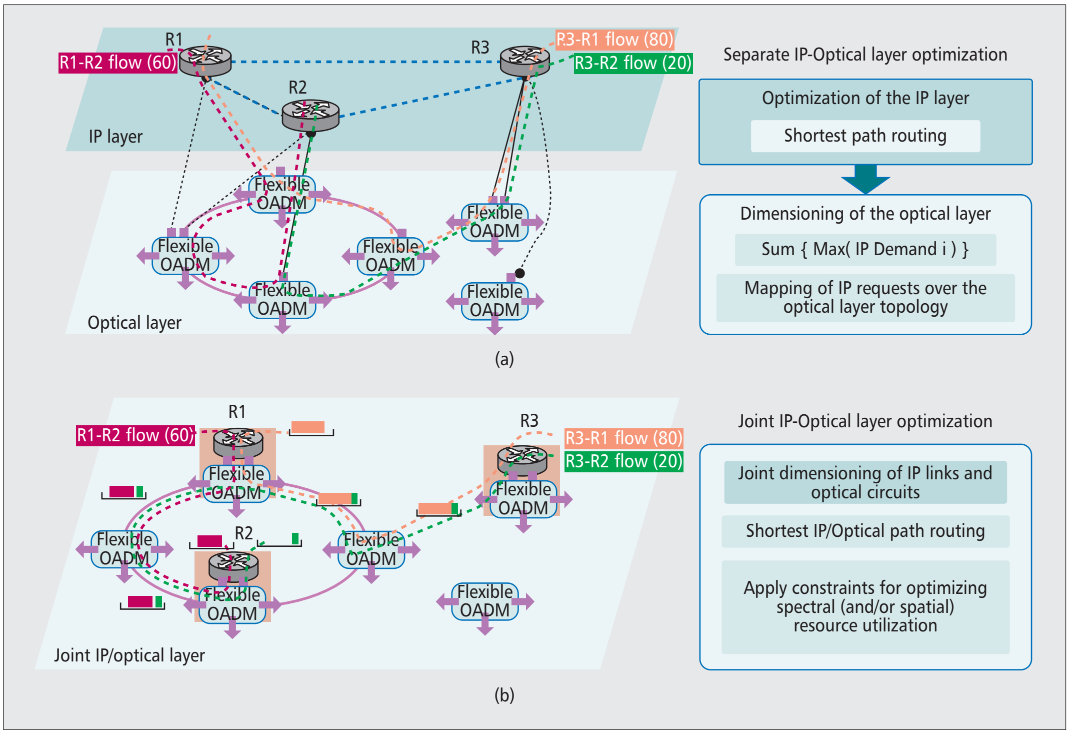


Figure 5. Multi-layer network planning approaches: a) the current approach, which treats IP and optical layers separately, providing optimized IP flow routing and then mapping of requests over optical layer topology; b) the joint IP-optical layer optimization approach, in which IP routing and optical resources are jointly considered and optimized for each demand.

The integration and joint optimization of the IP and optical layer has gained significant interest with the introduction of flexible optical networking. This is attributed to the flexible use of the optical layer resources which can be adapted to the capacity needs of the client (i.e. IP) layer.

ing (DSP). A list of spectrum allocation and defragmentation polices are presented in [12], where cooperative polices are also evaluated.

The introduction of the space dimension in flexible optical networking provides an extra degree of freedom with respect to the defragmentation options. Thus, in addition to the optical path rerouting and spectrum slot reassignment options, one could consider the spatial reassignment option. In this case, the fragmented spectral slots per space (e.g., fiber cores) can be freed by shifting the signals to other available cores in the same link. Also, the same spectral slots can be maintained (provided that spectral resources are available in different cores), thus avoiding the use of advanced laser tuning and DSP schemes. Referring to the allocation options presented earlier, the spatial reassignment defragmentation method is applicable to cases A, D, and E. Case C uses the space dimension to expand the SCh capacity and therefore has the same defragmentation needs as the spectrally flexible networking schemes. Case B faces no fragmentation issues since it resembles the WDM allocation method over fixed spectral slots, but expanded in space. It is noted that the potential benefits related to the spatial reassignment option still remain to be explored.

IP/Optical Integration

The integration and joint optimization of the IP and optical layer has gained significant interest with the introduction of flexible optical networking. This is attributed to the flexible use of the optical layer resources, which can be adapted to the capacity needs of the client (i.e., IP) layer.

Until now the IP and optical layers were managed independently, targeting primarily the routing optimization of the IP flows, while both layers were overdimensioned in an effort to absorb sporadic bursts and traffic peaks, as well as failures, without exploiting the switching capability of the optical layer. This approach is depicted in Fig. 5a, in which first the IP layer provides the optimized IP flow routing and policy management, followed next by an efficient mapping process of the IP traffic demands over the optical layer.

On a rather longer term, one could imagine much tighter integration of the flexible optical interface inside the IP router. Instead of having IP flow routing performed most of the time on the shortest path criterion, the path calculation process could directly integrate the flexible optical channel configuration and characteristics inside its IP policies, in such a way that routing and optical resources will be directly optimized (the approach depicted in Fig. 5b). If the IP router controls both the physical routing in the network topology and the optical channel configuration, it will be able to solve all operational issues presented earlier, linked to the multi-layer aspect of the transport network. In such a network management paradigm, the operator would have to manage optimized IP adjacencies over a "virtualized" optical infrastructure able to flexibly adapt its resources according to the IP layer demands (or even on a longer term based on the type of applications and services). This in turn enables the adoption of optical layer resource optimization criteria (e.g., spectrum/space utilization, cost and energy consumption reduction, network virtualization policies).

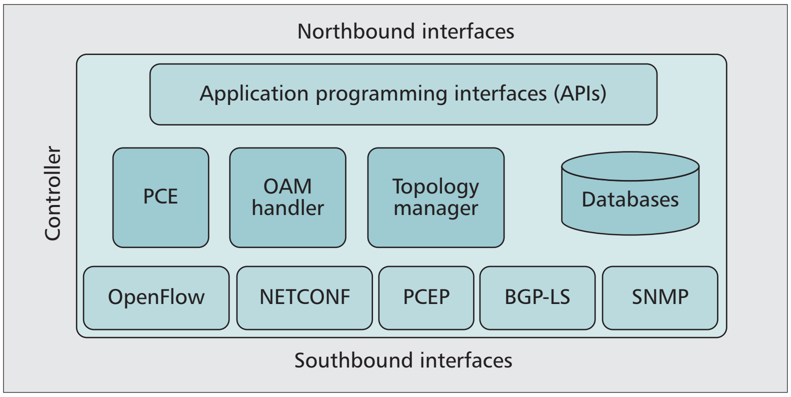


Figure 6. Converged ABNO and SDN controller architecture including the most relevant modules and interfaces.

The use of the space dimension introduces an additional degree of flexibility with respect to joint IP and optical layer resource optimization. For example, the space dimension can be used to separate different IP demands (e.g., according to classification or priority criteria) that are commonly allocated over the same source-destination path, while using the spectrum dimension to serve the end-to-end path allocation criteria over the network. Moreover, new potentials are introduced with respect to network virtualization by allowing the creation of spatially separated virtual network segments, with capacity flexibility maintained in the spectrum domain.

NETWORK CONTROL SOLUTIONS

Control plane solutions for optical transport networks have been steadily evolving in the last decade in order to adapt to technological breakthroughs and offer a wider set of network functionalities, while guaranteeing effective management of network resources. In fact, besides basic functionalities including interaction with clients, admission control, routing algorithms, lightpath provisioning, and survivability, future network controllers/orchestrators are expected to provide advanced functionalities such as real-time service modification, in-operation planning and optimization, programmability, virtualization, and refined monitoring procedures.

The possibility to exploit the space dimension as a further degree of freedom poses new challenges on the optical control plane side. Alongside the offered capacity increase and the additional networking features mentioned in the previous sections, the use of the space dimension in flexible networks mainly implies a significant change in the network model (related to the representation and configuration of nodes, switches, fibers, amplifiers, etc.), plus additional switching constraints to be carefully taken into account. Moreover, one of the functionalities that could benefit from the introduction of spatial flexibility in networks is virtualization. In fact, it would be possible to instantiate virtual slices embedded on top of different space dimensions of the same physical substrate. From the isolation perspective, with some flavors of SDM (e.g., bundles of SMFs or MCFs) it would also be possible to provide physical isolation by matching logical partitions to spatial dimensions.

Control plane architectures able to realize the aforementioned functionalities as well as the envisioned extensions from the introduction of the space dimension in optical networking range from legacy solutions based on generalized multiprotocol label switching (GMPLS), adequately extended, to more modern software defined networking (SDN)-based approaches.

GMPLS-based architectures investigated so far can be differentiated mainly on the basis of their computation and orchestration capabilities. They range from fully distributed solutions relying on source routing to centralized ones based on the path computation element (PCE). Centralized architectures are able to exploit dedicated computational resources to compute paths, store information about network status, and, in the most recent flavors, modify already active connections (i.e., active stateful PCE), or even autonomously set up or release connections independent of clients, thus behaving as a full network controller (i.e., PCE with instantiation capabilities). Despite all the different PCE/GMPLS flavors, functionalities such as programmability and network virtualization are difficult to implement in practice using the aforementioned architectures. The Internet Engineering Task Force (IETF) PCE-based architecture for application-based network operations (ABNO) [13] tries to fill this gap by proposing a modular framework composed of multiple existing building blocks (e.g., PCE). ABNO aims at providing advanced functionalities able to optimize traffic flows between applications, coordinate programming and provisioning of network resources, facilitate grooming, scheduling, and optimization, and provide virtual private network planning.

In parallel with the aforementioned PCE/GMPLS evolution, the SDN technology has been introduced at the beginning in the context of packet switching (e.g., within data centers). In SDN, a centralized controller is in charge of the orchestration and effective abstraction of the underlying infrastructure. The interface between the application and control layers goes under the name of northbound interface and is specified by application programming interfaces (APIs). Furthermore, SDN is designed to provide direct programmability of forwarding functions through the southbound interface, the OpenFlow protocol being the most popular. This way, no distributed protocols are adopted, and vendor interoperability can be better guaranteed. Moreover, a large community is actively contributing to SDN development through open source implementations (e.g., OpenDayLight).

Both the PCE/ABNO and SDN solutions are slowly converging toward a common controller architecture for optical networks. Figure 6 shows the most relevant components and interfaces that are typically included within such architecture: databases, topology manager, PCE, and a dedicated module for operation, administration, and maintenance (OAM), called the OAM handler. The architecture also includes northbound interfaces for network function virtualization and southbound interfaces for network configurations, where the GMPLS protocol suite is considered an alternative solution to the direct programmability of forwarding functions implemented by SDN.

Extensions to flexible optical networks have been proposed within GMPLS routing and signaling protocols, as well as within OpenFlow. These extensions include the definition of the occupied frequency resources (i.e., frequency slot) and the configuration of the most relevant transmission parameters, such as modulation format and code. However, because of the intrinsic complexity of data plane infrastructures, and the need to guarantee accurate monitoring and management capabilities, these protocols do not seem to be fully adequate for addressing all expected networking requirements. For example, GMPLS and OpenFlow solutions for handling transmission impairments and accurate filter configurations are still far from being standardized. To address this issue, additional effort is also in place to define the formal model of optical devices such that additional protocol solutions (e.g., XML-based protocols like NETCONF) may be utilized to directly program optical devices.

From the performance perspective, the use of GMPLS in flexible optical networks has demonstrated its capability to guarantee adequate recovery time in case of optical network failures, while some concern is still present when SDN is considered. On the other hand, SDN enables more accurate device configurations and facilitates advanced data plane solutions for network defragmentation and dynamic adaptation [14].

From a medium-/long-term perspective, the introduction of SDM will require further adaptations, as preliminarily proposed in [15]. The node models will need to take into account new designs and switching capabilities (and restrictions) of SDM WSSs and ROADMs. Also, the link models will need to include fiber types and physical characteristics, while dynamic algorithms for routing and spectrum allocation will have to consider the new switching features and constraints introduced by SDM. Such innovations will impact virtually all the protocols currently running in an optical control plane.

Conclusions and Remarks

The use of the space dimension in addition to WDM is introduced primarily in an effort to increase the overall capacity of optical transmission systems. However, by also extending the well studied spectral flexibility concept in the space dimension, the spectrally and spatially flexible networking concept is introduced, in which the spatial resources (fibers, cores, or modes) are considered in addition to spectral resources (wavelength, bandwidth, and modulation format) for overall network planning and resource optimization. This new concept generates several opportunities (compared to traditional reconfigurable WDM and GMPLS-based approaches) primarily with respect to channel allocation, resource optimization, defragmentation, joint IP/optical layer optimization, and virtualization, but also significant challenges in terms of complexity and network control issues that need to be explored in order to fully exploit the potential benefits of a spatially-spectrally flexible network solution. The different allocation options, as well as the potential benefits and challenges, are identified in this article.

Finally, it is worth noticing that the realization of the spectrally and spatially flexible networking concept requires the support of key technology innovations in data transmission, switching, and network elements in order to provide the ability to route traffic demands that are adaptively allocated over multiple data dimensions. Currently, significant research efforts have shown impressive results in new MCF, MMF, and FM-MCF types for SDM, ultra-high-capacity data transmission over MCF/MMF systems, and novel designs of SDM-WSS switches. Despite these efforts, a major practical limitation is still evident related to the enormous cost of replacing the existing fiber infrastructure with the new SDM fiber types. A shorter-term approach considers the use of fiber bundles for increasing the link capacity in networks. However, even in this case the spectrally-spatially flexible networking concept could be applicable, since it can expand the spectral flexibility in the space dimension, providing a powerful solution for overall resource optimization and the introduction of advanced network operating functions such as defragmentation, IP/optical layer integration, and spatial virtualization.

From a medium-/long-term perspective, the introduction of SDM will require further adaptations, as preliminarily proposed in [15]. The node models will need to take into account new designs and switching capabilities (and restrictions) of SDM WSSs and ROADMs.

Acknowledgments

The work presented here has received partially funding from the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreements no. 318415 FOX-C, no. 317999 IDEALIST, and no. 619732 INSPACE projects.

References

[1] O. Gerstel, M. Jinno, A. Lord, and S. J. B. Yoo, "Elastic Optical Networking: A New Dawn for the Optical Layer?," *IEEE Commun. Mag.*, vol. 50, no. 2, Feb. 2012, pp. S12–S20.

[2] M. Cvijetic *et al.*, "Dynamic Multidimensional Optical Networking Based on Spatial and Spectral Processing," *Optics Express*, vol. 20, no. 8, 2012, pp. 9144–50.

[3] P. J. Winzer, "Spatial Multiplexing: The Next Frontier in Network Capacity Scaling," *Proc. ECOC '13*, Sept. 2013, paper We.1.D.1.

[4] R. Rudnick *et al.*, "One GHz Resolution Arrayed Waveguide Grating Filter with LCoS Phase Compensation," *Proc. OFC/NFOEC '14*, Mar. 2014, paper Th3F.7.

[5] S. Sygletos *et al.*, "A Novel Architecture for All-Optical Add-Drop Multiplexing of OFDM Signals," *Proc. ECOC '14*, Sept. 2014, paper We.1.5.4.

[6] L.E. Nelson *et al.*, "Spatial Superchannel Routing in a Two-Span ROADM System for Space Division Multiplexing," *IEEE JLT*, vol. 32, no. 4, Feb. 15, 2014, pp. 783–89.

[7] Y. Sasaki *et al.*, "Dynamic Multidimensional Optical Networking Based on Spatial and Spectral Processing," *Optics Express*, vol. 20, no. 26, Dec. 2012, pp. B77–B84.

[8] I. Tomkos *et al.*, "A Tutorial on the Flexible Optical Networking Paradigm: State of the Art, Trends, and Research Challenges," *Proc. IEEE*, vol. 102, no. 9, Sept. 2014, pp. 1317–37.

[9] S. Fujii *et al.*, "On-Demand Spectrum and Core Allocation for Multi-Core Fibers in Elastic Optical Network," *J. Optical Commun. and Net.*, vol. 6, no. 12, Nov. 2014, pp. 1059–71.

[10] A. Muhammad *et al.*, "Routing, Spectrum and Core Allocation in Flexgrid SDM Networks with Multi-Core Fibers," *Proc. Int'l. Conf. Optical Network Design and Modeling '14*, May 2014, pp.192–97.

Significant research efforts have shown impressive results in new MCF, MMF, and FM-MCF types for SDM, ultra-high-capacity data transmission over MCF/MMF systems, and novel designs of SDM-WSS switches. Despite these efforts, a major practical limitation is still evident related to the enormous cost of replacing the existing fiber infrastructure with the new SDM fiber types.

[11] Xi.Wang *et al.*, "A Hitless Defragmentation Method for Self-Optimizing Flexible Grid Optical Networks," *Proc. ECOC '12*, Sept. 2012, p. P5.04.

[12] I. Stiakogiannakis *et al.*, "Dynamic Cooperative Spectrum Sharing and Defragmentation for Elastic Optical Networks," *J. Opt. Commun. and Net.*, vol. 6, no. 3, Feb. 2014, pp. 259–69.

[13] D. King and A. Farrel, "A PCE-Based Architecture for Application-Based Network Operations," IETF draft, draft-farrkingel-pce-abno-architecture.

[14] N. Sambo *et al.*, "Programmable Transponder, Code and Differentiated Filter Configuration in Elastic Optical Networks," *IEEE JLT*, vol. 32, no. 11, June 2014, pp. 2079–86.

[15] N. Amaya *et al.*, "Software Defined Networking (SDN) Over Space Division Multiplexing (SDM) Optical Networks: Features, Benefits and Experimental Demonstration," *Optics Express*, vol. 22, no. 3, 2014, pp. 3638–47.

Biographies

Dimitrios Klonidis (dikl@ait.gr) has been a senior research fellow at Athens Information Technology Centre (AIT) since September 2005. He received his Ph.D. degree from the University of Essex, United Kingdom, in 2005. He has been involved in the management of a large number of EU and national funded research projects in the areas of optical network design and optimization for access, metro, and core networks, transmission and switching technologies, network planning solutions, and experimental performance evaluation. He has co-authored over 110 publications in international journals and conferences.

Filippo Cugini (filippo.cugini@cnit.it) received his M.S. degree in telecommunication engineering from the University of Parma, Italy. Since 2001, he has been with the National Laboratory of Photonic Networks, CNIT, Pisa, Italy. His main research interests include theoretical and experimental studies in the field of optical networking. He is a co-author of 12 international patents and more than 100 IEEE publications.

Ori Gerstel [M'95, SM'05, F'08] (ori@ieee.org) is the founder/CTO of Sedona Systems. Until 2014, he was a principal engineer at Cisco, where he was responsible for the architecture of router-transport integration. Until 2002, he held senior architecture positions at Tellabs, Nortel, and IBM. He has published ~100 papers in the main conferences and journals in the field, and several book chapters. He holds over 35 patents, and a similar number of pending patents. He serves as Editor-in-Chief for *JOCN* and a Steering Committee Member for OFC. He received a Ph.D. degree from the Technion and is an OSA Fellow.

Masahiko Jinno [M'90-SM'12] (jinno.masahiko@lab.ntt.co.jp) currently serves as a professor of electronics and information engineering at Kagawa University. Prior to joining Kagawa University in October 2012, he was a senior research engineer, supervisor at Network Innovation Laboratories, Nippon Telegraph and Telephone Corporation (NTT). He has authored or co-authored over 180 peer-reviewed journal and conference papers in the fields of ultra-fast optical signal processing, ultra-wideband DWDM transmission systems in L-band and S-band, ROADM systems, GMPLS and application-aware optical networking, and elastic optical networks. He is a Fellow of the Institute of Electronics, Information and Communication Engineers (IEICE).

Victor Lopez (victor.lopezalvarez@telefonica.com) is a telecommunication engineer who received his M.Sc degree from Universidad de Alcalá de Henares in 2005 and his Ph.D. from Universidad Autónoma de Madrid (UAM) in 2009. In 2004, he joined Telefónica I+D as a researcher, where he was involved in transport networks projects. In 2006, he worked as a researcher and an assistant professor at UAM. In 2011, he joined Telefónica I+D as a technology specialist. He has co-authored more than 100 publications and contributed to IETF drafts. His research interests include the integration of Internet services over IP/MPLS and optical networks and control plane technologies (PCE, SDN, GMPLS).

Eleni Palkopoulou (epalkopo@cisco.com) is with Cisco Systems. Previously she was with AIT, Nokia Siemens Networks Research, and Siemens, Corporate Technology. She received her Diploma degree in electrical and computer engineering from the National Technical University of Athens, Greece, her M.Sc. from Munich University of Technology, Germany, and her Ph.D. degree from Chemnitz University of Technology, Germany. Her research interests include network architectures, network planning and optimization, and resilience mechanisms.

Motoyoshi Sekiya (sekiya.motoyosh@jp.fujitsu.com) is a director of Fujitsu Labs of America. His current research interests include optical networks, SDNs, network virtualization, and high-capacity optical transmission systems. He received his B.S. and M.E. degrees from Keio University in 1988 and 1990 respectively. He has served as a Technical Program Committee member for OFC/NFOEC and other conferences.

Domenico Siracusa (domenico.siracusa@create-net.org) is a senior researcher in the Future Networks (FuN) group at CREATE-NET. He received his Ph.D. in information technology from Politecnico di Milano in March 2012. His current research interests include control plane solutions for converged future networks, spatially-spectrally flexible optical networks, software-defined networking, and network virtualization. He has authored a number of publications in international conferences and journals on optical and transport networking technologies.

Gilles Thouénon (gilles.thouenon@orange.com) received his M.Sc. degree in electronics and telecommunication systems from the University of Brittany, France. He joined Orange Labs as a research engineer, where he has worked on transport network technologies for 16 years. His current research activity is focused on multi-layer transport network architecture optimization issues with the integration of new emerging networking concepts with optical flexibility. He has been involved in several French and European collaborative projects, and participated in a number of publications in journals on optical networking technologies.

Christophe Betoule (christophe.betoule@orange.com) is a research project manager at Orange Labs. He graduated from the French engineering school ISEN in June 2001. He is leading research activities on flexible optical networking and, more generally, projects on multilayer optical transport network architecture optimization issues. He has been involved in several French and European collaborative projects, and participated in a number of publications in journals on optical networking technologies.

3D Elastic Optical Networking in the Temporal, Spectral, and Spatial Domains

Roberto Proietti, Lei Liu, Ryan P. Scott, Binbin Guan, Chuan Qin, Tiehui Su, Francesco Giannone, and S. J. B. Yoo

Roberto Proietti, Lei Liu, Ryan P. Scott, Binbin Guan, Chuan Qin, Tiehui Su, and S. J. B. Yoo are with the University of California, Davis.

Francesco Giannone is with Altran Italia S.p.A.

ABSTRACT

Conventional elastic optical networking, EON, uses elasticity in two domains, time and frequency, to optimize utilization of optical network resources in the presence of fluctuating traffic demand and link quality. Currently, networking exploiting a third domain, space, is the focus of significant research efforts since space-division multiplexing, SDM, has the potential to substantially improve future network capacity and spectral efficiency. This article extends 2D-EON to include elasticity in all three domains: time, frequency, and space. We introduce enabling technologies, architectures, and algorithms for 3D-EONs. Based on sample network topologies, we investigate algorithms for routing, spectrum, spatial mode, and modulation format assignment — RSSMA. In particular, we investigate fragmentation-aware RSSMA and how the constraints in the formation of superchannels in MIMO-based SDM systems can impact the network performance in terms of blocking probability.

INTRODUCTION

While the rapid deployment of wavelength-division multiplexing (WDM) technologies sustained the explosive and exponential network traffic growths in the past decade, the continuing trend of exponential traffic growth driven by data centers and emerging new services is demanding deployment of more scalable and flexible networking technologies. The legacy WDM technologies can support traffic up to multiple terabits per seconds on a single fiber, but this is not sufficient to support future traffic demands with peak link capacity beyond 10 Tb/s. It is expected that commercial systems will need to support link capacity as high as 100 Tb/s by 2018 [1]. The recent renaissance of coherent optical communications with polarization-division multiplexing has enabled capacity of single fiber links up to 20 Tb/s. Further increases in capacity and spectral efficiency are very challenging because of the nonlinear Shannon limit [2]. The high optical power required for a high signal-to-noise ratio (SNR) starts to degrade the transmitted signal quality due to nonlinear optical effects in the transmission link [2] even for moderate (~500 km) transmission distances. Hence, transmission capacities beyond 100 Tb/s must explore a new and final frontier in optical communications — the space domain — by using space-division multiplexing (SDM). SDM allows the nonlinear Shannon capacity limit over a single-mode fiber to be overcome, providing a pathway toward petabit per second link capacities and spectrum efficiencies beyond 10 b/s/Hz for practical transmission distances.

Several papers in the literature have already discussed the advances and technological challenges in actual deployments of SDM. In addition to the straightforward bundled fiber approach, where the SDM system is composed of *N* independent single-mode fiber (SMF) systems, an SDM system can exploit few-mode fibers (FMFs) [3], multi-core fibers (MCF), and multi-(many)-mode fibers (MMFs) supporting orbital angular momentum (OAM) [4, 5], or other eigenstates. Each approach has benefits and drawbacks related to some particular technological challenges. FMF systems strongly rely on multiple-input multiple-output (MIMO) digital signal processing (DSP) due to strong mode coupling, but can benefit from the availability of FMF amplifiers [3]. MCF systems require complex fiber fabrication and the use of *N* separated amplification stages, but compared to FMF, the crosstalk and coupling between cores can be carefully controlled to levels that do not require or strongly limit the use of MIMO DSP. OAM states have relatively well defined azimuthal orthogonality, leading to possibly simpler MIMO DSP for propagation through ring-core FMFs or MMFs compared to other multimode propagation methods.

Figure 1 shows some representations of the three physical domains used to increase the capacity in fiber optic communication systems. In addition to the capacity increase, future networks must achieve flexible and agile utilization of network resources with scalable network control and management (NC&M). Recently, researchers proposed a new optical networking

The flexibility in spectrum and bandwidth allocation together with variable modulation formats allow more optimized utilization of network resources according to the given traffic demand and the link conditions.

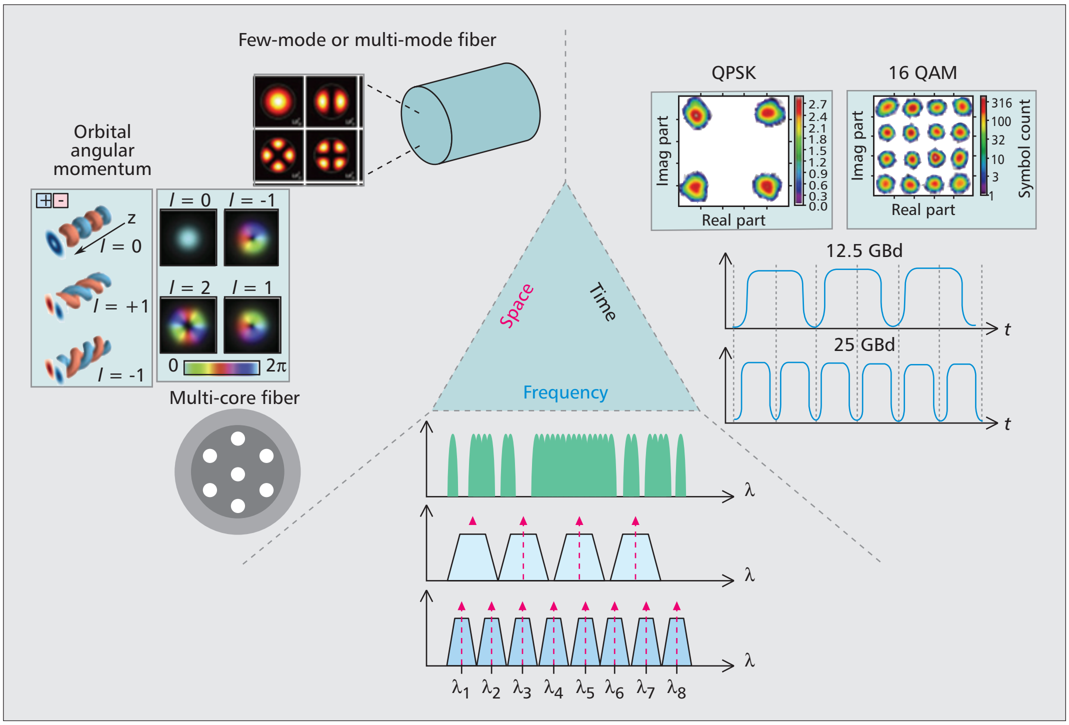


Figure 1. Three physical domains (frequency, time-modulation format, and space) to support elastic optical networking in the spectral, temporal, and spatial domains (3D-EON).

technology called elastic optical networking (EON) [6], in which each data link uses flexible spectral bandwidths and variable modulation formats. The flexibility in spectrum and bandwidth allocation together with variable modulation formats allow more optimized utilization of network resources according to the given traffic demand and the link conditions. Network resources optimization spans the temporal, spectral, and spatial domains, and efficient and effective algorithms need to be part of NC&M to allocate the available resources optimally. For instance, dynamic adaptation to varying traffic demand can lead to stranded and fragmented spectral resources in EON where there is no fixed spectral grid. Recent studies investigated spectral and temporal domain solutions for fragmentation-aware routing, spectrum, and modulation format assignment (RSMA) [7] in 2D-EONs. New studies in the temporal, spectral, and spatial domains for 3D-EONs are necessary.

At the physical layer, 2D-EON can employ coherent optical orthogonal frequency-division multiplexing (CO-OFDM), coherent optical WDM (CO-WDM), Nyquist WDM [8], or dynamic optical arbitrary waveform generation (OAWG) and optical arbitrary waveform measurement (OAWM) technologies [9], adopting various modulation formats across the flexible bandwidth depending on the reach. The 2D-EONs have achieved ~1.6× improvements compared to the fixed-grid networks in capacity, utilization, and availability in the network. To achieve more than 10× improvements we must introduce a new degree of freedom in the spatial domain using SDM. The elasticity concepts discussed above can then also be extended to the spatial domain. In [10] the authors proposed spatial-spectral processing and coded modulation to exploit elasticity in the spectral and spatial domains. Reference [11] reports the first demonstration of an elastic SDM system based on MCFs capable of switching the traffic in the spatial, spectral, and time domains. Reference [12] discusses a new routing, spectrum and core assignment (RSCA) for SDM systems based on multicore fibers, taking into account the crosstalk between cores. In 3D-EONs, the reliance on electronic DSP (with or without MIMO) can limit the total capacity.

This article discusses EON in the temporal, spectral, and spatial domains (3D-EON) for future optical fiber communication networks. In the following sections, this article:

- Introduces the concept of elasticity in the time, spectral, and spatial domains
- Presents enabling technologies for SDM systems
- Discusses the routing, spectral, spatial mode, and modulation format assignment (RSSMA) algorithm, which extends RSMA by exploiting the additional space domain for resource allocation

Elasticity in the Temporal, Spectral, and Spatial Domains

Figure 2 illustrates the concept of elasticity in time, frequency, and space for 3D-EON. Let us assume that we have an SDM system with four spatial modes and that each node in the network has a certain number of transceivers which can tune in wavelength, line rate, and modulation format. If the network needs to provision a 1 Tb/s connection between two nodes (e.g., from

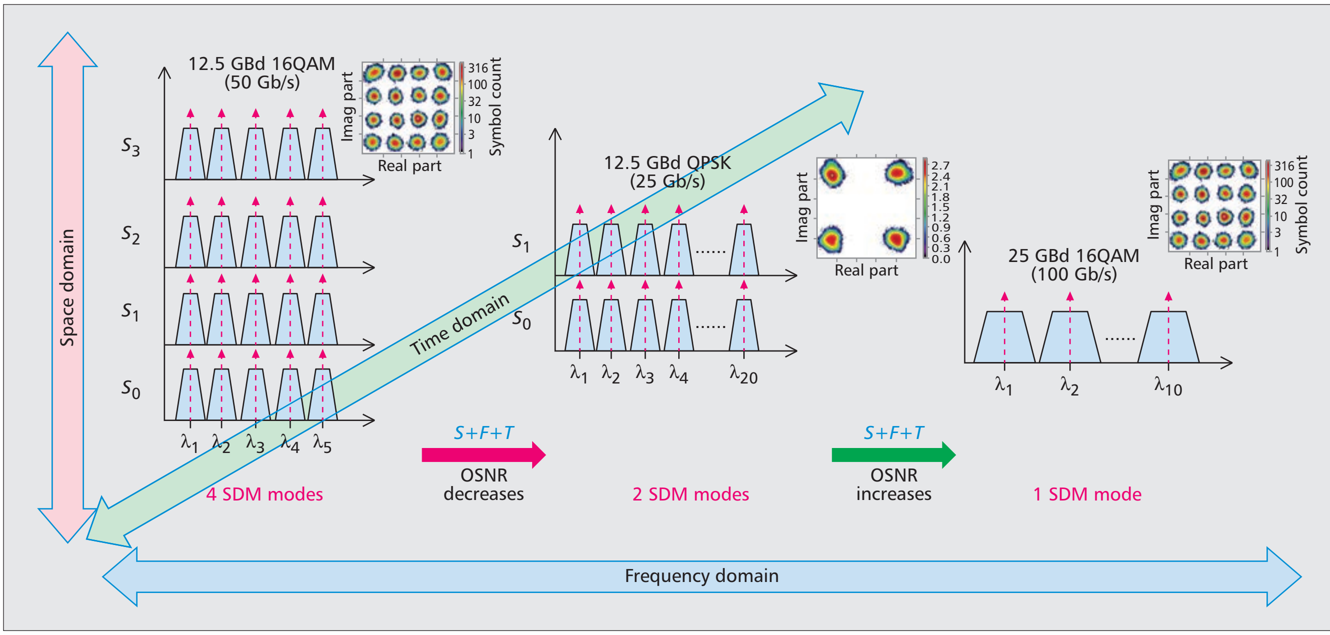


Figure 2. Three different possibilities to provision a 1 Tb/s connection exploiting flexibility in space, frequency, and time (S+F+T).

node A to node D in Fig. 4a), the flexibility provided by the three domains (temporal, spectral, and spatial) may allow this 1 Tb/s connection to be established in different ways. At the beginning, the network control plane could decide to use the solution in Fig. 2 (left) by creating a superchannel that spans across all four spatial modes occupying five 12.5 GHz spectral slots on each mode. Each spectral slot is filled in this case with a 12.5 Gbaud, 16-quadrature amplitude modulated (QAM) signal. If the network condition changes, for instance, the optical SNR (OSNR) at the receivers decreases, the network may be able to reconfigure the connection as shown in Fig. 2 (center) by using only two spatial modes, reducing the complexity of the modulation format to quadrature phase shift keying (QPSK) and occupying 20 spectral slots on each mode. This reconfiguration exploits the elasticity in all three domains (S+F+T). Figure 2 (right) shows another example of elasticity in space, frequency, and time. In this case the OSNR at the receiver is high enough to provision the 1 Tb/s connection by using one single spatial mode and 10 25-GHz spectral slots filled with 25 Gbaud 16-QAM signals. In this case, the flexibility in the three domains allows saving 50 percent of the network resources in terms of number of transceivers used.

While the scenario depicted in Fig. 2 is ideal, an actual implementation of elasticity in the time, frequency, and space domains could be constrained by:
- Crosstalk among different spatial modes
- The need for MIMO processing to combat crosstalk impairments and the associated constraints on the formation of superchannels
- MIMO DSP complexity
- Intermodal nonlinear impairments
- The SDM node architecture and its switch and add/drop traffic granularity (trade-off between flexibility and complexity)
- The number of TRXs available in the transponders

All of the above aspects are active research topics in spectral and spatial EON.

Enabling Technologies for Space-Division Multiplexing

Although practical SDM technologies are still under active development, we must understand the current status of SDM technologies when considering 3D-EON. Various technological approaches to SDM fibers include MCFs, FMFs, and MMFs. An MCF contains multiple cores and operates in one of the two regimes, independent cores with sufficiently large spacing to limit the crosstalk or coupled cores to support super modes. An FMF has a single core with a dimension that restricts the number of modes (e.g., less than a dozen distinct modes). An MMF includes ring-core or "vortex" fibers that support OAM modes with limited crosstalk. At the research level, a few experiments have already demonstrated long-distance (> 1000 km) transmission of multiple spatial modes using MCFs or FMFs (using either commercial single-mode Erbium-doped fiber amplifiers or FMF amplifiers [3]), while demonstrations of OAM-based SDM fiber transmission systems are limited to a few kilometers so far. However, the deployment of SDM systems can be conceivable only if they provide benefit in terms of cost and performance compared to their counterpart using multiple stranded fibers. To this aim, as discussed in [1, 3], optical amplification and integration are essential for practical implementation of SDM in actual systems.

A compact and integrated MUX/DEMUX represents an important building block toward the realization of compact and energy-efficient transponders and reconfigurable optical add/drop multiplexers. In the case of FMF, pho-

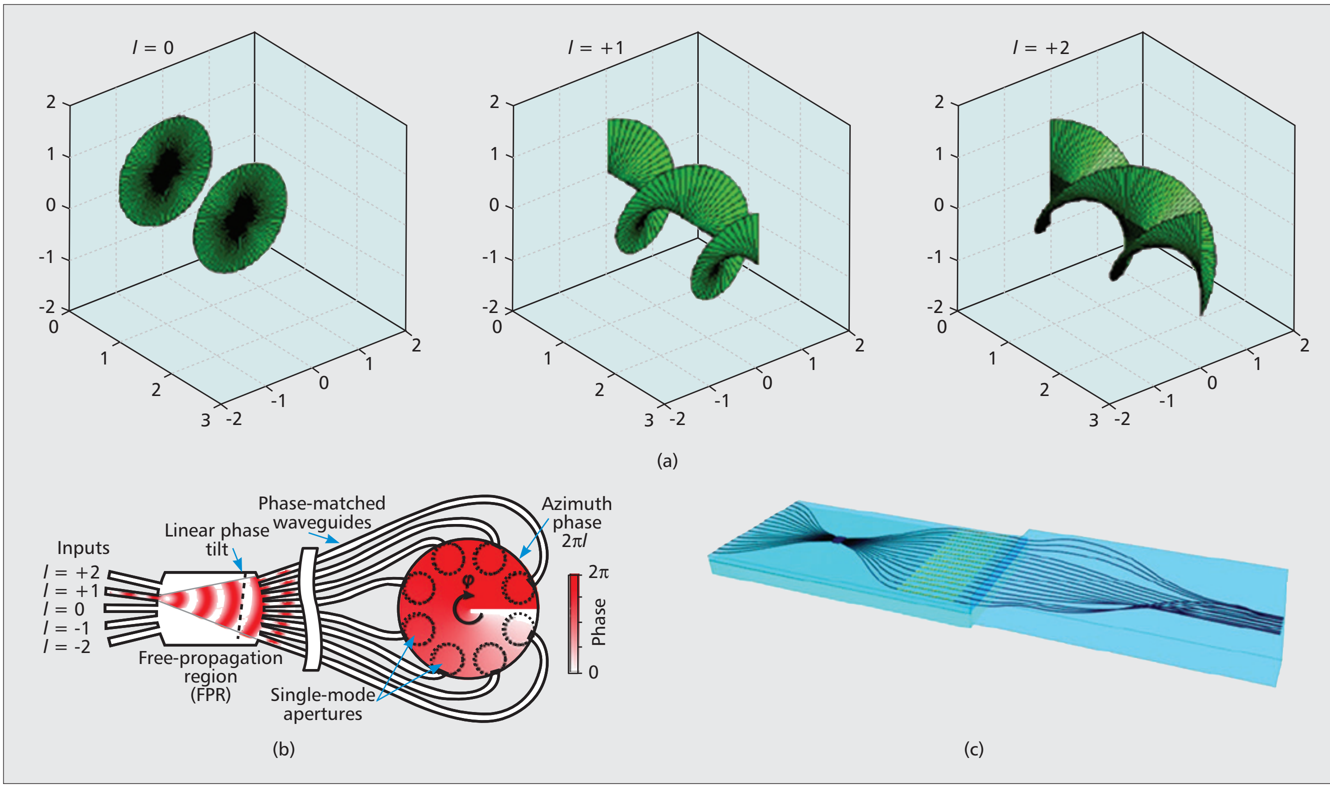


Figure 3. a) 3D illustration of OAM mode propagation for state numbers $\ell = 0, +1, -1$; b) integrated device conceptual configuration for an OAM multiplexer/demultiplexer. The OAM multiplexer converts the linear phase front tilt to azimuthal varying phase. Input OAM states at $\ell = +2, \ldots, -2$ are each projected to an array of phase-matched waveguides to excite apertures with azimuthally varying phase dependent on the OAM state number ℓ. This way, multiple OAM states can be multiplexed spatially (SDM); c) 3D illustration of the hybrid integrated device.

tonic lanterns implemented from all-fiber or 3D-waveguide technologies have shown promise as mode-group-selective spatial multiplexers [13]. Alternatively, OAM multiplexers based on hybrid 3D photonic integrated circuits (PICs), when coupled with vortex fibers, may also provide limited crosstalk data paths through a single fiber. As an example, we provide some details of an OAM MUX/DEMUX device fabricated at the University of California, Davis (UC Davis).

As Figure 3a illustrates, a light beam carrying OAM exhibits a specific azimuthal phase variation determined by OAM state number, l, which is an integer. OAM states with differing charge numbers are orthogonal, and the sign of l determines the handedness of the helical phase front. The light beam can simultaneously support multiple OAM states (potentially an infinite number, subject to the Shannon limitation imposed by the SNR). Therefore, OAM has the potential to significantly improve the spectral efficiency or photon efficiency of free-space and fiber optical communications. Recent experimental results [5] show successful terabit-per-second transmission by multiple OAM modes propagating over a specially designed fiber without using computationally intensive DSP MIMO algorithms.

Figures 3b and 3c show a device fabricated at UC Davis performing OAM multiplexing and demultiplexing based on PICs containing a free propagation region (FPR). The FPR induces a linear phase tilt on the input signals based on their input waveguide position. After geometrical transformation by the phase-matched waveguides, this phase tilt provides progressive phase evolution in the azimuthal angle as required for the OAM state generation (Figure 3b). The superposition principle of waves indicates that multiplexing of multiple OAM states is possible, while the propagation of the waves in reverse will achieve OAM state demultiplexing. The device supports multiplexing and demultiplexing of up to 15 OAM states, with both TE and TM polarizations and relatively low loss performance at 1.55 μm and crosstalk better than –15 dB. The compact design and single-mode interface can guarantee easy connection with other high-speed optical components for future on-chip integration. The device was fabricated using 3D direct laser inscribing technology [4] that has also been used to fabricate photonic lantern mode-group selective multiplexers for FMF [13].

RSSMA in the Temporal, Spectral, and Spatial Domains

Although the details of particular SDM implementations are an active research topic, our discussion below on routing, spectrum, spatial mode, and modulation format assignment (RSSMA) is relatively agnostic to particular SDM techniques, so the implications of the work are general.

In a 3D-EON, an RSSMA algorithm is necessary in order to route a connection request and

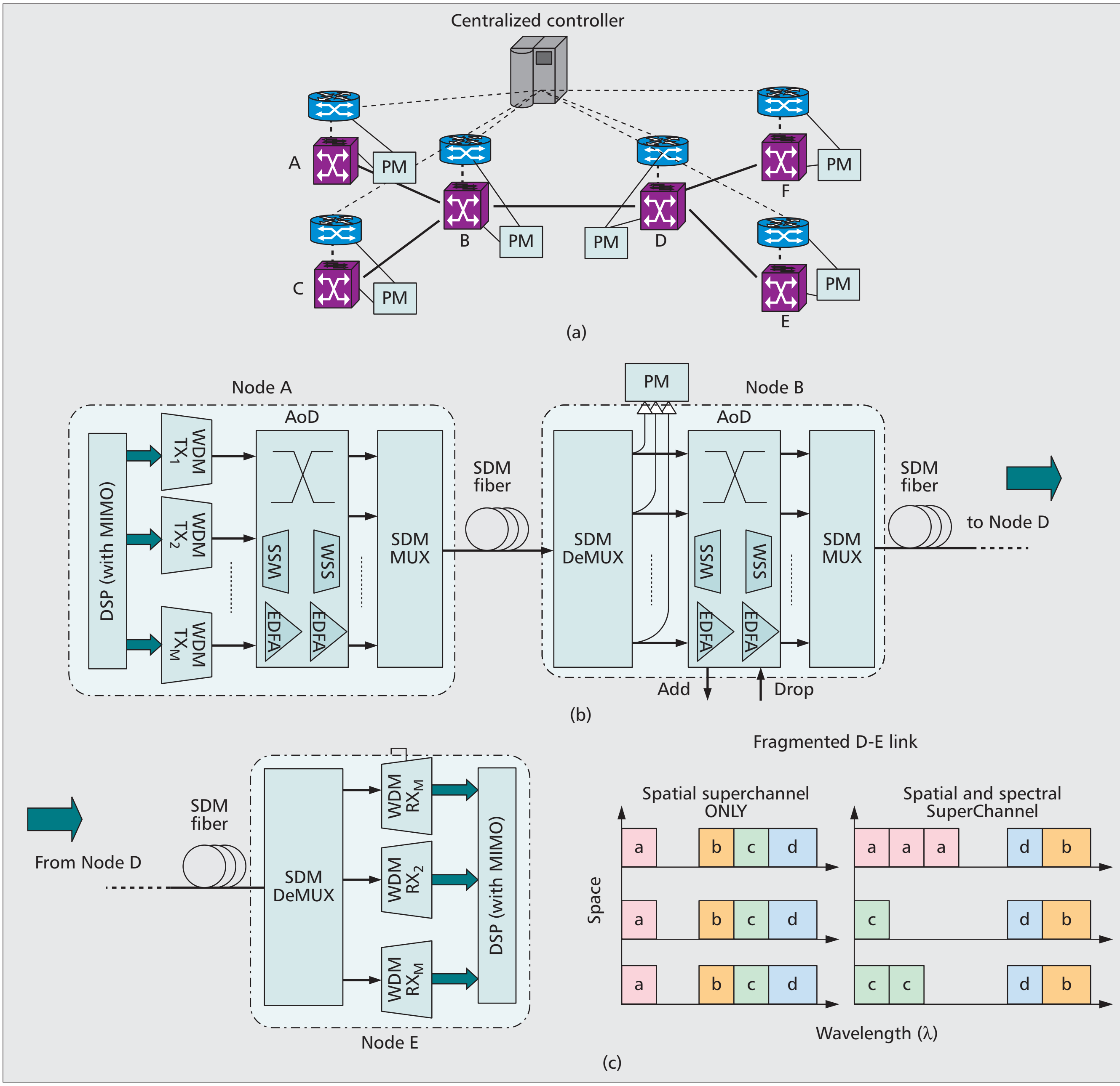


Figure 4. a) Example of a six-node 3D-EON network with centralized control plane and performance monitoring (PM); b) node *A*–node *E* link system diagram; c) Fragmentation of *D-E* link status: (left) only spatial superchannels are used, leading to spectral fragmentation only; (right) both spatial and spectral superchannel are used, leading to spectral and spatial fragmentation. WDM TX/RX: multi-wavelength transmitters/receivers, implementing elasticity in the spectral-/temporal domains; AoD: architecture on demand; WSS: wavelength selective switch; EDFA: Erbium-doped fiber amplifier; MUX/DeMUX: multiplexer/demultiplexer.

assign resources for this request in the network. Figure 4a shows a six-node 3D-EON network with a centralized control plane. The control plane runs RSSMA algorithms to maintain efficient network level utilization. Each node can have an impairment monitoring system that transmits the information on the status of the links to a centralized controller, which will use this information for the RSSMA algorithm.

Figure 4b shows a system-level schematic of the *N*-mode SDM links used in the 3D-EON network of Figure 4a (i.e., the node A–node E link). *M* multi-wavelength transmitters, implementing elasticity in the spectral-temporal domains, are spatially multiplexed. In addition to the required DSP to implement CO-OFDM, Nyquist WDM, or OAWG, it might also be necessary to use MIMO DSP in the case of mode mixing during transmission through the optical link. An architecture on demand (AoD) programmable node [11] is placed between the SDM MUX and the *M* TXs to guarantee reconfigurability and switching granularity at the wavelength level, and fully exploit the spectral and spatial domain. The AoD consists of a large port-count optical backplane (e.g., micro-electro-

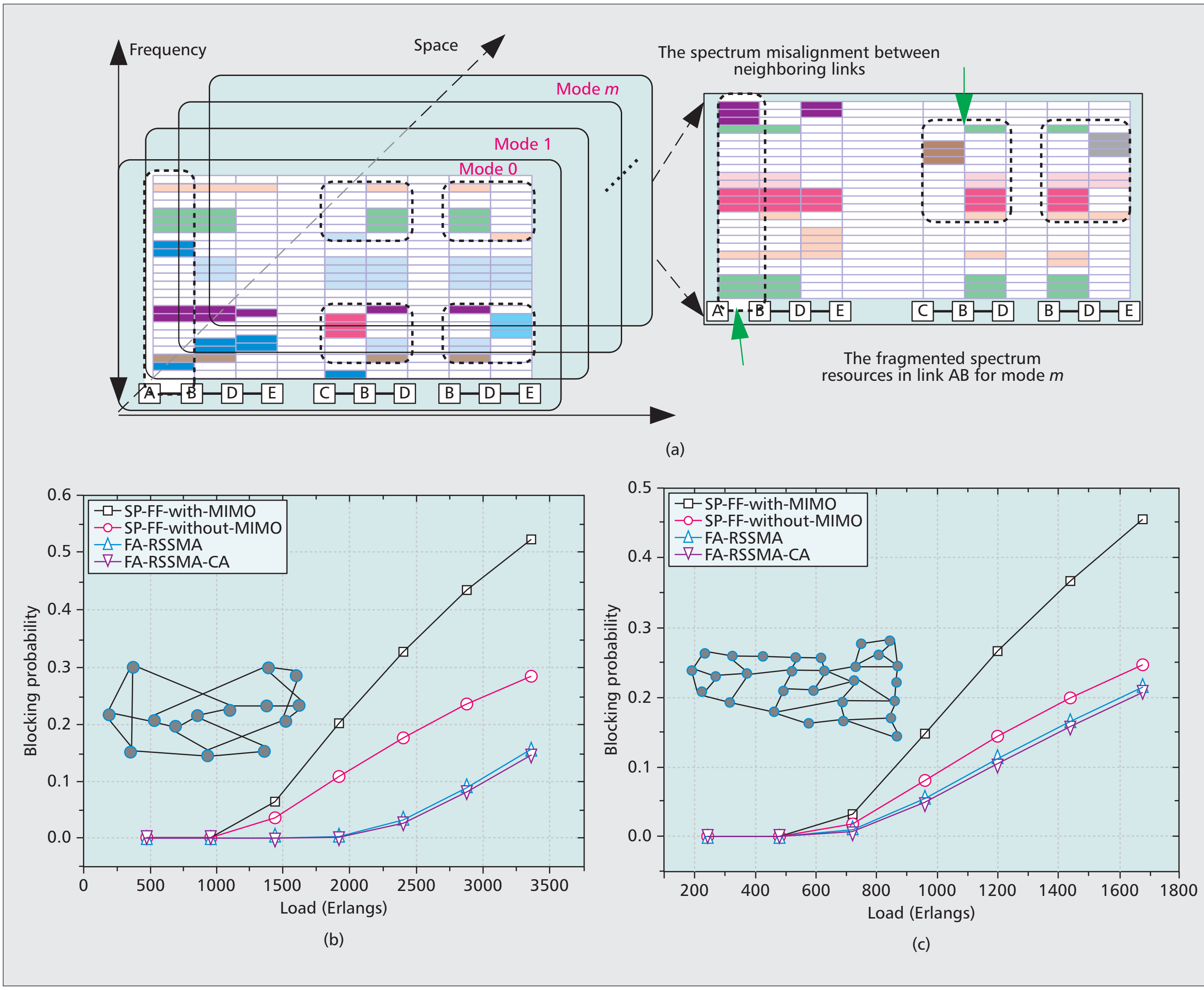


Figure 5. a) Spectral fragmentation on the spectral-spatial-temporal dimensions; b) blocking probability comparison of different algorithms in the 14-node NSFNET network; c) blocking probability comparison of different algorithms in the 28-node USBN network.

mechanical systems, MEMS), and several plug-in modules such as flexible-grid wavelength selective switches (WSSs) and optical amplifiers. After the SDM MUX, the SDM signal is launched through an SDM-compatible fiber. At node B, the different spatial modes are demultiplexed. A small amount of optical power is tapped from each of the SDM DEMUX outputs to feed the impairment monitor system, which provides information on the link quality to the centralized controller. The *N* SDM DEMUX outputs then connect to an AoD node that serves as a reconfigurable optical add-drop multiplexer (ROADM), allowing adding, dropping and switching of different flexible channels with granularity down to the wavelength level. *N* AoD outputs connect to another SDM MUX to launch the SDM signal into the next fiber span and reach the next nodes all the way to the final node, E. At node E, an SDM DEMUX spatially demultiplexes the different spatial mode. Each mode is then spectrally demultiplexed and fed into a bank of coherent receivers with DSP.

The RSSMA Algorithm and Simulation Results

The amount of crosstalk among different spatial modes and the need for MIMO processing will not only impact the physical layer performance, but also the way superchannels are formed and routed. It has previously been pointed out [1, 3] that in the case of an SDM system relying heavily on MIMO, it is necessary to create a super-channel in the space domain with an end-to-end signal that occupies all of the spatial modes for a given wavelength slot (e.g., connections **a**, **b**, **c**, and **d** in Figure 4c on the left). It is evident that this constraint on the creation of spatial super-channels will limit the flexibility to switch, add, and drop single SDM channels. However, at the same time, it will simplify the network control since fragmentation issues are confined to only the wavelength domain (Fig. 4, left). Conversely, in the case of an SDM system not relying on MIMO (this is possible for certain MCF designs and potentially for OAM systems), the creation of super-channels and resource utilization would

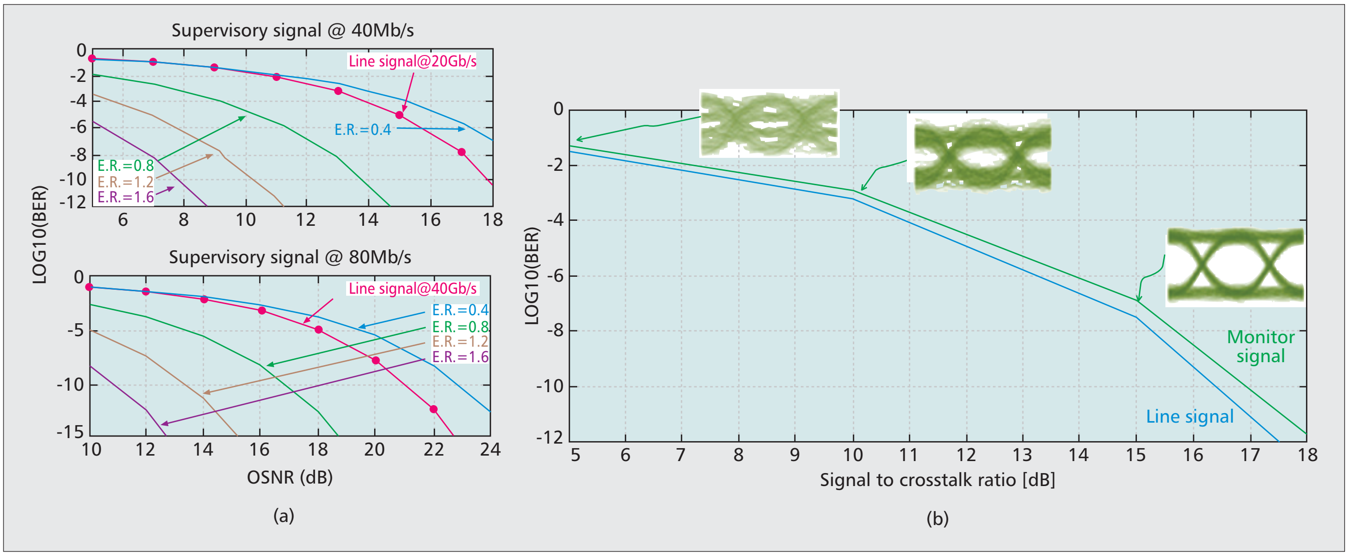


Figure 6. a) Simulated correlation between the bit error rate of the data and supervisory signals as a function of the extinction ratio (ER) parameter of the intensity modulator used to overmodulate the data signal (data signal is QPSK-modulated at 20 Gb/s, top, and 40 Gb/s, bottom); b) simulation results show strong correlation between the bit error rate of data and monitor signals as a function of the signal-to-crosstalk ratio for a 10 Gbaud QPSK signal with a 20 Mb/s monitor signal. The results suggest that the supervisory channel technique can do performance monitoring for inband crosstalk in SDM systems.

be more flexible. In this case, however, fragmentation in both the space and wavelength domains might occur (Fig. 4c, right).

In conventional 2D-EONs, routing, spectral, and modulation format assignment (RSMA) [7] along with defragmentation in the spectral domain [14] have been widely investigated. Considering that defragmentation operations are complex, we investigated fragmentation-aware and alignment-aware RSMA algorithms to minimize the need for defragmentation. For 3D EONs, we investigate routing, spectral, spatial mode, and modulation format assignment (RSSMA) with similar fragmentation awareness and alignment awareness. Using an example network topology of Fig. 4a, the dashed rectangle in Fig. 5a indicates the fragmented spectrum resources in link *AB*, which is the fragmentation in the spectral dimension. In addition, considering that spectrum misalignment between neighboring links (dashed rectangle in Fig. 5a) will most likely increase the end-to-end blocking probability, we propose two RSSMA algorithms. The first is fragmentation-aware RSSMA (FA-RSSMA), and the other is FA-RSSMA with congestion avoidance (FA-RSSMA-CA). The operational procedures for both algorithms are different, but their objective is the same: any new connection is set up to fragment the least number of continuous spectral blocks on candidate links, while it fills up as many misaligned spectral slots as possible on neighboring links for a given spatial mode.

We have conducted simulations on sample network topologies to compare the blocking probability of the proposed fragmentation-aware RSSMA algorithms with the commonly used benchmark algorithm, that is, the shortest path routing and first-fit spectrum (and spatial mode) assignment (SP-FF) algorithm. For comparison purposes, we measured the performance of SP-FF for the 3D EON deployed with an SDM system relying on MIMO and without MIMO (i.e., SP-FF-with-MIMO and SP-FF-without-MIMO, respectively). We simulated dynamic connection arrival and departure events on a 14-node NSFNET network and a 24-node U.S. backbone network (USBN), respectively. Each fiber link has 400 spectral slots, and each slot is set to be 12.5 GHz. In the simulation, connection requests are randomly generated between each source-destination pair according to a Poisson process. The holding time of each connection follows a negative exponential distribution with an average of five time units. The connection bandwidth is randomly selected in the range of [1, 10] slots. The number of spatial modes is 8. Figures 5b and 5c compare the blocking probability of the different algorithms. As expected, the use of MIMO has a significant impact on the network blocking performance, leading to the worst blocking probability caused by an inflexible resource allocation due to the constraint mentioned above on the creation of spatial super-channels. The results in Fig. 5 (b–c) also show that both the FA-RSSMA and FA-RSSMA-CA algorithms can greatly reduce the blocking probability, compared to that of the non-fragmentation-aware SP-FF algorithm. The FA-RSSMA-CA outperforms FA-RSSMA, but the difference is relatively small.

Supervisory Channel and Performance Monitoring for 3D EON

In addition to the impairments already present in a 2D-EON (optical noise accumulation, four-wave mixing, self-phase modulation, etc.), the spatial domain may introduce a significant amount of in-band crosstalk due to the mode coupling in the transmission media and crosstalk in the spatial MUX/DEMUX elements distributed along the fiber links.

As we already experimentally demonstrated

The operational procedures for both algorithms are different, but their objective is the same: any new connection is setup to fragment the least number of continuous spectral blocks on candidate links, while it fills up as many misaligned spectral slots as possible on neighboring links, for a given spatial mode.

in [15], a real-time performance monitoring method across the spread spectrum allows the network to dynamically and adaptively adjust the modulation format to maximize spectral efficiency while maintaining the required quality of transmission (QoT) and bit error rate (BER) performance even for signals that experience time-varying physical layer impairments. The monitoring technique consists of a low-speed supervisory channel modulating the high-speed data with a low modulation index (e.g., ~0.1). Since the data and the supervisory channel signals follow that same path, the BER of the data is strongly correlated with that of the supervisory channel, as shown in Fig. 6a. This correlation allows estimation of the BER of the data at different modulation formats, and the network control plane can choose the modulation format that maximizes the spectral bandwidth while meeting the transmission requirements. The same spread-spectrum supervisory channel can also be applied to each spatial mode in a 3D-EON network to monitor the QoT of each single spatial mode.

The simulation results in Fig. 6 show how the supervisory channel technique described above can do performance monitoring in case of inband crosstalk, which is a very critical impairment in SDM systems due to mode coupling in multi-modeMMFs or MCFs. Figure 6a shows the simulated correlation between the BER of the data and supervisory signals, as a function of the extinction ratio (ER) parameter of the intensity modulator used to overmodulate the data signal. The results are obtained for a data signal with QPSK modulation format at 20 Gb/s (top) and 40 Gb/s (bottom). As expected, the ratio between the line rate of the data and supervisory signals remains constant (40 Mb/s for 20 Gb/s and 80 Mb/s for 40 Gb/s). More importantly, the simulation results in Fig. 6b show the simulated correlation between the BER of the data and supervisory signals as a function of the signal to inband crosstalk ratio in an SDM system for a 10 Gbaud QPSK signal with a 20 Mb/s supervisory channel. The results show strong correlation between the BERs of the data signal and supervisory channel.

Conclusion

We discuss temporal, spectral, and spatial elastic optical networking, and related physical layer technologies and networking aspects. We introduce the concept of elasticity in the time, frequency, and space domains, and discuss recent technological advances toward integrated photonic SDM multiplexers and demultiplexers. We further discuss RSSMA algorithms, which take advantage of the spatial dimension to reduce the blocking probability. The study considers the possible constraints on the allocation of spectral and spatial slots given the use of MIMO processing, which may be needed to combat in-band crosstalk in SDM systems. Finally, we discuss the application of a supervisory channel performance monitoring technique in the new 3D-EON to monitor in-band crosstalk resulting from mode coupling in SDM systems.

Acknowledgments

This work was supported in part by DOE under grant DE-FC02-13ER26154, by NSF under EECS grant 1028729, and by DARPA DSO under grants HR0011-11-1-0005 and W911NF-12-1-0311.

References

[1] P. J. Winzer, "Making Spatial Multiplexing A Reality," *Nature Photonics*, vol. 8, 2014, pp. 345–48.

[2] R. Essiambre *et al.*, "Capacity Limits of Optical Fiber Networks," *J. Lightwave Tech.*, vol. 28, 2010, pp. 662–701.

[3] D. Richardson, J. Fini, and L. Nelson, "Space-Division Multiplexing in Optical Fibres," *Nature Photonics*, vol. 7, 2013, pp. 354–62.

[4] B. Guan *et al.*, "Free-Space Coherent Optical Communication with Orbital Angular, Momentum Multiplexing/Demultiplexing Using a Hybrid 3D Photonic Integrated Circuit," *Optics Express*, vol. 22, 2014/01/13 2014, pp. 145–56.

[5] N. Bozinovic *et al.*, "Terabit-Scale Orbital Angular Momentum Mode Division Multiplexing in Fibers," *Science*, vol. 340, June 28, 2013 2013, pp. 1545–48.

[6] M. Jinno *et al.*, "Spectrum-Efficient and Scalable Elastic Optical Path Network: Architecture, Benefits, and Enabling Technologies," *IEEE Commun. Mag.*, vol. 47, 2009, pp. 66–73.

[7] Y. Yin *et al.*, "Spectral and Spatial 2D Fragmentation-Aware Routing and Spectrum Assignment Algorithms in Elastic Optical Networks [Invited]," *J. Optical Commun. and Networking*, vol. 5, 2013/10/01 2013, pp. A100–A106.

[8] G. Bosco *et al.*, "On the Performance of Nyquist-WDM Terabit Superchannels based on PM-BPSK, PM-QPSK, PM-8QAM or PM-16QAM Subcarriers," *J. Lightwave Technology*, vol. 29, 2011, pp. 53–61.

[9] R. P. Scott *et al.*, "Dynamic Optical Arbitrary Waveform Generation and Measurement," *Optics Express*, vol. 18, Aug. 30 2010, pp. 18,655–70.

[10] M. Cvijetic, I. B. Djordjevic, and N. Cvijetic, "Dynamic Multidimensional Optical Networking based on Spatial and Spectral Processing," *Optics Express*, vol. 20, Apr. 2012, pp. 9144–50.

[11] N. Amaya *et al.*, "First Fully-Elastic Multi-Granular Network with Space/Frequency/Time Switching Using Multi-core Fibres and Programmable Optical Nodes," *ECOC*, Amsterdam, The Netherlands, 2012, Th.3.D.3.

[12] A. Muhammad *et al.*, "Routing, Spectrum and Core Allocation in Flexgrid SDM Networks With Multi-Core Fibers," *2014 Int'l. Conf. Optical Network Design and Modeling*, 2014, pp. 192–97.

[13] R. Ryf *et al.*, "1705-km Transmission over Coupled-Core Fibre Supporting 6 Spatial Modes," *ECOC*, 2014, pp. 1–3.

[14] R. Proietti *et al.*, "Rapid and Complete Hitless Defragmentation Method Using a Coherent RX LO with Fast Wavelength Tracking in Elastic Optical Networks," *Optics Express*, vol. 20, Nov. 2012, pp. 26,958–68.

[15] D. J. Geisler *et al.*, "Experimental Demonstration of Flexible Bandwidth Networking with Real-Time Impairment Awareness," Optics Express, vol. 19, Dec. 2011, pp. B736–B745.

Biographies

Roberto Proietti (rproietti@ucdavis.edu) received his M.S. degree in telecommunications engineering from the University of Pisa, Italy, in 2004, and his Ph.D. in fiber optical communications from Scuola Superiore Sant'Anna, Pisa, Italy, in 2009. He is a project scientist at UC Davis. His research interests include high-spectrum-efficiency coherent transmission systems and elastic optical networking, access optical networks and radio over fiber, and optical switching technologies and architectures for supercomputing and data center networks.

Lei Liu [S'08, M'11] (liulei@ieee.org) received B.E. and Ph.D. degrees from Beijing University of Posts and Telecommunications, China, in 2004 and 2009, respectively. From 2009 to 2012, he was a research engineer at KDDI R&D Laboratories Inc., Japan, where he was engaged in research and development on control planes for optical networks. He is now with UC Davis. He has coauthored more than 100 peer-reviewed papers in prestigious international journals and conferences.

RYAN P. SCOTT [M'03] (rpscott@ucdavis.edu) received his Ph.D. degree in electrical and computer engineering from UC Davis in 2009. He is currently an assistant project scientist at UC Davis, and his present research interests include optical arbitrary waveform generation and measurement, optical comb generation, spatial division multiplexing, and direct laser inscription of 3D optical waveguides.

BINBIN GUAN (rayguan@ucdavis.edu) received his B.S. degree in optics engineering from Zhejiang University, China, in 2008. He is currently working toward his Ph.D. degree in electrical and computer engineering at UC Davis. His current graduate research focuses on optical coherent communications and optical signal processing.

CHUAN QIN (chaqin@ucdavis.edu) received his B.S. degree in optical sciences and engineering from Zhejiang University, Hangzhou, China, in 2011. He is currently working toward his Ph.D. degree in electrical and computer engineering from UC Davis. His research interests include high-speed optical arbitrary waveform generation and measurement, spatial division multiplexing, and coherent optical communications.

TIEHUI SU (tiesu@ucdavis.edu) received his B.S. degree in electric engineering from Peking University, Beijing, China, in 2008. He is currently working toward his Ph.D. degree in electrical engineering at UC Davis. His research interests include photonic integrated circuits design, orbital angular momentum (OAM), and AWG devices design on various material platforms.

FRANCESCO GIANNONE (frcgnn@gmail.com) received his M.S. degree in telecommunications engineering from the University of Pisa in 2014. He was a visiting research scholar at UC Davis from September 2013 to June 2014. He is currently working as a consultant engineer at Altran Italia S.p.A.

S. J. BEN YOO [S'82, M'84, SM'97, F'07] received his B.S. degree in electrical engineering with distinction, his M.S. degree in electrical engineering, and his Ph.D. degree in electrical engineering with a minor in physics, all from Stanford University, California, in 1984, 1986, and 1991, respectively. He currently serves as a professor of electrical engineering at UC Davis. His research at UC Davis includes optical switching devices, systems, and networking technologies for future computing and communications.

NETWORK AND SERVICE VIRTUALIZATION

Kostas Pentikousis

Catalin Meirosu

Diego R. Lopez

Spyros Denazis

Kohei Shiomoto

Fritz-Joachim Westphal

The introduction of software-defined networking (SDN) and network function virtualization (NFV) has altered, in a wholesale manner, the way to plan for network infrastructure evolution in the forthcoming decade. This Feature Topic aims to provide a concise reference entry point to a wider audience with respect to carrier-grade networking and service virtualization with emphasis on automating the entire networking and cloud infrastructure. The first steps in network virtualization are already taking place today, although more often than not through a piecemeal approach that simply replaces hardware-based network appliances with software-based alternatives. This may help to reduce operator costs in the mid-term but does not alter the full life cycle of service creation and deployment. The real potential for network and service virtualization lies in upgrading the entire toolbox network operators have at their disposal, as state-of-the-art research and development efforts already indicate. For example, the European FP7 UNIFY project defines an architecture where the entire network, from home devices to data centers, forms a unified production environment, a dynamic service creation platform able to distribute functions and state anywhere in the network, aided by automated orchestration engines; see www.fp7-unify.eu for more details.While drafting the Call for Papers for this Feature Topic, our first goal was to attract high-quality contributions from operator and industry research labs as this topic is particularly pertinent to practitioners in the field. Carriers, in particular, can be the main beneficiaries from the emerging infrastructures based on NFV, SDN, and cloud technologies. Therefore, articles by authors working at global operators currently developing, evaluating, and standardizing solutions for network and service virtualization were particularly welcome and encouraged. In this sense, we are glad that all five selected papers for publication in this issue are penned by experts affiliated with European, American, and Asian carriers.

Our second goal in this Feature Topic is to cover developments in the three major standardization fora: the European Telecommunications Standards Institute (ETSI) NFV Industry Specification Group, the Internet Engineering Task Force (IETF), and the Open Networking Foundation (ONF), as well as articles that capitalize on, and contribute to, major industry-led open-source projects in this area such as OpenStack and OpenDaylight. As such, our intent was to include articles that report on practical experiences with the aforementioned technologies, for instance, by sharing results from proof-of-concept (PoC) implementations.

After a first editorial check to ensure the goodness of fit to the scope of the Call for Papers, 37 manuscripts entered the peer review process. Each paper was reviewed by at least three domain experts, often many more. With the support of the Editor-in-Chief, given the quality of the submitted articles and the increased interest in this area, a second installment on this Feature Topic is scheduled to appear in April 2015.

This Feature Topic opens with the article "Network Functions Virtualization: Challenges and Opportunities for Innovations" by Bo *et al.*, which provides a comprehensive introduction to NFV and can serve as a primer in this area for researchers and practitioners alike. The authors review the key requirements for NFV solutions based on the work undertaken by the ETSI NFV ISG. The article includes a synopsis of the ETSI NFV architecture framework and its application to two prominent use cases, infrastructure virtualization for home access networks and the mobile core network. The article concludes with a discussion of the current research challenges and future directions for NFV. Some of these challenges are tackled in the remaining four articles in this edition of the Feature Topic.

The second article, by Soares *et al.*, delves into the design considerations of SFC for telecom operators. "Toward a Telco Cloud Environment for Service Functions" presents a platform that can enable carriers to provide virtualized service functions through the federated management and joint optimization of WAN and cloud resources. The work considers the developments in IETF with respect to SFC standardization alongside the industry-led CloudStack and OpenDaylight open source projects. The authors model service functions virtualization, and discuss the pros and cons of the two major approaches to SFC packet classification and forwarding. The article concludes with the description of the associated PoC implementation along with the experience of provisioning an access network service function chain comprising virtual customer premises equipment (CPE), VPN server and firewall components.

The third article by Sama *et al.*, "Software-Defined Control of Virtualized Mobile Packet Core," reports on the work currently undertaken within the ONF Wireless and Mobile

Group to address the homonymous use case indentified by the ETSI NFV ISG. Specifically, the article presents the architecture developed in the ONF Mobile Packet Core project and the OpenFlow extensions under consideration for standardization. The core of the article comprises the introduction of a mobile packet core architecture that combines NFV and SDN and ties in with the ETSI Management and Orchestration (MANO). A possible realization of the proposed architecture is subsequently presented based on a layered model compatible with RFC 7426. The article includes a handy comparison of the proposed approach with earlier seminal work in this area.

The fourth article, by Shen *et al.*, "vConductor: Enabler for Achieving Virtual Network as a Service," provides a preview of what is to come in the area of virtual enterprise networks, or what the authors refer to as advanced virtual network integration as a service (VNIaaS). The authors argue that VNIaaS can become a significant new revenue stream for carriers and present how vConductor can be employed as a MANO solution by enterprise customers. The authors illustrate vConductor operation, a turnkey solution that includes a portal for user interaction, network editor, resource inventory, and service monitoring functionality along with its integration with the OpenDaylight and OpenStack network and compute resource provisioning frameworks. The data model used for resource provisioning is discussed in detail, along with the implementation of the associated PoC prototype.

The final article in this edition, "An Instrumentation and Analytics Framework for Optimal and Robust NFV Deployment" by Veitch, Mcgrath, and Bayon, describes a framework for testing, instrumenting, and executing local analytics tasks on telecom-grade virtual network functions (VNFs) executed on general-purpose compute platforms. The framework allows customization and configuration for particular VNFs, and the authors show how it can useful for both developers, who collect results from automated testing, and operators, who investigate the true cause of performance degradation during production deployments. The power of the visualization and analytics framework is illustrated in two use cases. The first use case investigates how an operator can use the proposed framework to determine the optimal resources needed for a virtual CDN deployment .The second use case illustrates the impact of instrumentation and analytics in correlating information during a fault diagnostic process for a WAN acceleration VNF.

As we have seen, virtualization technologies are closely related to software-defined approaches to the management and control of virtualized infrastructure and services. Virtualized infrastructure and services are expected to be integrated in self-adaptation processes that adjust in moments to the quantity of resources consumed according to user demands. From a carrier operations perspective, widespread use of self-adaptation mechanisms changes the nature of the tasks in which human operators are engaged. On one hand, many tasks are simplified through automation, which capitalizes on new software-defined capabilities of operation support systems (OSSs). On the other hand, the focus of the tasks shifts as human operators play a key role in solving more complex problems. In this context, operators will need new tools in this effort as the bar of expectations raises many-fold in terms of reaction times compared to the current generation of network deployments. Expert knowledge of the actual infrastructure elements or services and their interactions constitutes the new baseline but is no longer sufficient. Considering the needs for cross-functional expertise and the necessity of rapid analysis and comprehension of the large amounts of data generated by monitoring processes, we expect that DevOps, as a representative approach that addresses such cross-cultural gaps, will transition swiftly from the domain of IT startups to enterprise network operations and the telecom carrier world.

We hope that academics and practitioners alike will find this Feature Topic handy as they delve deeper into the emergent area of infrastructure network and service virtualization. We close this short introduction by thanking all reviewers for their thorough and constructive comments. We also thank the numerous authors who submitted their work to our Call for Papers and worked diligently to improve their manuscripts throughout the peer-review process. We gratefully acknowledge the magazine's Editor-in-Chief at the time this Feature Topic was being prepared, Dr. Sean Moore, and the Editorial Board for their continuous help. Finally, we say a big thank you to the ComSoc final production editors and staff for their professionalism, and in particular to Charis Scoggins for her guidance throughout the entire process.

BIOGRAPHIES

KOSTAS PENTIKOUSIS (k.pentikousis@eict.de) is the head of IT Infrastructure at the European Center for Information and Communication Technologies (EICT GmbH) in Berlin, Germany, and is currently leading the FP7 UNIFY dissemination and standardization work package. He graduated first in his class at the University of Thessaloniki School of Informatics, and went on to earn his Master's (2000) and doctoral (2004) degrees in computer science from Stony Brook University. His research interests include network architecture, and system and protocol design and implementation. Visit http://linkedin.com/in/kostas for more biographical details and to get in touch.

CATALIN MEIROSU (catalin.meirosu@ericsson.com) is a master researcher with Ericsson Research in Stockholm, Sweden, which he joined in 2007. He holds a Ph.D. in telecommunications (2005) from Politehnica University, Bucharest, Romania, and was a project associate of the ATLAS experiment at the Large Hadron Collider at CERN, Geneva, Switzerland. He has three granted patents and has co-authored over 30 scientific papers. He is currently working to develop DevOps capabilities for service providers in the FP7 UNIFY project.

DIEGO R. LOPEZ (diego@tid.es) joined Telefonica I+D in 2011 as a senior technology expert and is currently in charge of the Technology Exploration activities within the GCTO Unit. He is focused on network virtualization, infrastructural services, network management, new network architectures, and network security. He actively participates in the ETSI ISG on Network Function Virtualization (chairing its Technical Steering Committee), the ONF, and the IETF WGs connected to these activities.

SPYROS DENAZIS (sdena@upatras.gr) received his B.Sc. in mathematics from the University of Ioannina, Greece, in 1987, and in 1993 his Ph.D. in computer science from the University of Bradford, United Kingdom. He worked in European industry for eight years, and is now an associate professor in the Department of Electrical and Computer Engineering, University of Patras, Greece. His current research interests include P2P, SDN, and Future Internet. He is currently technical manager of the STEER EU project. He has co-authored more than 50 papers and is a co-author of the book *Programmable Networks for IP Service Deployment*.

KOHEI SHIOMOTO [M] (shiomoto.kohei@lab.ntt.co.jp) is a senior manager at NTT Network Technology Laboratories. His research fields include IP/MPLS, IP+Optical, GMPLS/PCE, network virtualization, and traffic/QoE management. He has been engaged in standardization at the IETF and the organization of international conferences including WTC, MPLS, and iPOP. He is a Fellow of IEICE and a member of ACM. He co-authored *GMPLS Technologies: Broadband Backbone Networks and Systems*. He received his B.E., M.E., and Ph.D. degrees from Osaka University.

FRITZ-JOACHIM WESTPHAL (fritz-joachim.westphal@telekom.de) has over 25 years of experience in the field of telecommunications. For 15 years he has been a member of different research departments of Deutsche Telekom. His activities are focused on strategies of new network architectures and infrastructure-based services. In 2009 he joined Telekom Innovation Laboratories (T-Labs), and is currently responsible for the research group "Network Architecture & Modelling." He has been involved in different projects on SDN and network virtualization, recently being active in the European FP7 project UNIFY. He has authored or coauthored more than 100 technical conference or journal papers.

Network Function Virtualization: Challenges and Opportunities for Innovations

Bo Han, Vijay Gopalakrishnan, Lusheng Ji, and Seungjoon Lee

ABSTRACT

Network function virtualization was recently proposed to improve the flexibility of network service provisioning and reduce the time to market of new services. By leveraging virtualization technologies and commercial off-the-shelf programmable hardware, such as general-purpose servers, storage, and switches, NFV decouples the software implementation of network functions from the underlying hardware. As an emerging technology, NFV brings several challenges to network operators, such as the guarantee of network performance for virtual appliances, their dynamic instantiation and migration, and their efficient placement. In this article, we provide a brief overview of NFV, explain its requirements and architectural framework, present several use cases, and discuss the challenges and future directions in this burgeoning research area.

INTRODUCTION

It is well known that bringing new services into today's networks is becoming increasingly difficult due to the proprietary nature of existing hardware appliances, the cost of offering the space and energy for a variety of middle-boxes, and the lack of skilled professionals to integrate and maintain these services. Network function virtualization (NFV) was recently proposed to alleviate these problems, along with other emerging technologies, such as software defined networking (SDN) and cloud computing.[1]

NFV transforms how network operators architect their infrastructure by leveraging the full-blown virtualization technology to separate software instance from hardware platform, and by decoupling functionality from location for faster networking service provisioning [3]. Essentially, NFV implements network functions through software virtualization techniques and runs them on commodity hardware (i.e., industry standard servers, storage, and switches), as shown in Fig. 1. These virtual appliances can be instantiated on demand without the installation of new equipment. For example, network operators may run an open source software-based firewall in a virtual machine (VM) on an x86 platform. Recent trials have demonstrated that it is feasible to implement network functions on general-purpose processor-based platforms, for example, for physical layer signal processing [2] and components in cellular core networks [9].

As an innovative step toward implementing a lower-cost agile network infrastructure, NFV can potentially bring several benefits to network carriers, dramatically changing the landscape of the telecommunications industry. It may reduce capital investment and energy consumption by consolidating networking appliances, decrease the time to market of a new service by changing the typical innovation cycle of network operators (e.g., through software-based service deployment), and rapidly introduce targeted and tailored services based on customer needs, just to list a few.

Along with the benefits of NFV, network operators also face several technical challenges when deploying virtual appliances. A frequently raised issue about virtualized network functions (VNFs)[2] is their network performance. Previous work has shown that virtualization may lead to abnormal latency variations and significant throughput instability even when the underlying network is only lightly utilized [14]. Therefore, ensuring that network performance remains at least as good as that of purpose-built hardware implementations will be one of the key challenges in realizing NFV. Besides the network performance issue, another major problem network carriers are confronted with is how to smoothly migrate from the existing network infrastructure to NFV-based solutions, given the former's large scale and tight coupling among its components. Moreover, the separation of functionality from location also creates the problem of how to efficiently place the virtual appliances and dynamically instantiate them on demand.

These facts all impose the need to investigate open research issues brought by NFV in order to ensure its successful adoption. However, there are very limited prior efforts in the literature to offer an overview of aspects to be considered and issues to be addressed when adopting NFV. Our goal is to bridge this gap by identifying critical research challenges involved in the evolution toward NFV.

Bo Han, Vijay Gopalakrishnan, and Lusheng Ji are with AT&T Labs Research.

Seungjoon Lee is currently with Two Sigma Investments, LLC. This work was done when he was employed at AT&T Labs Research.

[1] We discuss the relationship between NFV, SDN, and cloud computing later.

[2] A VNF is the software instance in NFV that consists of some number or portion of VMs running different processes for a network function.

 0163-6804/15/$25.00 © 2015 IEEE

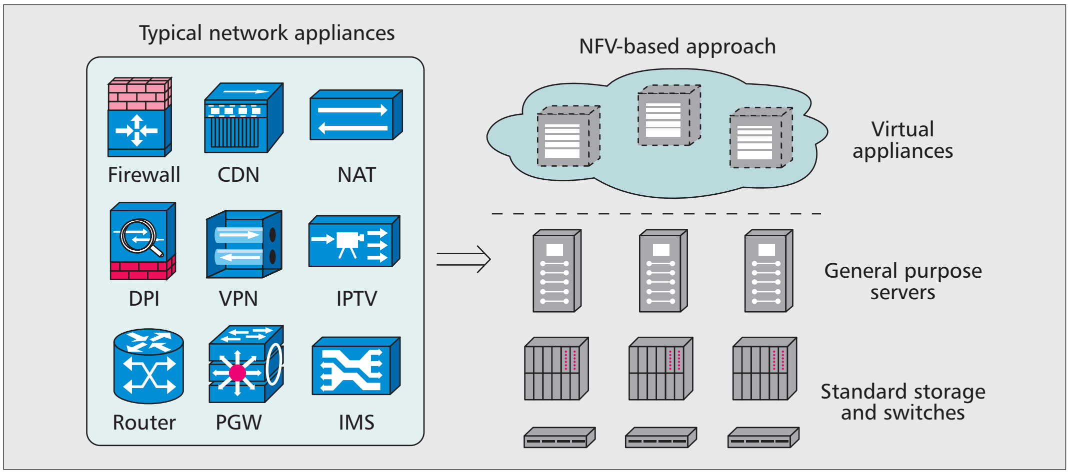


Figure 1. From dedicated hardware-based appliances for network services, such as firewalls, content delivery networks (CDNs), network address translation (NAT), deep packet inspection (DPI), virtual private networks (VPNs), IPTV, routers, packet data network gateways (PDN-GWs or PGWs), and IP multimedia subsystems (IMSs), to software-based NFV solutions.

When talking about software-based implementation of network functions through virtualization technologies on general-purpose servers, the first question we may ask is whether the performance, such as throughput and latency, will be affected.

In this article, we first present the related work and key technical requirements of NFV. We then introduce its architectural framework. We also describe several use cases of NFV, including the virtualization of the cellular core network and home network. Finally, we discuss the open research issues and point out future directions for NFV, focusing on the network performance of virtualized appliances, and their efficient instantiation, placement, and migration.

Related Work

The European Telecommunications Standards Institute (ETSI) has created an Industry Specification Group (ISG) for NFV to achieve the common architecture required to support VNFs through a consistent approach. This ISG was initiated by several leading telecommunication carriers, including AT&T, BT, China Mobile, Deutsche Telekom, Orange, Telefónica, and Verizon. It has quickly attracted broad industry support, and had over 150 members and participants by the end of 2013, ranging from network operators to equipment vendors and IT vendors.

The ETSI NFV ISG currently has four working groups: Infrastructure Architecture, Management and Orchestration, Software Architecture, and Reliability & Availability; and two expert groups: Security and Performance & Portability. Although it is not a standards development organization, it seeks to define the requirements that network operators may adopt and tailor for their commercial deployment. Part of this article (e.g., the architectural framework) is based on the NFV white paper [3] and several related specifications [12, 13] published by this ISG.

Besides ETSI, the Third Generation Partnership Project (3GPP) and Internet Engineering Task Force (IETF) have also been actively involved in NFV. The 3GPP Telecom Management working group (SA5) created a Study Item on the management of virtualized 3GPP network functions. The goal is to investigate whether the architectural framework proposed by ETSI NFV impacts the existing management reference model of 3GPP when all or some instances of 3GPP-defined network elements are virtualized. IETF has formed the Service Function Chaining (SFC) working group to study how to dynamically steer data traffic through a series of network functions, either physical or virtualized. In this article, we review some of the existing work and offer deeper insights on the research challenges of NFV. There are also several multivendor proofs of concept (PoCs) to build the confidence that NFV is a viable technology. For example, CloudNFV[3] is an open platform to implement NFV by leveraging cloud computing and SDN technologies in a multivendor environment. Services, functions, and resources in CloudNFV are represented in an "active virtualization" data model with two key components, the active contract and active resource. When it manages NFV-based services, CloudNFV integrates resource commitments in the active contract with resource state from the active resource.

Technical Requirements

In this section, we summarize the technical requirements when implementing VNFs, including their network performance, and manageability, reliability, and security.

Performance

When talking about software-based implementation of network functions through virtualization technologies on general-purpose servers, the first question we may ask is whether the performance, such as throughput and latency, will be affected. The per-instance capacity of a VNF may be less than the corresponding physical version on dedicated hardware.

Although it is hard to completely avoid performance degradation, we should keep it as small as possible while not impacting the portability of VNFs on heterogeneous hardware platforms. One possible solution is to leverage clustered VNF instances and modern software technologies, such as Linux New API (NAPI)[4] and Intel's Data Plane Development Kit

[3] http://cloudnfv.com/

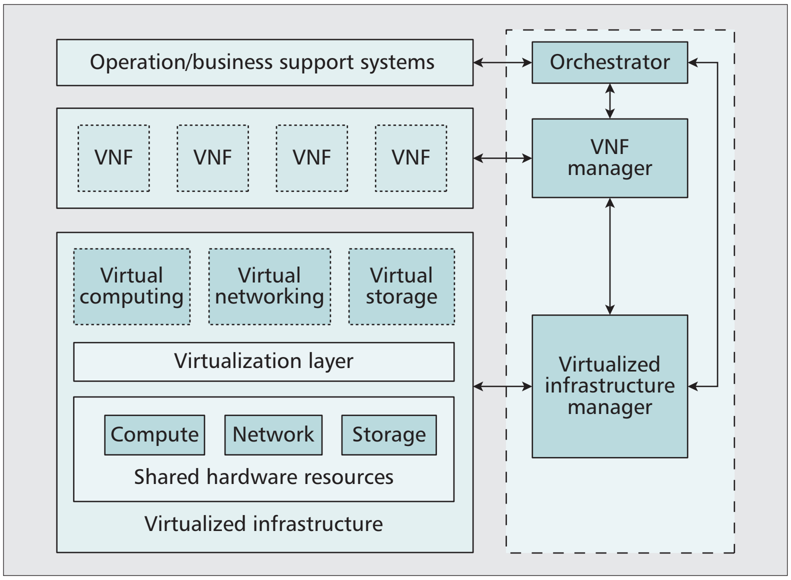


Figure 2. NFV architectural framework [12].

(DPDK).[5] When deploying VNF instances, we need to design efficient algorithms to split network load across a number of distributed and clustered VMs while keeping the latency requirement in mind. Moreover, the underlying NFV infrastructure should be able to gather network performance information at different levels (e.g., hypervisor, virtual switch, and network adapter). We discuss the research challenges related to NFV performance later.

The bottom line is that when designing NFV systems, we should understand the maximum achievable performance of the underlying programmable hardware platforms. Based on this information, we can make the proper design decisions.

Manageability

The NFV infrastructure should be able to instantiate VNFs in the right locations at the right time, dynamically allocate and scale hardware resources for them, and interconnect them to achieve service chaining.[6] This flexibility of service provisioning poses new requirements to manage both virtual and legacy appliances. The manageability in NFV is quite different from that in data center networking, where the hardware resources are almost equivalent, which makes their coordination easier. However, the cost and value of resources may vary significantly between network points of presence and customers' premises. The management functionality should take the variations into account and optimize resource usage across the wide area.

Since service unavailability is typically thought unacceptable, network carriers usually overprovision their services [5]; thus, the utilization of resources allocated to these services is normally low due to the offered redundancy for unexpected traffic increase or service element failure. If we share cloud resources across multiple services, and their failure modes are independent, we can leverage the pool of spare resources to provide the necessary redundancy across them and dynamically create VNFs to appropriately handle traffic increase or failure. In addition, NFV can potentially improve resource utilization through the elasticity feature of cloud computing, for example, by consolidating the workload on a small number of servers during overnight hours and turning the rest off (or using them for services such as online gaming). The management functionality should be able to support sharing spare resources and elastic provisioning of network services effectively.

Although NFV may make planned maintenance relatively easy [15], it presents new requirements for service quality management. Network operators should be able to obtain and process actionable information from various service impacting events, determine and correlate faults, and recover from them by monitoring compute, storage, and network resource usage during the life cycle of a VNF. Since VNFs can be dynamically created/migrated, it brings an additional dimension of complexity in terms of keeping track of where a given VNF is running. Moreover, a VNF can behave erratically even if the underlying infrastructure is running fine, which makes the detection of issues nontrivial.

Reliability and Stability

Reliability is an important requirement for network operators when offering specific services (e.g., voice call and video on demand), whether through physical or virtual network appliances. Carriers need to guarantee that service reliability and service level agreements are not affected when evolving to NFV. Purpose-built network equipment can provide the traditional five-nines reliability in the telecommunications industry. To meet the same reliability requirement, NFV needs to build resilience into software when moving to error-prone hardware platforms. Moreover, as mentioned above, the elasticity of service provisioning may require the consolidation and migration of VNFs based on traffic load and user demand. All these operations create new points of failure that should be handled automatically.

In addition, ensuring service stability poses another challenge to NFV, especially when reconfiguring or relocating a large number of software-based virtual appliances from different vendors and running on different hypervisors. Network operators should be able to move VNF components from one hardware platform onto a different platform while still satisfying the service continuity requirement. They also need to specify the values of several key performance indicators to achieve service stability and continuity, including maximum unintentional packet loss rate and call/session drop rate, maximum per-flow delay and latency variation, and maximum time to detect and recover from failures.

Security

When deploying VNFs, operators need to make sure that the security features of their network will not be affected. NFV may bring new security concerns along with its benefits. Virtual appliances may run in data centers that are not owned by network operators directly. These VNFs may even be outsourced to third parties [11]. The introduction of new elements, such as orchestra-

[4] http://www.linuxfoundation.org/collaborate/workgroups/networking/napi

[5] http://dpdk.org/

[6] Service chaining describes a method for the delivery of network services based on their function associations, and enables the ordering and topological independence of the network functions.

tors and hypervisors, may generate additional security vulnerabilities that increase the load of intrusion detection systems. The underlying shared networking and storage can also introduce new security threats, for example, when running a software router in a VM that shares the physical resources with other network appliances. Moreover, these software-based components may be offered by different vendors, potentially creating security holes due to integration complexity. All these changes require us to rethink security issues when designing and building NFV systems.

DESIGN AND ARCHITECTURAL FRAMEWORK

Virtualization provides us the opportunity for a flexible software design. Existing networking services are supported by diverse network functions that are connected in a static way. NFV enables additional *dynamic* schemes to create and manage network functions. Its key concept is the VNF forwarding graph, which simplifies the service chain provisioning by quickly and inexpensively creating, modifying, and removing service chains. On one hand, we can compose several VNFs together to reduce management complexity, for instance, by merging the serving gateway (SGW) and PGW of a 4G core network into a single box. On the other hand, we can decompose a VNF into smaller functional blocks for reusability and faster response time. However, we note that the actual carrier-grade deployment of VNF instances should be transparent to end-to-end services.

Compared to current practice, NFV introduces the following three major differences [12]:

- *Separation of software from hardware*: This separation enables the software to evolve independent from the hardware, and vice versa.
- *Flexible deployment of network functions*: NFV can automatically deploy network-function software on a pool of hardware resources that may run different functions at different times in different data centers.
- *Dynamic service provisioning*: Network operators can scale the NFV performance dynamically and on a grow-as-you-need basis with fine granularity control based on the current network conditions.

We illustrate the high-level architectural framework of NFV in Fig. 2. Its four major functional blocks are the orchestrator, VNF manager, virtualization layer, and virtualized infrastructure manager. The *orchestrator* is responsible for the management and orchestration of software resources and the virtualized hardware infrastructure to realize networking services. The *VNF manager* is in charge of the instantiation, scaling, termination, and update events during the life cycle of a VNF, and supports zero-touch automation. The *virtualization layer* abstracts the physical resources and anchors the VNFs to the virtualized infrastructure. It ensures that the VNF life cycle is independent of the underlying hardware platforms by offering standardized interfaces. This type of functionality is typically provided in the form of virtual machines (VMs) and their hypervisors. The *virtualized infrastructure manager* is used to virtualize and manage the configurable compute, network, and storage resources, and control their interaction with VNFs. It allocates VMs onto hypervisors and manages their network connectivity. It also analyzes the root cause of performance issues and collects information about infrastructure fault for capacity planning and optimization.

As we can see from this architectural framework, the two major enablers of NFV are industry-standard servers and technologies developed for cloud computing. A common feature of industry-standard servers is that their high volume makes it easy to find interchangeable components inside them at a competitive price, compared to network appliances based on bespoke application-specific integrated circuits (ASICs). Using these general-purpose servers can also reduce the number of different hardware architectures in operators' networks and prolong the life cycle of hardware when technologies evolve (e.g., running different software versions on the same platform). Recent developments of cloud computing, such as various hypervisors, OpenStack, and Open vSwitch, also make NFV achievable in reality. For example, the cloud management and orchestration schemes enable the automatic instantiation and migration of VMs running specific network services.

NFV is closely related to other emerging technologies, such as SDN. SDN is a networking technology that decouples the control plane from the underlying data plane and consolidates the control functions into a logically centralized controller. NFV and SDN are mutually beneficial, highly complementary to each other, and share the same feature of promoting innovation, creativity, openness, and competitiveness. These two solutions can be combined to create greater value. For example, SDN can support NFV to enhance its performance, facilitate its operation, and simplify the compatibility with legacy deployments. However, we emphasize that the virtualization and deployment of network functions do not rely on SDN technologies, and vice versa.

NFV enables additional dynamic schemes to create and manage network functions. Its key concept is the VNF forwarding graph, which simplifies the service chain provisioning by quickly and inexpensively creating, modifying, and removing service chains.

USE CASES

In this section, we describe two use cases of NFV, the virtualization of a mobile core network and a home network. We focus on the problems of existing architecture and the benefits of NFV-based solutions. NFV is applicable to both data plane processing and control plane function. We refer interested readers to the specification of ETSI [13] for more use cases, such as the virtualization of the content delivery network (CDN) and fixed access network.

VIRTUALIZATION OF MOBILE CORE NETWORK

Today's mobile core networks suffer from a huge variety of expensive and proprietary equipment, as well as inflexible hard-state signaling protocols [9]. When a specific function is not available, cellular operators have to replace existing equipment even if it is still sufficient for most purposes, which reveals the difficulty of

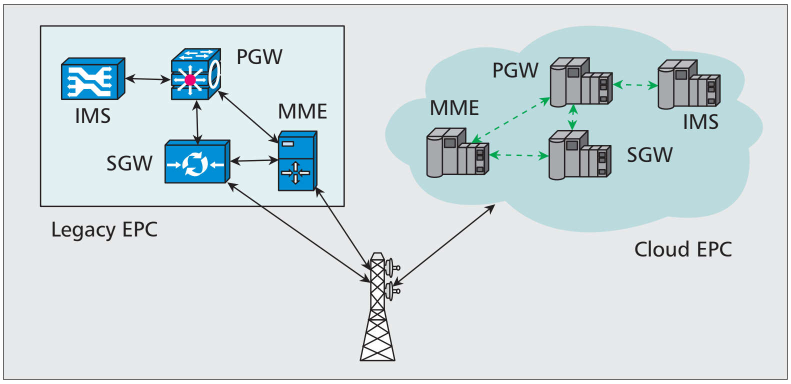


Figure 3. Virtualization of EPC and its coexistence with legacy EPC.

scaling offered services up and down rapidly as required. Moreover, the mobile core network leverages the tunneling mechanism over lower-layer transport protocols to and from a few centralized gateways (PGWs in case of 4G Evolved Packet Core, EPC) for the delivery of user data traffic. These long-distance permanent tunnels are very expensive to control and maintain for cellular operators.

Cloud EPC can potentially address these problems by virtualizing the mobile core network to meet changing market requirements. The virtualization targets of EPC include the mobility management entity (MME), home subscriber server (HSS), SGW, PGW, and policy and charging rules function (PCRF). We illustrate the virtualization of EPC for 4G LTE networks and its coexistence with the legacy EPC in Fig. 3. The coexistence is made possible through technologies such as MME pooling. We note that it is possible to virtualize only part of the mobile core network, such as SGW and PGW, and use physical appliances for other components.

Benefits: By virtualizing the aforementioned network functions, Cloud EPC allows us to move toward a more intelligent, resilient, and scalable core architecture. It enables flexible distribution of hardware resources to eliminate performance bottlenecks and rapid launch of innovative services to generate new revenue sources (e.g., machine-to-machine, M2M, communications). The virtualization of EPC frees distributed network resources from their geographic limitations to ensure service reliability and stability in the event of local resource failure, and reduce the total cost of ownership (TCO). It also makes the flexible deployment of SGW and PGW possible, for example, co-locating them with an eNodeB[7] and thus eliminating long-distance tunnels. With Cloud EPC, cellular carriers can not only expand their current horizontal market business, but also capitalize on previously untouched vertical markets.

Challenges: One of the challenging issues of Cloud EPC is that carriers need to dynamically redirect user traffic when scaling offered services. Early work has shown that SDN could enable the service chaining of various components in cellular core networks [9]. However, as SDN has primarily focused on data center networking in the past, it is still not clear how existing SDN controllers perform in the wide area in terms of scalability and manageability, especially for cellular networks, which have strict latency requirements. Another interesting topic for Cloud EPC is the support of M2M and Internet of Things (IoT) applications where there are a huge number of devices carrying very limited traffic but consuming the bearer resources in the core network.

Virtualization of Home Network

Network service providers offer home services through dedicated customer premises equipment (CPE) supported by network-located back-end systems. Typical CPE devices include residential gateways (RGs) for Internet access and set-top boxes (STBs) for multimedia services. Under this architecture, the delivery of time-shifted IPTV services is known to be complicated due to the interactive stream control functions (e.g., rewind and fast forward) [1]. The emerging NFV technology, with the availability of high-throughput last-mile access, facilitates the virtualization of the home network and brings down the complexity of IPTV services.

We depict the architecture of virtualized home networks in Fig. 4. The virtualization targets are STBs and a range of components of RGs, such as firewall, DHCP server, VPN gateway, and NAT router. By moving them to data centers, network and service operators need to provide only low-cost devices to customers for physical connectivity with low maintenance requirements, demonstrated by the three gray boxes at the bottom left corner of Fig. 4. These devices need to provide only layer 2 functionality for Internet access, as the layer 3 and above functions of RGs are moved into the operators' network. We note that with this virtual architecture, it is possible to share some functionalities of RGs and STBs among customers.

Benefits: This virtualized architecture presents numerous advantages to network operators and end users. First, it reduces the operating expense by avoiding the constant maintenance and updating of the CPE devices, and alleviating the call center and product return burdens. Second, it improves the quality of experience by offering near unlimited storage capacity and enabling access to all services and shared content from different locations and multiple devices, such as smartphones and tablets. Third, it allows dynamic service quality management and controlled sharing among user application streams which helps content providers programmatically provision capacity to end users via open APIs. Finally, it introduces new services more smoothly and less cumbersome by minimizing the dependency on the CPE functions.

Challenges: A fundamental issue in this area is the performance of virtualized packet processing on standard high volume servers. To achieve the same or comparable performance in a virtual environment as in bare metal, we need to carefully design the software architecture and configure the system parameters correctly, as indicated in the testing of virtualized Broadband Remote Access Server (BRAS) PoC from Intel.[8] Moreover, security and privacy issue will be another research challenge when sharing virtualized resource among customers to minimize operating cost.

[7] Although similar schemes have been proposed in the context of mobility management in cellular networks, NFV enables the global orchestration of these entities and their flexible migration, which may further improve mobility management.

[8] Available at http://networkbuilders.intel.com/

Discussion

In various areas where NFV is expected to deliver benefits, different network functions may generate different value and face different challenges and difficulties when moving towards virtualization. For example, based on a recent analysis from Ericsson,[9] there may be higher value and less of a challenge to virtualize home network and media distribution network, but lower value and more of a challenge to virtualize access network and core routers. However, we note that the trade-off between value and challenge may change as the underlying technologies evolve.

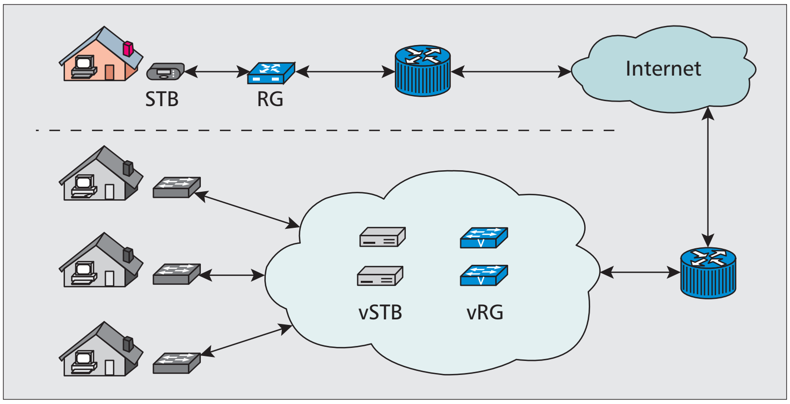


Figure 4. Virtualization of home network.

Research Challenges and Future Directions

In this section, we discuss some of the research challenges and future directions for NFV, including the network performance of virtualization, the placement, instantiation and migration of virtual appliances, and the outsourcing of VNFs.

Network Performance of VNF

The recent effort from the telecommunications industry has been centered on the software virtualization framework (e.g,. management and orchestration). However, it is challenging to offer guaranteed network performance for virtual appliances. Wang and Ng [14] measured the end-to-end networking performance of the Amazon EC2 cloud service. They found that the sharing of processors may lead to very unstable TCP/UDP throughput, fluctuating between zero and 1 Gb/s at the tens of milliseconds time granularity, and the delay variations among Amazon EC2 instances can be 100 times larger than most propagation delays, which are smaller than 0.2 ms, even when the network is not heavily loaded. The unstable networking characteristics caused by virtualization can obviously affect the performance and deployment of virtual appliances.

As mentioned earlier, it may be possible to leverage Linux NAPI and Intel's DPDK to improve the network performance of VNFs. NAPI is a modification of the packet processing framework in Linux device drivers, aiming to improve the performance of high-speed networking. It achieves this goal by disabling some interrupts when the network traffic load is high and switching to polling the devices instead, and thus avoids frequent interruptions sharing the same message that there are lots of packets to process. Another advantage of this polling-based approach is that when the kernel is overwhelmed, the packets that cannot be handled in time are simply dropped in the device queues (i.e., overwritten in the incoming buffer). Intel's DPDK is another software-based acceleration for high-speed networking applications that also uses polling to avoid the overhead of interrupt processing. Recent work by Hwang *et al.* [6] extends the DPDK libraries to provide low latency and high throughput networking in virtualized environments.

Placement of Virtual Appliances

Ideally, network operators should place VNFs where they will be used most effectively and least expensively. Although the virtualization of certain network functions is straightforward, there are a number of network functions that have strict delay requirements. For example, network functions offered by middle-boxes usually depend on the network topology, and these boxes are placed on the direct path between two endpoints. When virtualizing these functions and moving their software implementations into data centers, data traffic may go through indirect paths, causing a potential delay of packets. Therefore, the placement of VMs that carry VNFs is crucial to the performance of offered services. For these services, it would be advantageous and efficient to run some network functions at the edge of the network [7].

Using a mobile core network as an example, we could place a PGW, which currently sits in the cellular core network, right next to an eNodeB, and forward user traffic to the Internet as early as possible. However, the co-location of the PGW and eNodeB will make mobility management difficult, as neighboring eNodeBs will no longer share the same PGW as the anchor point. A possible solution would be to install virtualized PGWs that handle traffic for a small geographical area at the mobile telephone switching office (MTSO) or some other network points of presence in the metro area. Future work regarding low-latency operation should be based on investigation of the redirection architecture and the carrier's data center footprint. Placement problems usually involve optimization through linear programming, integer programming, or a mix, which works on a snapshot of the network and may take a long time to solve an instance. Thus, the online approximation algorithms for these optimization problems are challenging, given the dynamic nature of user traffic.

Instantiation and Migration of Virtual Appliances

Network infrastructure will become more fluid when deploying VNFs. To consolidate VNFs running in VMs based on traffic demand, network operators need to instantiate and migrate virtual appliances dynamically and efficiently. The native solution of running VNFs in Linux or other commodity OS VMs has a slow instantiation time (several seconds) and a relatively large memory footprint. The carrier-grade deployment of VNFs requires a lightweight VM implementa-

[9] http://www-ipv6.ericsson.com/ericsson/industryanalysts/telebriefings/

It is envisioned that NFV, along with cloud computing and SDN, will become a critical enabling technology to radically revolutionize the way network operators architect and monetize their infrastructure. NFV is prospectively the unifying revolution among the three, offering more revenue opportunities in the services value chain.

tion. For instance, Martins *et al.* [8] recently proposed ClickOS, a tiny Xen-based VM to facilitate NFV. ClickOS can be instantiated within around 30 ms and requires about 5 MB memory when running. However, optimizing the performance of this type of lightweight simplified VM, especially during wide-area migration, is still an open research issue.

Take virtual routers as an example, by enabling their free movement, carriers can separate the logical configurations (e.g., packet forwarding functions) from physical routers, and simplify management tasks such as planned maintenance [15]. However, it is challenging to keep the packet forwarding uninterrupted, and the migration disruptions and operating expenses minimized, while at the same time guaranteeing the stringent throughput and latency requirements and other service level agreements. To solve this problem, FreeFlow [10] has been proposed to offer efficient, transparent, and balanced elasticity for virtual middle-boxes, building on top of a state-centric system-level abstraction of network functions. OpenNF [4] is a control plane framework that provides coordinated control of network forwarding state and internal state of network functions through a set of APIs to export and import the middle-box state. A common requirement of these approaches is that they all need to modify the middlebox implementations to achieve efficient migration of virtual appliances. Hence, they cannot be applied to existing implementations as is.

VNF Outsourcing

The end-to-end principle of initial Internet architecture that does not modify packets on the fly is no longer valid in current networks with the deployment of a variety of middle-boxes. Based on a study of 57 enterprise networks with different sizes, ranging from fewer than 1000 hosts to more than 100,000 hosts, Sherry *et al.* [11] found that the number of middle-boxes in a typical enterprise is comparable to its number of hosted routers. In the last five years, surveyed large networks have paid more than US$ 1 million for their middle-box equipment. Moreover, a network with about 100 middle-boxes may need a management team of 100–500 personnel for tasks such as configuration, upgrades, monitoring, diagnostics, training, and vendor interaction [11].

By advocating the split of network functions and their locations, NFV makes the outsourcing of middle-boxes to a third party [11] easier, which may release network carriers from some of the cumbersome operation and maintenance tasks. With the help of *VNF service providers* (e.g., cloud service providers or their partners), end users and small businesses may also be able to enjoy more diverse networking services previously not affordable due to their associated complexity and costs. However, the charging rules and policy interactions between carrier network infrastructure and outsourced VNFs need to be carefully investigated before taking actual actions. Another open question along this direction is to identify what types of VNFs can be outsourced to third parties and how to do it efficiently.

There are also several other open research issues for NFV. For example, using dedicated hardware appliances, it is relatively easy to identify which component is malfunctioning and isolate it when a failure occurs. When deploying network functions in software at different locations, *troubleshooting* and *fault isolation* become harder. Moreover, as the creation of VMs is easy, when the number of VNFs increases, so-called VM sprawl could happen. There may be a large amount of VNFs sprawled across the network even if they are seldom used. As a result, the same management inefficiency problem that NFV was proposed to solve may recur. The efficient *management* and *orchestration* of VNFs, especially in the wide area, is another challenging issue.

Conclusion

In this article, we present an overview of the emerging network function virtualization technology, illustrate its architectural framework, summarize several use cases, and discuss some interesting future research directions. NFV extracts the functionality in specialized appliances and replicates it in virtual form. It is envisioned that NFV, along with cloud computing and SDN, will become a critical enabling technology to radically revolutionize the way network operators architect and monetize their infrastructure. NFV is prospectively the unifying revolution among the three, offering more revenue opportunities in the services value chain. We are looking forward to more initiatives from the networking research community to tackle various challenging issues introduced by NFV and its widespread and successful adoption.

Acknowledgment

We thank the Guest Editors and anonymous reviewers for their insightful comments and suggestions, which improved the overall quality of this article. We thank Jennifer Yates, Changbin Liu, and Feng Qian for discussions and suggestions.

References

[1] V. Aggarwal *et al.*, "Optimizing Cloud Resources for Delivering IPTV Services Through Virtualization," *IEEE Trans. Multimedia*, vol. 15, no. 4, June 2013, pp. 789–801.

[2] China Mobile Research Institute, "C-RAN The Road towards Green RAN," China Mobile White Paper, Oct. 2011.

[3] M. Chiosi *et al.*, "Network Functions Virtualisation: An Introduction, Benefits, Enablers, Challenges & Call for Action," ETSI White Paper, Oct. 2012.

[4] A. Gember-Jacobson *et al.*, "OpenNF: Enabling Innovation in Network Function Control," *Proc. SIGCOMM 2014*, Aug. 2014, pp. 163–74.

[5] A. Greenberg *et al.*, "The Cost of a Cloud: Research Problems in Data Center Networks," *ACM SIGCOMM Computer Commun. Review*, vol. 39, no. 1, Jan. 2009, pp. 68–73.

[6] J. Hwang, K. K. Ramakrishnan, and T. Wood, "NetVM: High Performance and Flexible Networking Using Virtualization on Commodity Platforms," *Proc. NSDI '14*, Apr. 2014, pp. 445–58.

[7] A. Manzalini *et al.*, "Clouds of Virtual Machines in Edge Networks," *IEEE Commun. Mag.*, vol. 51, no. 7, July 2013, pp. 63–70.

[8] J. Martins *et al.*, "ClickOS and the Art of Network Function Virtualization," *Proc. NSDI 2014*, Apr. 2014, pp. 459–73.

[9] K. Pentikousis, Y. Wang, and W. Hu, "MobileFlow: Toward Software-Defined Mobile Networks," *IEEE Commun. Mag.*, vol. 51, no. 7, July 2013, pp. 44–53.

[10] S. Rajagopalan *et al.*, "Split/Merge: System Support for Elastic Execution in Virtual Middleboxes," *Proc. NSDI '13*, Apr. 2013, pp. 227–40.

[11] J. Sherry *et al.*, "Making Middleboxes Someone Else's Problem: Network Processing as a Cloud Service," *Proc. SIGCOMM '12*, Aug. 2012, pp. 13–24.

[12] ETSI, "Network Functions Virtualisation (NFV); Architectural Framework," GS NFV 002 (v. 1.1.1), Oct. 2013.

[13] ETSI, "Network Functions Virtualisation (NFV); Use Cases," GS NFV 001 (v. 1.1.1), Oct. 2013.

[14] G. Wang and T. S. E. Ng, "The Impact of Virtualization on Network Performance of Amazon EC2 Data Center," *Proc. INFOCOM '10*, Mar. 2010, pp. 1163–71.

[15] Y. Wang *et al.*, "Virtual Routers on the Move: Live Router Migration as a Network- Management Primitive," *Proc. SIGCOMM '08*, Aug. 2008, pp. 231–42.

Biographies

BO HAN received his Bachelor's degree in computer science and technology from Tsinghua University in 2000, his M.Phil. degree in computer science from City University of Hong Kong in 2006, and his Ph.D. degree in computer science from the University of Maryland in 2012. He is currently a senior inventive scientist at AT&T Labs Research. His research interests are in the areas of wireless networking, mobile computing, software defined networking, and network functions virtualization, with a focus on developing simple yet efficient and elegant solutions for real-world networking and systems problems.

VIJAY GOPALAKRISHNAN [M] is a director in the Network Evolution Department in AT&T Labs–Research, leading a small team focused on systems challenges in the architecture, protocols, and management of networks. His research interests lie broadly in the area of networked systems where he has worked on topics of network management, content delivery, and the mobile web. Prior to joining AT&T, he got his Master's and Ph.D. degrees in computer science from the University of Maryland, College Park in 2003 and 2006, respectively. He is a member of ACM.

LUSHENG JI received his Ph.D. degree in computer science from the University of Maryland, College Park, in 2001. He is a Principal Member of Technical Staff, Research with the AT&T Shannon Laboratory, Bedminster, New Jersey. His research interests include wireless networking, mobile computing, wireless sensor networks, and networking security.

SEUNGJOON LEE is currently with Two Sigma Investments, LLC. Before joining Two Sigma in 2014, he was with AT&T Research, New Jersey, for more than eight years. He received his Ph.D in computer science from the University of Maryland, College Park in 2006. He also received Bachelor's and Master's degrees in computer science from Seoul National University, Korea, in 1996 and 2000. His research interests include large-scale systems, network management, content distribution, cloud computing, and mobile computing.

Toward a Telco Cloud Environment for Service Functions

João Soares, Carlos Gonçalves, Bruno Parreira, Paulo Tavares, Jorge Carapinha, João Paulo Barraca, Rui L. Aguiar, and Susana Sargento

João Soares is with Ericsson Research.

Carlos Gonçalves is with NEC Laboratories.

Bruno Parreira, Paulo Teveres, João Paulo Barraca, Rui L. Aguiar, and Susana Sargento are with Instituto de Telecomunicações.

Jorge Carapinha is with Portugal Telecom Inovação e Sistemas.

João Soares and Carlos Gonçalves were with Instituto de Telecomunicações at the time this work was carried out.

ABSTRACT

Deploying service functions, SFs, is an essential action for a network provider. However, the action of creating, modifying and removing network SFs is traditionally very costly in time and effort, requiring the acquisition and placement of specialized hardware devices and their interconnection. Fortunately, the emergence of concepts like cloud computing, SDN, and ultimately NFV is expected to raise new possibilities for the management of SFs with a positive impact in terms of agility and cost. From a telco viewpoint these concepts can help to both reduce OPEX and open the door to new business opportunities. In this article, we identify how telcos can benefit from the abovementioned paradigms, and explore some of the aspects that still need to be addressed in the NFV domain. We focus on two major aspects: enabling telco infrastructures to adopt this new paradigm, and orchestrating and managing SFs toward telco-ready cloud infrastructures. The technologies we describe enable a telco to deploy and manage SFs in a distributed cloud infrastructure. In this context, the Cloud4NFV platform is presented. Special attention is given to the way SFs are modeled toward cloud infrastructure resources. In addition, we explore the ability to perform service function chaining as one of the fundamental features in the composition of SFs. Finally, we describe a proof of concept that demonstrates how a telco can benefit from the described technologies.

INTRODUCTION

The emergence of the cloud concept, its ongoing evolution, and the opportunities it brings have led many businesses to adapt in order to get the most utility out of it. One can say that the telco sector is today one of the most active business sectors exploring the opportunities offered by the cloud. The relationship and interdependence between clouds and telecommunications can be analyzed from two distinct perspectives.

Telcos *supporting* the cloud: In a cloud environment, communication endpoints are user devices and virtual machines (VMs) that can be hosted in different physical locations according to varying conditions. Compared to traditional networking environments, network capacity requirements are no longer static, but are likely to change as the associated computing and storage resources expand and reduce. This poses a whole new set of challenges to the network, now jointly including the data center (DC) and the wide area network (WAN) segments. To provide assured levels of performance to cloud services, cloud and telco services need to be provisioned, managed, controlled, and monitored in an integrated way.

Telcos *using* the cloud: Today, the establishment, management, and composition of service functions (SFs) (e.g., router, firewall) follow a rigid, static, and time consuming process. For example, resource overprovisioning is usually necessary to cope with estimated peak demand; and a fault in a single function can disrupt an entire network, imposing the need for faster disaster recovery methods. As virtualization technologies reach maturity and are able to provide carrier-grade performance and reliability, it becomes feasible to consolidate multiple network equipment types, traditionally running on specialized hardware platforms, onto industry standard hardware, which minimizes costs, reduces time to market, and facilitates open innovation. Cloud computing, combined with software defined networking (SDN) [1] and network function virtualization (NFV) [2], promises to make SF management processes much more agile. Cloud computing represents a paradigm for information technology (IT) services, which can now be delivered in an on-demand and self-service manner. SDN brings new capabilities in terms of network automation, programmability, and agility that facilitate integration with the cloud. On the other hand, NFV, from a high-level perspective, accelerates the innovation of networks and services, allowing new operational approaches, novel services, faster service deployment (shorter time to market), increased service assurance, and stronger security.

Conceptually, an SF is a functional block responsible for a specific treatment of received packets and has well defined external interfaces [3]. An SF can be embedded in a virtual instance or directly in a physical element (the usual situation until recently). Virtual SFs offer the opportunity to compose and organize virtual SFs dynamically, opening a new set of business opportunities — and technical challenges. One of the topics that arise from the combination of

 0163-6804/15/$25.00

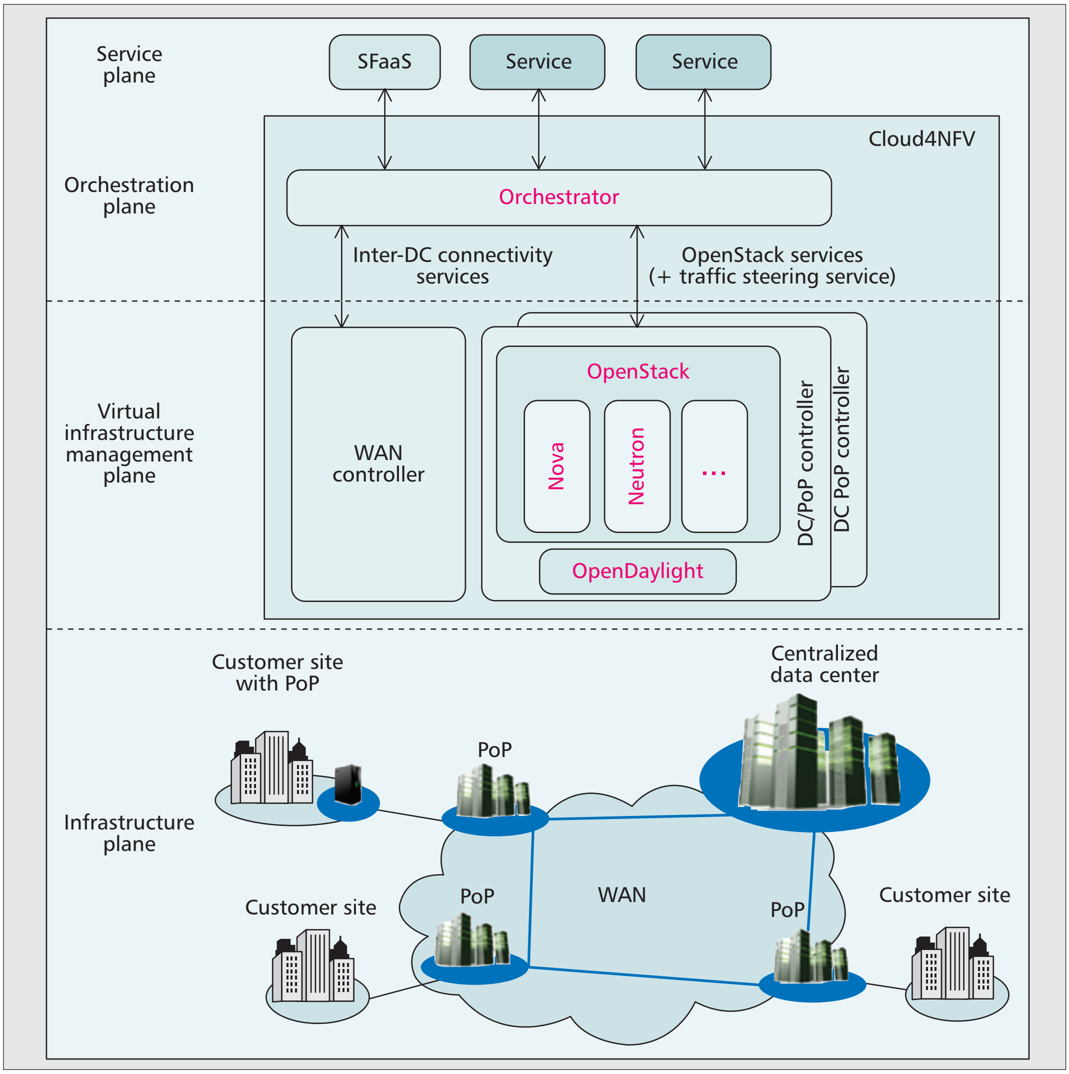


Figure 1. Cloud4NFV platform — overview.

Telcos, with their already established distributed network infrastructure and hosting centers, are ideally positioned to take the lead in this area, as they can easily create a compelling end-to-end cloud proposition that integrates their network management capabilities, adapted to a more agile and cloud service-oriented operation model.

SFs is SF chaining (SFC). SFC is loosely defined as "an ordered set of service functions that must be applied to packets and/or frames selected as a result of classification" [3]. It can be considered as a particular case of service composition. It requires the placement of SFs and the adaptation of traffic forwarding policies of the underlying network to steer packets through an ordered chain of service components. However, the lack of automatic configuration and customization capabilities increases the operational complexity.

In this article, we explore how telecom operators can take advantage of the above concepts to improve the management of SFs and potentially build new business models. First, we highlight the telcos' privileged position in this area compared to traditional cloud providers. We then present Cloud4NFV, a platform for managing SFs in a telco cloud environment. Later, we focus on SF modeling toward cloud infrastructure resources. Special attention is given to the ability to perform SFC. To emphasize possible application scenarios of the solution presented in this article, a proof of concept (POC) is then detailed. Finally, we point out future work directions and conclusions.

The Carrier Cloud Opportunity

Traditional cloud infrastructures are far from suitable for all types of businesses, especially when referring to network SFs. Most network SFs have carrier grade requirements, from guaranteed quality of service (QoS) in terms of IT resources and network connectivity, to high availability (e.g. perform detection and forecast of operational anomalies, support fault mitigation procedures such as VM migration and network replanning) and fast fault recovery through redundancy.

Telcos, with their already established distributed network infrastructure and hosting centers, are ideally positioned to take the lead in this area, as they can easily create a compelling end-to-end cloud proposition that integrates their network management capabilities, adapted to a more agile and cloud service-oriented operation model (on-demand, self-service, elastic).

We envision a near-future telco cloud infrastructure that comprises not only the traditional centralized DC domains, but also the WAN domain. In such a scenario, the telco can take

The Cloud4NFV platform builds upon Cloud, SDN and WAN technologies to allow SFs to be managed on an as-a-Service basis. The platform is targeted for Telcos to improve the management of SFs within their environment, but can also be used to build new services based on the concept of SFaaS.

	Cloud4NFV	UNIFY [4]	T-NOVA [5]	CloNe [6]	StEERING [7]	CloudBand
Distributed infrastructure	Yes	Yes (conceptually)	Yes (conceptually)	Yes	No	Yes
End-to-end service management	Yes	Yes (conceptually)	Yes (conceptually)	Yes	No	Yes
Multi-domain architecture	Yes	Yes (conceptually)	Yes (conceptually)	Yes	No	Yes
Network QoS support	Yes	Yes (conceptually)	Yes (conceptually)	Yes	No	Yes (partially)
SF management	Yes	Yes (conceptually)	Yes (conceptually)	No	Yes (partially)	Yes
Traffic steering support	Yes	Yes (conceptually)	No	No	Yes	No
SFC support	Yes	Yes (conceptually)	No	No	Yes (partially)	No

Table 1. Summary of existing approaches.

advantage of its already established distributed facilities (sometimes referred as points of presence, PoPs) to host small cloud environments. It is also possible for this distributed cloud infrastructure to extend itself into the customer site. Figure 1 depicts this scenario.

Although there are important contributions ongoing in this area, work is still required, namely when it comes to the definition of a true telco cloud platform and the details of how to model and actually realize SFCs. Table 1 presents a summary of the features supported by some existing solutions that more closely relate to the scope of this work. The information presented reflects the publically available information at the time of writing, and may meanwhile have been superseded. The recent UNIFY [4] and T-NOVA [5] projects seem to share a similar vision; however, these projects have recently started and have only provided conceptual approaches to some extent. CloNe [6] has support for the infrastructure features; however, it lacks SF management, traffic steering, and SFC. StEERING [7] supports traffic steering, and partially supports SFC and SF management ("partially" because the SFC service model and SF management features do not seem fully mature). Finally, we also consider the Alcatel Lucent CloudBand[1] solution, which supports some of the envisioned features.

Cloud4NFV Platform

The Cloud4NFV platform builds on cloud, SDN, and WAN technologies to allow SFs to be managed on an as-a-service basis. The platform is targeted for telcos to improve the management of SFs within their environment, but can also be used to build new services based on the concept of SF-as-a-Service (SFaaS), in which case SFs or bundles containing a combination of SFs can be offered as a service to customers.

Functionalities

The most relevant functionalities of Cloud4NFV are:

- Automated deployment, configuration, and life cycle management (instantiation, configuration, update, scale up/down, termination, etc.) of SFs
- Exposure of functionalities such as service deployment and provisioning, service monitoring and reconfiguration, and service teardown
- Federated management and optimization of WAN and cloud resources for accommodating SFs
- Support of SF composition through SFC

All the above mentioned functionalities are essential in the scope of an NFV platform; however, we highlight the last two due to their novelty. These two functionalities are seen as key differentiation factors from other available solutions, taking this platform closer to being fully carrier-grade compliant. The federated management and optimization of WAN and cloud resources gives the platform a broad and distributed scope. It allows the establishment of end-to-end services over a distributed physical infrastructure. The ability to perform SFC gives the platform unprecedented flexibility with respect to SF management and composition, allowing the definition and establishment of advanced services in a much more efficient and flexible way.

Architecture

Figure 1 provides an overview of the system, organized in four major planes: infrastructure plane, virtual infrastructure management (VIM) plane, orchestration plane, and service plane. The service plane handles the services that are built on Cloud4NFV, and the infrastructure plane comprises all physical resources. Special attention should be given to the VIM and

[1] CloudBand, http://www.alcatel-lucent.com/solutions/cloudband.

orchestration planes, since we consider them to be the major lever for enabling SFC. It is important to note that this architecture is aligned with the ETSI NFV architectural guidelines [8]. This fact is highlighted along the description of the platform.

Orchestrator

The orchestrator is responsible for the automated provision, management, and monitoring of SFs over the virtual infrastructure. It exposes the ability to create and delete SFs, as well as the ability to chain SFs. It relies on the VIM plane to provision the infrastructure resources where SFs run (VMs, virtual networks, etc.). Looking to the ETSI NFV reference architectural framework [8], this component considers the *orchestrator* and *VNF manager(s)* entities. The orchestrator has an interface (REST) that exposes the ability to create and delete SFs as well as to chain SFs.

Virtual Infrastructure Management Plane

The VIM plane includes the components for management of infrastructure resources. It includes cloud DC controllers (one per DC) and a WAN controller that is able to establish inter-DC connectivity services. The VIM plane can be seen as the *virtual infrastructure manager(s)* in the European Telecommunications Standards Institute (ETSI) NFV reference architectural framework [8]. However, the current ETSI specification does not take into consideration the WAN component. This is considered by ETSI to be the subject of future analysis.

Data Center Controller(s) — Although the cloud model may require, to a large extent, the redefinition of SFs and the way they are managed, SFs also require adaptation from today's cloud solutions to cope with their requirements, especially in terms of networking features. A clear evidence of this fact is the OpenStack[2] project, a reference open-source cloud management platform, which has been witnessing a tremendous evolution of its networking features in its networking project mostly known by the codename, *Neutron*. It is also important to note that Neutron provides network service logics, and relies on different backends called *drivers* to interact with different networking technologies. Among these drivers is the recent OpenDaylight[3] SDN controller. OpenDaylight is today seen as an initiative equivalent to OpenStack in the SDN domain. With this in mind, our DC controller is based on OpenStack and OpenDaylight.

OpenStack — From a networking perspective, OpenStack allows the creation and management of *networks* (L2 network segments) and *ports* (attachment points for devices connecting to networks, e.g., virtual network interface cards, vNICs, in VMs). The OpenStack community has been making a considerable effort to keep up with users' demands by introducing new Neutron network service types: L3 routing, firewall as a service (FWaaS), load balancer as a service (LBaaS), and VPN as a service (VPNaaS); however, it is infeasible (and probably unwise) in the long run to keep up with demands at this pace in a timely manner. Therefore, we argue that OpenStack should focus on offering the basic tools for network services to be orchestrated at a higher level and be deployed as VMs.

With the orchestration and composition of SFs in mind, it is easy to identify the need to fill a gap in OpenStack: steering traffic between OpenStack elements (e.g. VMs, routers). We envision a new OpenStack service abstraction that extends and relies on current OpenStack networking features, allowing traffic steering between *Neutron ports* according to classification criteria. New entities are introduced into the OpenStack Neutron data model: *port steering* and *classifier*. Both entities have a set of common OpenStack data model attributes (i.e., *id, name*, *description*, and *tenant_id*). Port steering adds to this common set a list of ports (ports attribute) and a list of classifiers (*classifiers* attribute). The former lists the sets of ports that must be targeted for classification and then steered. The classifier entity adds the following attributes: *type*, *protocol*, *port_min*, *port_max*, *src_ip* and *dst_ip*.

This functionality is very useful as it provides the means to realize, among other things, SFC, as described later. Furthermore, the primitive is seen as a foundation for future (higher-level) abstractions within OpenStack.

OpenDaylight — OpenDaylight has a module that integrates with OpenStack Neutron for the enforcement of services in the infrastructure. This module was extended in order to support and enforce the previously mentioned OpenStack traffic steering feature. It is important to highlight that this implementation relies on OpenFlow and Open vSwitch Database Management Protocol (OVSDB) for the management of network resources.

Wide Area Network Controller — The WAN controller is responsible for managing the operator network, and it exposes connectivity services to the upper layers (in this case the orchestrator). In this context, WAN services are used to support SFs (the SF is the client of the WAN service). Point-to-point and multipoint connections with guaranteed network QoS are provided. These are exposed through a service interface that, similar to cloud IaaS interfaces, is technology-agnostic. The details and mechanisms to manage the automatic establishment of connectivity services across different locations are detailed in [6].

Service Function Virtualization

This section elaborates on how SFs are modeled toward virtual infrastructure resources. Figure 2 depicts the correspondent data model, and each class is detailed below.

Service function: represents an instance of a functional block responsible for a specific treatment of received packets.

Service function endpoint (SFE): represents an external interface of one SF instance that is always associated with an SF. Each SFE can have associated information regarding layer 1

The ability to perform SFC gives the platform unprecedented flexibility with respect to SF management and composition, allowing the definition and establishment of advanced services in a much more efficient and flexible way.

[2] OpenStack, http://www.openstack.org/

[3] OpenDaylight, http://www.opendaylight.org/

(e.g., physical/virtual interface), layer 2 (e.g., medium access control, MAC, address), and/or layer 3 (e.g., IP address), or even regarding higher layers (e.g., HTTP).

From an infrastructure perspective, the resources considered to realize a SF are: *compute instance* (i.e., virtual or physical machines), *image* (disk image), *compute flavor* (hardware specification of a compute instance, i.e., CPU, memory, and root disk), *block storage* (additional disks), *port* (i.e., network interface), *network* (a network segment), and *link* (a connection between two ports from different compute instances), which has an associated *link flavor* (dedicated QoS in terms of bandwidth, delay, and jitter). An SF can be associated with multiple compute instances, while each compute instance has a single image and a single flavor, and can have multiple ports and block storages. A port can only be associated with a single network; however, it can be associated with multiple links. An SFE is directly associated with a port, but not all ports need to map to SFEs.

The network QoS, represented in the model by link and link flavor, is not considered in today's cloud infrastructure systems. However, for a carrier grade cloud this is a must, and OpenStack already has an ongoing project to support it.[4]

Figure 3 presents an example of how several SFs can be composed and organized. Furthermore, it also highlights how SFCs can be built and explored.

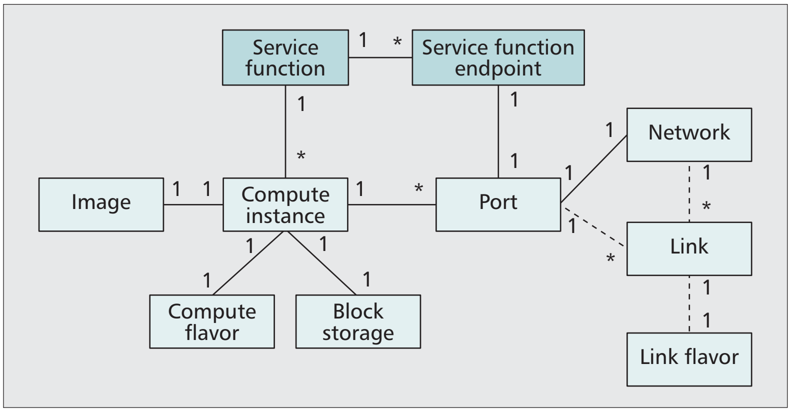


Figure 2. Service function data model toward a cloud infrastructure.

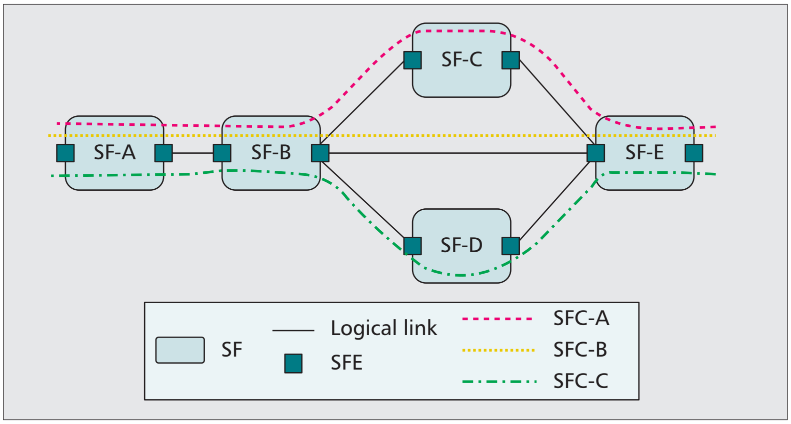


Figure 3. Service function composition — example.

Service Function Chaining

In this section we provide insights on the fundamentals and modeling aspects of SFC proposed in this work.

Fundamentals

In SFC two aspects are vital:

Classification: a policy for matching packets (e.g., HTTP traffic) used for the identification of appropriate actions (e.g., forwarding). It can be, for example, an explicit forwarding entry in a network device that forwards packets with a specific IP or MAC address into the SFC. (Re)Classification can also occur at each SF of the SFC independent from the previous SFs. In such cases, multiple classification policy entries should be allowed in an SFC system.

Traffic steering: the ability to manipulate the traffic route at the granularity of subscriber and traffic types [7]. The actual network topology should not be modified to accomplish this.

Moreover, the combination of classification and traffic steering can be done in two ways:

Tagged packet approach: classification can occur only at the initial redirection points to an SFC, if upon this classification packets are tagged. After that, packets are steered to the SFC and routed along it according to the embedded tags.

Non-tagged packet approach: classification occurs not only at the redirection points but also at each hop of the SFC. In this case, packets are not tagged and are subject to classification and steering at each SFC hop.

The consequences of following a tagged or non-tagged packet approach are felt at the VIM plane level. One of the benefits is that this choice is relatively well isolated from the higher planes. We believe the non-tagged approach to be the smoothest approach to follow due to its lower impact on SFs and virtual infrastructure management systems. The advantage of the tagged approach is that the traffic only needs to be classified and tagged (e.g., with a VLAN or other tag) once along the entire SFC. The drawback is that the SFs need to know how to handle the tags (in the simplest case, they should at least ignore them). Although we can add to the platform support for a tagged approach (e.g., classify only at one point, tag, and steer traffic according to the tag), it only makes sense if there is also support at the SF level. Hence, in this work we adopt the non-tagged packet approach.

Further aspects should be taken into account when elaborating an SFC solution, such as:

- No assumption should be done on how functions are deployed, that is, whether they are deployed on physical hardware, as one or more VMs, or any combination thereof.
- An SF can be part of multiple SFCs.
- An SF can be network-transport-independent.
- An SFC allows chaining of SFs that are in the same layer 3 subnet and of those that are not.
- Traffic must be forwarded without relying on the destination address of packets.
- Classification and steering policies should not need to be done by SFs themselves [10].

Service Function Categories

Two categories of SFs have been defined.

Active SFs: those that are in fact part of the main course of a packet, in which case two subtypes are considered:
- Functions that may drop packets or forward them, such as a firewall
- Functions that can actually change packets, such as an IPSec VPN server

Passive SFs: are considered to be out of the main course of the chain. These functions mainly inspect packets (e.g., a monitoring system or a deep packet inspection, DPI). In practice one can think of an SF in a physical device connected to a hub through a single network interface configured in promiscuous mode. Traffic is considered to be duplicated when having to reach a passive function.

These two categories are important because they impose constraints on how classification and steering can be implemented. In short, passive functions can rely on packet characteristics as packets are not modified, while active functions must be integrated at a service level because ingress and egress packets can be different (e.g., virtual private network, VPN). If an SFC has active functions that change packets, the classification may differ when passing one of these functions.

Service Function Chaining Abstraction Model

The ability to classify and steer traffic accordingly can be enough to implement low-level SFC functionality. However, it is important not to forget that the traffic steering functionality is a low-level functionality that does not explicitly express an SFC. Having in mind the considerations made so far, a base data model for SFC (that supports both tagged and non-tagged approaches) is now presented. Naturally, other SFC service abstraction proposals may appear in the future, but we consider that this model lays a strong foundation over which other service abstractions can easily be created by extending the model. Figure 4 depicts the model. Five main classes are considered: *service function chain*, *service function*, *service function endpoint*, *packet flow*, and *classifier*.

All classes have the following attributes: *id*, *name*, and *description*. The id refers to a unique identifier able to identify the class instance within the SFC system. The remaining two, name and description, are attributes that allow a human-readable characterization of the class instance. Below, we provide further detail about each class.

Service function chain: An SFC has a set of associated SFs and an attribute that defines the ordered sequence of functions (path). Since a function can have more than one SFE, the path attribute is specified by an ordered list of SFEs organized by hops. For example,

- "path= { hop={SF-A_E2, SF-B_E1}; hop={SF-B_E2, SF-D_E1}, passive={SF-C_E1}}"

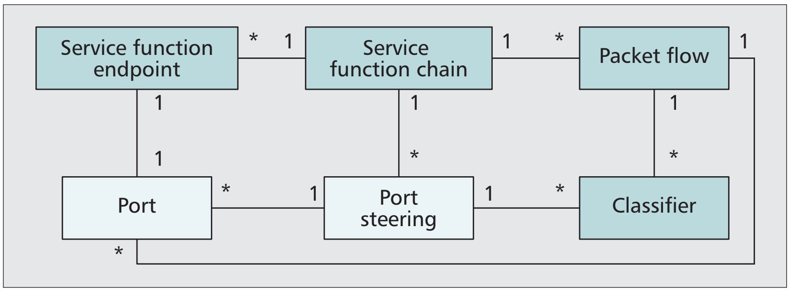


Figure 4. Service function data model toward a cloud infrastructure.

where the chain crosses SF-A, SF-B, and SF-D, and has SF-C as a passive function between SF-B and SF-D.

Classifier: A classifier represents a classification criterion applied to a packet, which determines if the packet matches that specific criterion or not. In this sense, a classifier has an attribute filter that contains the classification criteria;

- "filter={protocol='6'; port='80-90'; source_IP='192.168.10.20/32'; destination_IP='192.168.10.40/32'}" –

matches all TCP traffic using ports between 80 and 90 with source IP address 192.168.10.20 and destination IP address 192.168.10.40.

Packet flow: One *classifier* only identifies packets with a certain criterion, while a *packet flow* identifies a broader set of packets as it can aggregate packets associated with multiple classifiers. In this sense, a packet flow can have multiple classifiers, and a classifier can be associated with multiple packet flows. Moreover, a packet flow has a *source* and a *destination port*. The former identifies where the initial classification and redirection of the packet flow to the SFC takes place, while the latter identifies where packets are to be delivered after passing through the SFC. The attributes considered so far would be enough if the system realizing the SFC followed a tagged packet approach. For a non-tagged approach, an additional attribute is considered — *sfc_classifiers*. Due to the possibility of (active) SFs to modify packets, the classification initially done may not be the same along all hops of the SFC, and therefore, the *sfc_classifiers* attribute matches the classification criteria (classifiers) at each hop of the SFC.

Furthermore, the attribute *direction* is also considered to identify the direction of the SFC in which the *packet flow* must traverse; this attribute can assume one of two values: *forward* and *reverse*. We consider that multiple packet flows can be associated with a single SFC instance.

Port steering: This entity refers to the functionality presented above in OpenStack. This feature allows steering traffic between *ports*. Further details about the traffic steering functionality can be found in the OpenStack proposal,[5] for which we developed a prototype implementation.

In terms of operations, all classes are considered to allow create, read, update, and delete (CRUD) operations.

[4] OpenStack Neutron QoS support https://wiki.openstack.org/wiki/Neutron/QoS

The instantiation and configuration of SFs is done in a timescale of seconds/minutes (depending on the SF/VM and cloud infrastructure) and the SFC enforcement in a timescale of seconds. Furthermore, the user is able to control them through a dedicated SF management portal.

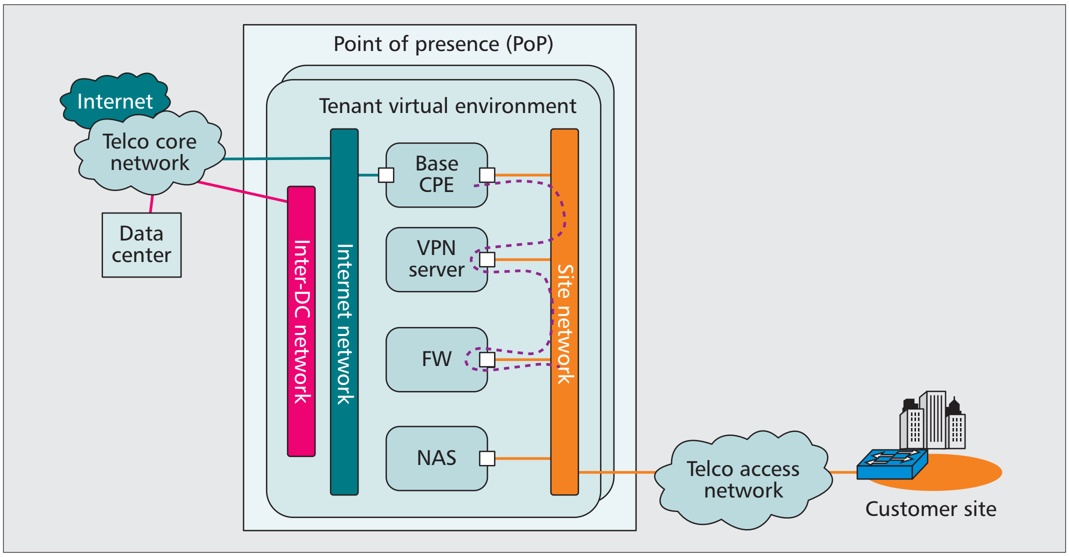


Figure 5. POC prototype setup.

Proof of Concept

A POC environment has been deployed to showcase how a telco can leverage the features described in this work. We highlight one of the most attractive use cases in the NFV scope and how it has been realized in this POC.

The testbed in place is depicted in Fig. 5, focusing on the PoP setup that is detailed later. At the core of the operator network (telco core network) there is an IP/multiprotocol label switching (MPLS) backbone composed of four provider (P) routers and four provider edge (PE) routers. The core network is managed by proprietary operations support systems (OSSs) that expose connectivity services through a service interface in a technology-agnostic manner (the WAN controller). The core network connects to two DC premises (managed by the OpenStack *IceHouse* release with traffic steering functionality), one of which represents a centralized DC and the other a PoP. Finally, the customer premises are represented by switching equipment, which is logically connected to the PoP over an access network (a simple switch-based network).

A prototype of the Cloud4NFV orchestrator, which interacts with the WAN and DC controllers, was developed using the Python language. Details regarding the orchestrator implementation (e.g., RESTful API) can be found in [11].

Customer Premises Equipment Use Case

Customer premises equipment (CPE) is often pointed out as one of the most suitable candidate SFs for virtualization [2, 12]. SFC will surely play a particularly relevant role in this case.

The CPE can be seen as a standard routing node enhanced by collection of SFs, such as network address translation (NAT), firewall (FW), voice over IP (VoIP) servers, VPN servers, network-attached storage (NAS), WAN optimization controllers (WOCs), DPI, or intrusion prevention system (IPS). These services are deployed for different scenarios, and not all traffic needs to traverse them, leaving room for optimization through SFC. It should be noted that some of the chains can even be temporary, which requires a model that enables the dynamic definition of chains.

Service Function as a Service

At the service layer we implemented a prototype of the SFaaS concept. This is exposed via a web portal. CPE functions are available in the SFaaS, and the ability to perform SFC is not exposed to the end user. The user requests CPE SFs, which already have a predetermined relation with other SFs, and associates them with one of the user's sites. The instantiation and configuration of SFs is done on a timescale of seconds/minutes (depending on the SF/VM and cloud infrastructure), and the SFC enforcement on a timescale of seconds. Furthermore, the user is able to control them through a dedicated SF management portal.

Note that the use of SFaaS requires a basic business relationship between the customer and the telco (customer sites registered and with connectivity services). In other words, the user is a customer of the telco that provides connectivity services (e.g., fiber, copper) from the client's sites (e.g., house, enterprise premises). On the site side, a layer 2 device (or a layer 3 device in bridge mode) is considered to be in place.

Currently, from a demonstration viewpoint, it is considered that the user, after having the physical connection in place, must first buy a base CPE function with routing, Dynamic Host Control Protocol (DHCP), and NAT functionalities (the POC relies on the OpenStack L3 native device). From that moment on, the user can acquire other CPE functions and services, such as Internet connection, firewall (POC relies on iptables), VPN server (POC relies on OpenVPN), and NAS (POC relies on Samba).

Prototype Setup

Figure 5 depicts the POC prototype setup with four functions as an example. Special attention is given to the setup at the PoP level.

[5] OpenStack Traffic Steering blueprint, https://review.openstack.org/#/c/92477/

Each customer has a dedicated virtual private environment in the PoP that is serving his/her site. This environment allows the creation of virtual networks and VMs (in OpenStack this is known as *tenant* or *project*). There is a point-to-point logical connection between the customer's premises (L2) device (currently we are using VLAN encapsulation to establish this connection, but others can be used). On the PoP side this logical connection is extended to a virtual network in the tenant virtual environment; in the figure, this is "site network," which has a private IP range (the OpenStack *provider network* concept is used to achieve this). Moreover, there is a virtual network shared among all tenants, which in the figure is the "Internet network" (in OpenStack this is achieved using the *external network* concept). This latter network is then connected to the core network, which provides the Internet access. Also depicted in the figure is the "inter-DC network," which provides access between the PoP and the DC over a telco VPN service in the core network (again, on the PoP side we rely on the OpenStack *provider network* concept to connect to the VPN). The processes explained so far are considered to be in place as soon as the customer establishes the basic business relationship with the telco.

All functions, when deployed upon request, are connected to the site network. When an Internet connection is requested, the base CPE is connected to the Internet network and configured to perform NAT. The figure also highlights an SFC that comprises the base CPE, VPN server, and firewall.

FUTURE WORK

Currently, the POC does not support the enforcement of network QoS in DC domains; this is only supported in the WAN connectivity services. We expect to add this support by the time OpenStack officially releases this feature. Furthermore, runtime management operations (e.g., scaling and migration of SFs) are yet to be included in the platform. On the WAN domain, we are currently adding an SDN-based network. The purpose is to have both legacy and SDN network technologies in place to better evaluate the advantages and disadvantages of each approach. Finally, we are working on exposing the ability of performing SFCs to the end user.

CONCLUSIONS

The orchestration and management of SFs is today a complex task that takes considerable time and effort. However, concepts like cloud computing, SDN, and NFV are paving the way to handling SFs in a much more flexible and agile manner. The telco will play a key role in this scenario, and we have given some insights on how that can be performed in the near future. Special attention has been given to the modeling of SFs toward cloud resources and to the combination of SFs through SFC. Finally, a platform for managing virtual SFs in a telco cloud infrastructure has been presented and a POC described that showcases how the platform and the principles here presented can be leveraged in a telco environment.

REFERENCES

[1] H. Kim and N. Feamster, "Improving Network Management with Software Defined Networking," *IEEE Commun. Mag.*, vol. 51, no. 2, Feb. 2013, pp. 114–19.

[2] ETSI, "Network Functions Virtualisation (NFV): Use Cases," tech. rep. ETSI GS NFV 001 v1.1.1, Oct. 2013.

[3] P. Quinn, Kumar, and T. Nadeau, "Service Function Chaining Problem Statement," IETF Internet draft, informational, Dec. 2013.

[4] A. Császár *et al.*, "Unifying Cloud and Carrier Network: EU FP7 Project UNIFY," *IEEE/ACM 6th Int'l. Conf. Utility and Cloud Computing*, Washington, DC, 2013.

[5] G. Xilourisl *et al.*, "T-NOVA: A Marketplace for Virtualized Network Functions," *EUCNC '14*, June 2014.

[6] H. Puthalath *et al.*, "Negotiating On-Demand Connectivity between Clouds and Wide Area Networks," *IEEE CloudNet*, Paris, France, Nov. 2012.

[7] Y.Zhang *et al.*, "StEERING: A Software-Defined Networking for Inline Service Chaining," *Proc. IEEE ICNP '13*, Göttingen, Germany, Oct. 2013.

[8] ETSI, "Network Functions Virtualisation (NFV): Architectural Framework," tech. rep. ETSI GS NFV 002 v. 1.1.1, Oct. 2013.

[9] ETSI, "Network Functions Virtualisation (NFV): Terminology for Main Concepts," tech. rep. ETSI GS NFV 003 v. 1.1.1, Oct. 2013.

[10] W. John *et al.*, "Research Directions in Network Service Chaining," *2013 IEEE SDN for Future Networks and Services*, 11–13 Nov. 2013, pp. 1–7.

[11] J. Soares *et al.*, "Cloud4NFV: A Platform for Virtual Network Functions," *IEEE CloudNet*, Oct. 2014.

[12] Alcatel Lucent, "Network Functions Virtualization — Challenges and Solutions," White Paper, 2013.

BIOGRAPHIES

JOÃO SOARES received an M.Sc. degree in electronics and telecommunications engineering from the University of Aveiro in 2009, and a Ph.D degree in 2015. He initiated his professional activity at the Institute of Telecommunications in 2009 and joined Portugal Telecom Inovação e Sistemas in 2010. He recently joined Ericsson Research as an experienced researcher in the area of cloud technologies. His interests cover the areas of cloud computing, cloud networking, SDN, and network virtualization.

CARLOS GONÇALVES received his Master of Science degree in computers and telematics engineering from the University of Aveiro in 2013. Currently, he is a research associate at NEC Laboratories Europe working in the areas of network function virtualization and carrier-cloud operation and management. He has participated in several open-source projects including OpenStack and Open Platform for NFV, serving as a member of the OpenStack Networking Advanced Services and Telco Working Groups.

BRUNO PARREIRA received his M.Sc. degree in electronic and telecommunications engineering from the University of Aveiro in 2012. He started his professional activity at the Institute of Telecommunications through a research scholarship in May 2012. He has participated in several European funded projects related with cloud networking. His main interests are: cloud computing, cloud networking, software defined networking, and network function virtualization.

PAULO TAVARES received his M.Sc. degree in engineering of computers and telematics from the University of Aveiro in 2011. He started his professional activity at the Institute of Telecommunications through a research scholarship in September 2011. His main interests include cloud computing, cloud networking, software defined networking, and network function virtualization.

JORGE CARAPINHA graduated in electronic engineering from the University of Coimbra in 1984 and got an M.Sc. degree in telecommunications from the University of Aveiro in 1998. He has been with Portugal Telecom Inovação e Sistemas (formerly CET) since 1985. Currently his main fields of interest are network virtualization, cloud networking, and software defined networking.

JOÃO PAULO BARRACA received a Ph.D. degree in informatics engineering from the University of Aveiro, where he developed work focused on network management functions. He is currently an invited lecturer at the University of Aveiro

Currently, the POC does not support the enforcement of network QoS in DC domains; this is only supported in the WAN connectivity services. We expect to add this support by the time OpenStack officially releases this feature. Furthermore, runtime management operations are yet to be included in the platform.

and a researcher at the Institute of Telecommunications in areas related to programming, networking, and security. He has published more than 40 papers in the areas of networking and computer systems, and has acted as a reviewer for tens of events and journals.

RUI L. AGUIAR [SM] received a Ph.D. degree in electrical engineering in 2001 from the University of Aveiro, where he is currently a professor. He is leading a research team at the Institute of Telecommunications, and is an invited researcher at Universidade Federal de Uberlandia, Brazil. His current research interests are centered on the implementation of advanced networks and systems with special emphasis on future Internet and 5G architectures, and he is currently involved in the 5G-PPP initiative. He is a member of ACM, with more than 350 published papers. He has served as Technical and General Chair of several conferences, such as (recently) Monami '12, ISCC '14, NTMS'2014, and Mobiarch '14. He has been invited as a keynote speaker to several events, both technical and for large generalist audiences.

SUSANA SARGENTO has been with the University of Aveiro and the Institute of Telecommunications since February 2004, where she leads the Network Architectures and Protocols (NAP) group (http://nap.av.it.pt). She is also a cofounder of Veniam (www.veniam.com), a spin-off that commercializes vehicular technology. She has been involved in several national, U.S., and European projects, taking leadership of several activities in projects, such as the QoS and ad hoc networks integration activity in the FP6 IST-Daidalos Project, and the deployment of 600 nodes in a vehicular network in the EU Future Cities. Her main research interests are in the areas of future networks, more specifically routing, QoS, mobility, and cloud integration.

Software-Defined Control of the Virtualized Mobile Packet Core

Malla Reddy Sama, Luis M. Contreras, John Kaippallimalil, Ippei Akiyoshi, Haiyang Qian, and Hui Ni

Malla Reddy Sama is with Orange Labs.

Luis M. Contreras is with Telefónica I+D.

John Kaippallimalil and Hui Ni are with Huawei.

Ippei Akiyoshi is with NEC.

Haiyang Qian is with China Mobile.

ABSTRACT

Mobile packet core networks are undergoing major changes to meet the requirements of the future data tsunami, enhance network flexibility, and reduce both CAPEX and OPEX. In this regard, SDN and NFV technologies have gained great momentum among the telcos, with the promise of interoperability, programmability, and on-demand dynamic provisioning. In this article, we present work being developed in the Mobile Packet Core project within the ONF Wireless & Mobile Working Group regarding SDN architecture for the Mobile Packet Core. In addition, we propose an SDN-based MPC in the NFV context in order to facilitate dynamic provisioning of MPC network functions. Finally, a potential control architecture considering both SDN and NFV is proposed.

INTRODUCTION

Over the last few years, cloud and virtualization technologies have gained great momentum among mobile operators and network equipment vendors. This is due to the flexibility these technologies introduce in the network, offering significant economic potential to reduce both capital expenditure (CAPEX) and operational expenditure (OPEX) costs. In addition, operators are also given the privilege to program the network functions (e.g., gateways, routers, load balancer) independent of the proprietary hardware substrate. This has not been the case over recent decades. For instance, operators could not change any software function or implement their own software function on vendor-specific hardware. Instead, they had to wait for vendors' developments to apply changes.

Innovations in the mobile field such as powerful new terminals (e.g., smartphone and tablets) and the proliferation of data-hungry mobile applications (mobile health, mobile education, context-awareness applications, cloud communication etc.) have dramatically increased data traffic usage because of the demand for rich and sophisticated services. It is estimated that mobile traffic will grow at a compound annual growth rate (CAGR) of 61 percent from 2013 to 2018 [1]. To sustain this data tsunami, operators are required to deploy a wide variety of hardware appliances such as gateways, routers, servers, intrusion detection systems and firewalls. These appliances run on dedicated proprietary hardware placed in different locations in the network. In the traditional mode of operation, telcos have populated the network with these monolithic physical devices, resulting in difficulties in rolling out new devices and allocating the required resources in a cost-effective manner. Under these circumstances, operators have to repeatedly invest to cope with severe data usage increments, resulting in increasing the network CAPEX in a context where the average revenue per user (ARPU) is declining.

In addition, the mobile network entities in the Evolved Packet Core (EPC), that is, the mobility management entity (MME), serving gateway (SGW), packet data network gateway (PGW), and home subscriber server (HSS), follow the same model, which is based on customized hardware that requires static deployment, provisioning, and configuring. As a consequence, the network architecture does not inherently define any flexibility, dynamicity, or on-demand features. Those network entities are tightly coupled in two dimensions: (i) software and hardware; and (ii) control and user planes. For example, to provision network capacity, operators typically dimension both the control and user plane capacity based on the load foreseen in peak hours. User plane processing is I/O bound and requires high capacity, while control plane processing is CPU bound. Significantly, static provisioning of both planes is not an optimal solution, causing waste of truly expensive resources.

On these ground, network function virtualization (NFV) and software defined networking (SDN) emerge as the latest incarnation of technological promises for reaching the necessary cost efficiency. SDN decouples the control and user planes, and logically centralizes the network intelligence (i.e., control plane), while the underlying network infrastructure (i.e., user plane) is abstracted for external applications requesting services through that control plane. This mechanism facilitates the programmability of the network resources by automatically and dynamically allowing the control of underlay network switches. The OpenFlow protocol is a key component of the SDN concept to model flow abstractions. In this context, last year, the Open Networking Foundation (ONF) chartered the Wireless &

 0163-6804/15/$25.00

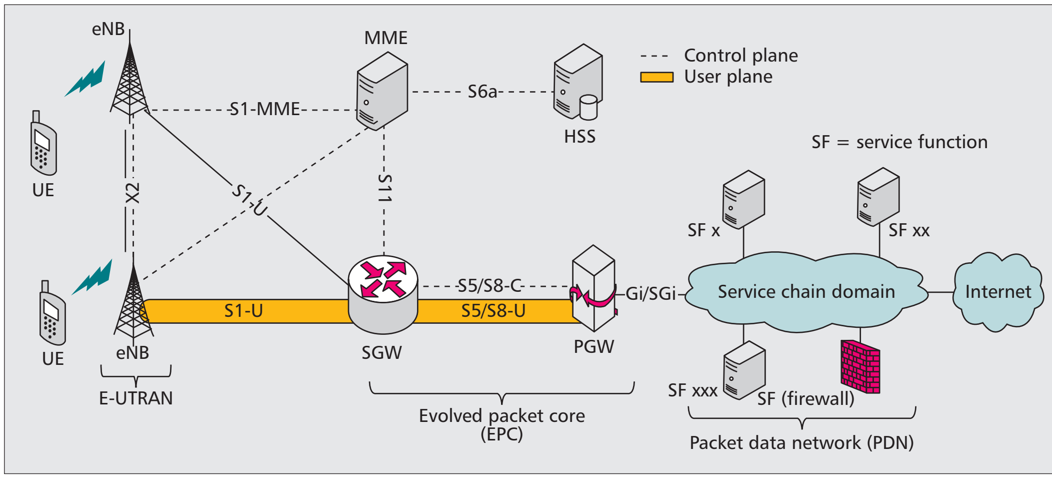


Figure 1. A mobile operator network.

The LTE/EPC architecture has been widely accepted. However, the architecture was not designed with elasticity in mind. Functionality is deployed on specific network entities, which execute specific functions. In addition, the current network entities are too prone to vendor locking, too complicated to manage, and too hard to change in behavior.

Mobile Working Group (WMWG)[1] to foster the adoption of OpenFlow-based SDN technology in mobile and wireless networks. This group studies and proposes simplified reference mobile architectures for transport by leveraging OpenFlow-based SDN. In addition, this group aims to simplify the interaction between wireless networks and fixed networks (i.e., physical layer to upper layers interaction and vice versa). Moreover, it is also responsible for proposing extensions to the OpenFlow protocol specification, such as extensions required to transfer the general packet radio service (GPRS) Tunneling Protocol (GTP) tunnel-ids to the switches, and the parameters for interaction between the control and user planes.

On the other hand, the NFV initiative was created by a group of tier-1 telcos within the European Telecommunications Standards Institute's (ETSI's) Industry Specification Group and released a white paper in October 2012 [3]. The scope of NFV is to propose consistent and common architecture to decouple mobile network function (e.g., firewalling, NAT, gateways) from proprietary hardware, making it possible to run such functions in general-purpose commodity servers, switches, and storage units, which could be deployed in an operator's data center. The main expected benefits with NFV are lower OPEX, greater flexibility (i.e., easily scaling network resources up and down), and easier network management.

This article addresses the evolution of the existing mobile packet core (MPC) architecture toward a new architecture where network functions are deployed in a virtualized manner with the separation of the control and user planes using a programmable infrastructure.

Software Defined Mobile Packet Core

LTE/EPC Architecture

The Third Generation Partnership Project (3GPP) Long Term Evolution/Evolved Packet Core (LTE/EPC) architecture has been designed to provide seamless IP connectivity between user equipment (UE) and external packet data networks (PDNs) as shown in Fig. 1. The EPC has a flat all-IP architecture [4] composed of a number of entities. The MME acts as the manager of network connectivity, and also for UE authentication and authorization, UE session setup, and intra-3GPP mobility management. The SGW routes data packets through the access network and is the local mobility anchor point for inter-eNB handover. The PGW is responsible for user plane quality of service (QoS) management and is the anchor point for external PDN Networks. Both GTP and proxy mobile IP (PMIP) are the main communication protocols within the LTE/EPC architecture.

The LTE/EPC architecture has been widely accepted. However, the architecture was not designed with elasticity in mind. Functionality is deployed on specific network entities, which execute specific functions. In addition, the current network entities are too prone to vendor locking, too complicated to manage, and too hard to change in behavior. With the introduction of SDN and NFV, greater flexibility can be achieved by network operators. In the first case, by usage of SDN capabilities, dynamic control of traffic flows can be performed, redirecting the traffic to gateways according to, say, workloads. In the second case, the introduction of NFV permits the separation of service functionalities from the capacity constraints limited boxes could suffer by allowing dynamic instantiation in commodity and powerful servers.

The SDN Approach to the LTE MPC

The Mobile Packet Core (MPC) project within the ONF is identifying the OpenFlow (OF) protocol extensions needed to support handling of packet data network (PDN) connections. Such extensions should be able to provision and manage the establishment, modification, and deletion of PDN connections while providing support for QoS, accounting, and online charging, which translates into programmable rate limiting, metering, and rating functionality configuration in the supporting switch. Forwarding abstractions for PDN connections in an SGW, a PGW, and a combined gateway (SGW and PGW combination) are being defined. These abstractions model the PDN connection segments using

[1] ONF Wireless and Mobile Group, https://www.opennetworking.org/technical-communities/areas/specification

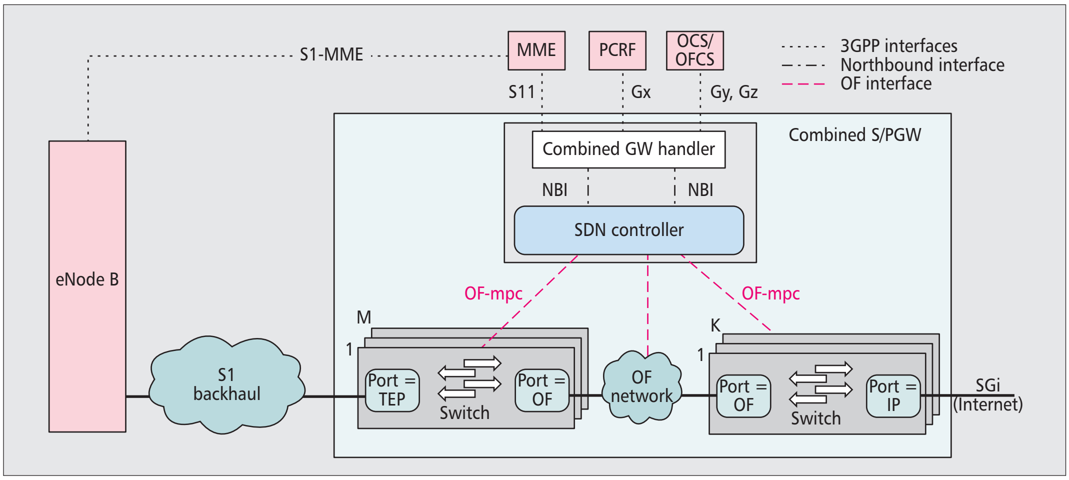


Figure 2. Combined SGW and PGW in the MPC architecture.

3GPP/LTE gateways need to support a number of capabilities, including IP address allocation and management, gating and rate control, event reporting and bearer binding, and support for handovers with end-markers and buffering in the SGW.

ports, tables, and control signaling in OpenFlow. On the other hand, no changes are necessary to standard 3GPP defined interfaces.

The SDN controller in this case may also be implemented as a plug-in in an NFV realization of 3GPP control functions. A description of an OF based combined SGW and PGW gateway and its relation to 3GPP interfaces is shown Fig. 2. Note that the dotted lines refer to control and signaling interface while solid lines represent data interface.

Figure 2 shows a combined SGW and PGW node with an SDN controller and OpenFlow channel (OF-mpc) extensions to control the switch. The necessary OF-mpc extensions are currently being defined by the ONF W&MWG with final specification expected during 2015. The combined gateways are common in current LTE deployments. However, there are scenarios where even in that case the S5 interface has to be implemented (not shown in Fig. 2 for simplicity). For instance, a roaming mobile terminal may attach via the SGW functionality of a combined gateway of the roaming network, and to a PGW in the home network via the S5 interface.

The combined gateway handler is the control software of the combined gateway and interfaces to the other 3GPP functional entities. In an NFV-based deployment with virtual machines (VMs), both the combined gateway handler and SDN controller may be implemented as a VM. The northbound interface (NBI) in Fig. 2 can be realized through controller plugins. Since the control plane scaling is critical in mobile infrastructure, the SDN controller itself may be implemented as a cluster.

Forwarding abstractions for the PDN connection consist of ports that serve as network interface points to connect the S1 interface (toward the eNB) and the SGi (connecting to external networks), as well as flow and group tables for OF switch components. Packets received on the ingress port are processed through a pipeline of flow tables that route them to the egress port. The SDN controller can manage multiple switches. In addition, these switches are able to encapsulate and decapsulate the GTP and generic routing encapsulation (GRE) packets based on fine-grained rules installed by the SDN controller. All the necessary extensions for such packet manipulation are currently being defined in ONF W&MWG.

3GPP/LTE gateways need to support a number of capabilities, including IP address allocation and management, gating and rate control, event reporting and bearer binding, and support for handovers with end-markers and buffering in the SGW. Support for these capabilities will be exposed by the OF-mpc channel extensions including metering, rating, and switch events.

OpenFlow Extension for Mobile Packet Core — In the MPC architecture OF switches need to handle GTP-U for actions like GTP-U encapsulation or decapsulation. In order to do that, the OpenFlow protocol needs to be extended. Possible extensions are:

- *Match field extension*: To distinguish evolved packet system (EPS) bearers and the corresponding GTP-U signal messages, the OF switch shall be able to match "message type" and "TEID" of a GTP-U header.
- *Action type extension*: The OF switch shall encapsulate and decapsulate GTP-U tunnels. The former action is used to encapsulate the user packet into an appropriate GTP-U tunnel, while the latter is used to decapsulate a GTP-U tunnel header from a GTP-U packet.

NFV- and SDN-Based MPC

The NFV framework in [3] consists of the NFV infrastructure (NFVI), which logically partitions the resources from underlying physical hardware resources. NFV management and orchestration (MANO) is the life cycle management of network service and the virtual network function (VNF) instance. The VNF is purely a software application (i.e., network functions like MME and PCRF) deployed in NFVI, and VNF is mapped into a VM running on top of NFVI. The MANO orchestrates other specific managers such as the virtual infrastructure manager (VIM) and the VNF Manager (VNFM). The VNFM is in charge of interacting with the VNFs, whereas the VIM is in charge of managing the NFVI, which includes computing, storage, and

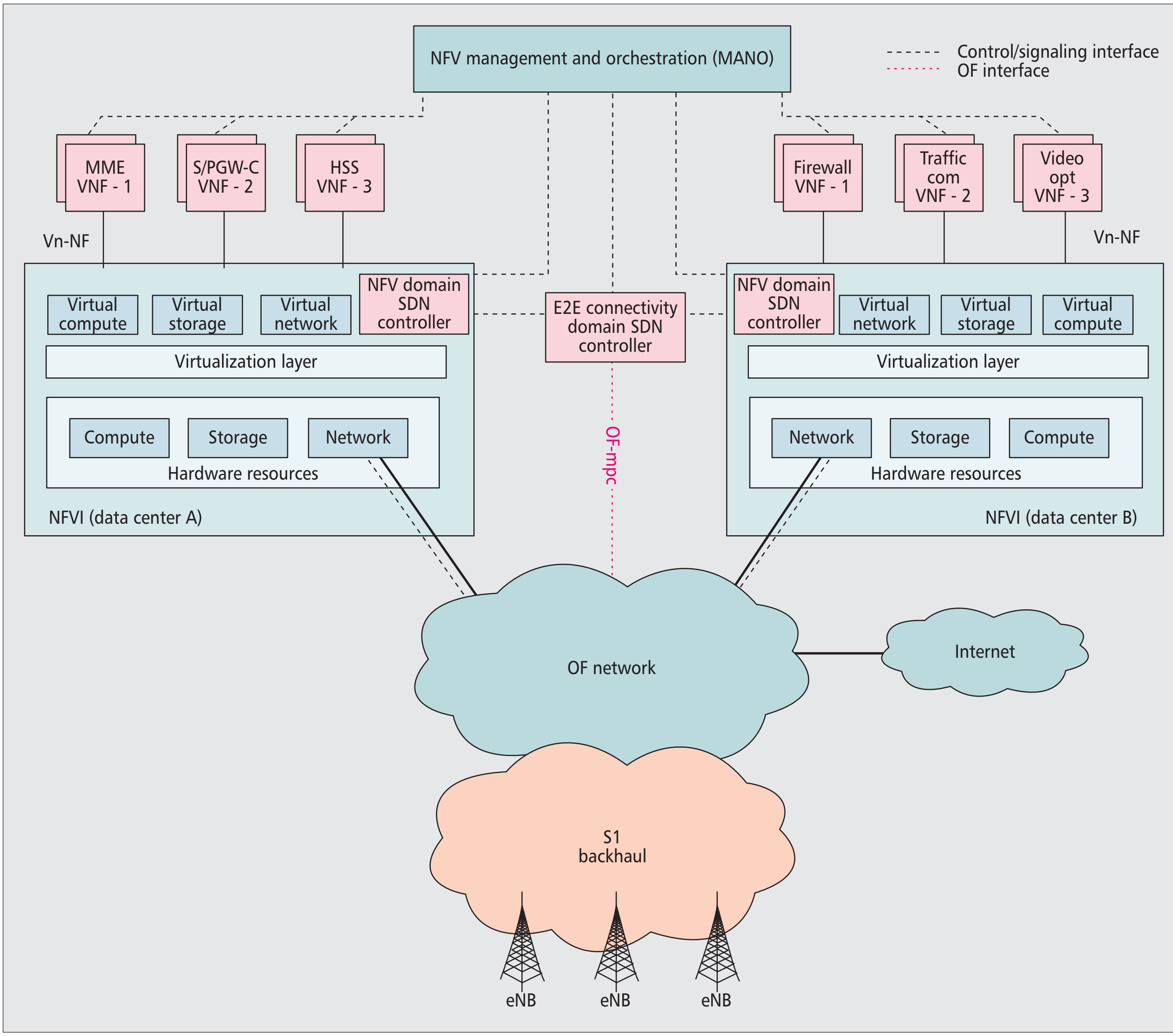


Figure 3. NFV- and SDN-based MPC architecture.

networking. The MANO has a total visibility of all VNFs running inside the NFVI. In addition, it is in-charge of the operation and configuration of VNFs, for example, through the operations support system (OSS)/base station subsystem (BSS).

One of the use cases outlined in the NFV group specifications is the mobile core [5]. A major motivation for enabling virtualization in this part of the network is the different pace of the scaling needs for each of the involved core entities, like the MME, SGW, and PGW. Those needs can be translated into higher or lower necessity for function distribution, or even higher or lower requirements in terms of resource usage (some entities limited by user plane traffic vs. others limited by control plane signaling).

NFV and SDN Combined Architecture

To enable flexibility, programmability, and optimization of network resources, we propose a logically centralized EPC control plane (i.e., MME, combined gateway handler, PCRF, etc.) running as a group of VNFs hosted in operators' data centers, as illustrated in Fig. 3.

In order to optimize resources in the user plane, the GTP tunnels are only used between the eNode B (eNB) and the first ingress switch of an OpenFlow-enabled network (i.e., S1 backhaul in Fig. 3). Note that substitutes for GTP tunnels are possible as long as they can manage UEs' localities and identities. Beyond the first ingress switch, the traffic is forwarded by an OpenFlow-enabled network infrastructure considering both the user equipment IP and destination addresses, and the QoS characteristics defined by the control entities through the NBI. In addition, if required, layer 2 (VLAN) tags, for instance, can be used to differentiate the transport bearers to differentiate QoS. Moreover, tags can also be used for overlapping UE addresses going to different access point names (APNs).

The "NFV domain SDN controller" is in charge of providing connectivity to the computing and storage resources locally to the data center. Those resources will attach as endpoints to transport capabilities in the OF-enabled network.

On the other hand, the "E2E connectivity domain SDN controller" is the main entity in our architecture as it manages the forwarding plane. It can consist of a network operating system (NetOS) running a collection of application modules, such as topology discovery, path computation, resource management, and load balancing. It also installs fine-grained packet handling rules in the switch. For instance, forwarding rules for UE flows can be dependent on the UE location, type of service, and network conditions. Note that the centralization of functions can introduce scalability issues (e.g., bottlenecks). For this scalability issue, there can be two possibilities:

- The controller is executed in a clustered environment (i.e., deploying active and backup controllers) using shared databases.
- The control functions (i.e., flow rules installation) are decentralized between the controller and switch.

In the latter case, the switch itself acts as a local controller and installs the rules. If the switch is unable to install the rules, the switch communicates with the centralized controller and gets the rules to improve scalability. In any case, the scalability issue in the SDN network is an extensive field of research, while in most cases that issue can be addressed without losing the benefits of SDN [6].

EPC Control Entities in an SDN Controlled NFV Environment

The EPC control plane entities can be instantiated into high-volume servers, while the data plane forwarding tasks can be performed by OpenFlow-enabled switches as shown in Fig. 3.

All these control entities are mapped as software applications in VNF instances running on top of NFVIs in a data center. For instance, the VNFs (e.g., MME, PGW) can run on top of kernel-based VM (KVM) hypervisors in a server. This KVM isolates the server hardware resources between the VNFs and interconnects the VNFs using virtual network ports and bridges inside the server. On the other hand, standard interfaces are maintained toward the other EPC control plane applications such as the MME, HSS, and PCRF, as shown in Fig. 2. In addition, the end-to-end (E2E) data plane is connected by means of SDN controllers by using communication mechanisms such as the REST application programming interface (API).

For instance, the combined control function of SGW and PGW (S/PGW-C in Fig. 3 or combined GW handler in Fig. 2) allocates the tunnel ids for UE flows, acting as a mobility anchor for inter-eNB communication. Furthermore, the S/PGW-C connects to the MME using GTP-C (i.e., the S11 interface, which is GTP-C-based).

The S/PGW-C function will use a northbound interface (e.g., REST API) with the "E2E connectivity domain SDN controller," which translates northbound messages into OpenFlow messages. For instance, the S/PGW-C sends the UE user plane bearer GTP TEIDs to the controller, which translates these IDs into OpenFlow messages (e.g., Packet Out) instructing the switches, in a similar manner as proposed by [14]. To do this, first the OpenFlow protocol needs an extension to transfer the GTP TEIDs; and second, the S/PGW-C function needs to send UE information (e.g., mobility updates) to the controller. The northbound interface needs to be extended to exchange the UE information (GTP TEIDs, eNB address, etc.).

The MANO is in charge of managing the hardware and virtual network resources within the data center as well as configuring the VNFs and their interfaces, defined in Fig. 3. For instance, if a new virtual MME is switched on in a data center, the MANO will notify and configure this new MME with the corresponding association of other virtual functions such as the PGW-C, HSS, and other MMEs. In addition, the MANO will interact with the referred SDN controllers for setting up the required connectivity, both internally and E2E.

On the other hand, the new MMEs are also required to configure and establish S1-AP interfaces with the eNBs, which are physical devices geographically distributed across the network. For instance, when an MME is scaled-in into the network, the MANO gives the pool of eNBs of interest and the information details to be configured. Based on this information, the new MME will start an initial handshake procedure and establish the S1-AP interface as specified in the 3GPP specification [7].

When an MME is scaled-in into the network, the MANO gives the pool of eNBs of interest and the information details to be configured. Based on this information, the new MME will start an initial handshake procedure and establish the S1-AP interface as specified in the 3GPP specification.

Realization of an Operational Control Architecture

The combination of virtual mobile control functions together with an underlying programmable infrastructure requires clear differentiation of actions to be implemented on the network. Some of those actions are intended for deploying and maintaining the network functionality, while others are specific for building pure connectivity. It seems difficult to develop programmatic applications capable of covering such a broad scope of actions in a unique controller.

Layered SDN Control Architecture

The centralized MPC control functions have to interact with different types of resources to accomplish service delivery. Obviously, control over the underlying transport resources is needed to forward the user plane traffic properly (e.g., the connectivity to the Internet for mobile data access), which relies on the bidirectional transport of user packets between the PGW and SGW. This will require the population of forwarding rules on OpenFlow-enabled switches in a coordinated way and the composition of an E2E connectivity path across the SDN domain. These are resources related to the transport function.

Notwithstanding, the packet core entities need to access databases and storage resources for a variety of functions, including security, logging, and registers, or even processing capacity (in terms of computing CPUs) for running their

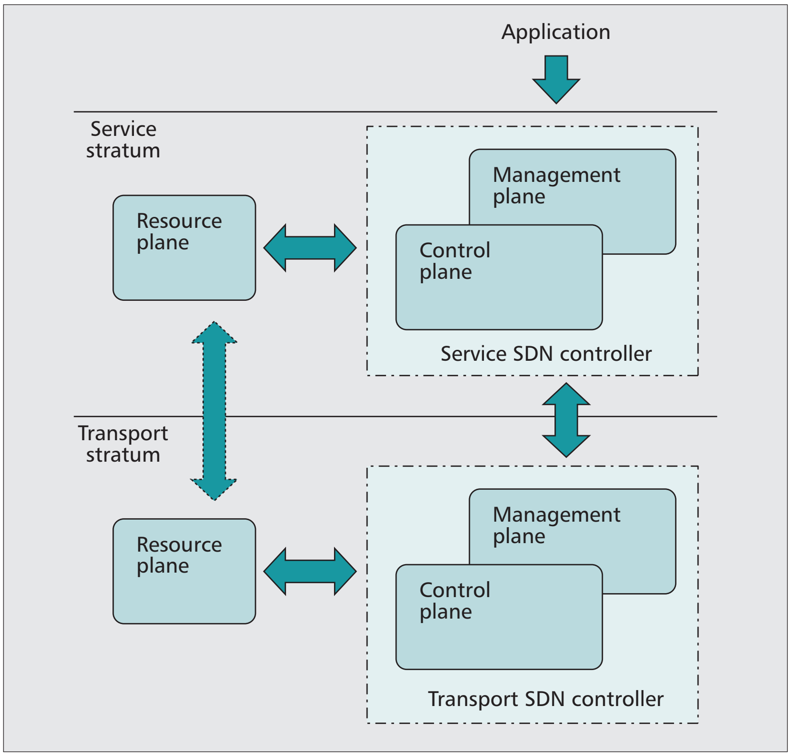


Figure 4. Layered control architecture [8].

own logic and finite state machine. These resources are related to the service itself.

There is then a clear differentiation between the kind of resources (for transport and service) involved in the service provision. However, all of those resources have to be orchestrated in a consistent way to perform the expected service. An approach is then needed that considers both kinds of resources at the same time.

Existing proposals for SDN-based control do not provide a clear separation between service and transport control.

The cooperating layered architecture for SDN (CLAS) [8] proposes a clear separation between both service and transport control, with both layers cooperating in a tight way for the final service provision. It is based on the functional separation in the NGN architecture defined by International Telecommunication Union — Telecommunication Standardization Sector (ITU-T) Recommendation Y.2011.

Basically, in CLAS both the control functions associated with transport and those related to services are differentiated by allocating them in separate control layers, which cooperate to build the final service. Figure 4 presents a schematic view of the layered architecture.

Two strata are defined, the transport stratum and the service stratum. The former includes the functions devoted to handling the data forwarding by building (and managing over time) connectivity between communication endpoints. Apart from dealing with distinct adapters for interacting with the infrastructure (OpenFlow, NETCONF, etc.), this may involve some transport abstraction capabilities.

The latter comprises the functions related to the provision of services as well as the control capabilities exposed to external applications that run on top for providing value added services.

Each stratum consists of three different planes. One of those planes refers to the resources targeted for the given stratum. For example, the service stratum could require storage capacity, while the transport stratum could require disjoint paths for ensuring network resiliency. In some cases the service resources can be connected to the transport ones (e.g., as the terminating points of a transport function), while in other cases they can be totally decoupled.

The control plane consists of the control functions of each stratum and directly implementing control actions (e.g., configuration) on the corresponding resources. Both control planes have to be tightly coordinated in order to provide consistent service.

The management plane is in charge of performing management procedures, including the collection of monitoring data, alarms, and so on for elements that are part of either the resource or control plane.

Both the control and management planes in CLAS can be assimilated to the homonym planes (plus the associated abstraction layer) in [2]. The separation of service and transport functionalities can be seen as a method of recursion. On the other end, the resource plane in CLAS can be homologated to the forwarding and operational planes in [2], including the corresponding abstraction layer thus defined.

Coordinated Service and Transport Control in the Virtualized Mobile Packet Core

The NFV architecture considers a number of components. While the VNFM is in charge of instantiating and controlling the EPC functions (and others distinct from EPC for complementing the service), the VIM controls the computing, storage, and networking resources associated with them.

Porting this idea to the CLAS layering, the VIM will interact with (or incorporate the capability of) the SDN controller in the service stratum when deploying the VNFs for configuring the computing and storage resources for the VNF of interest, but also for the networking part to attach those VNFs to the border of the underlying transport network to make them reachable from outside the data center. Those are the resources required from the service point of view when forming the resource plane of the service stratum.

On the other hand, the VIM has to prepare the paths in the pure transport network to make those VNFs reachable. For instance, looking at Fig. 1, the VIM has to prepare the transport network, for example, for ridding the S1 or S5 interfaces on top of it. Then the SDN controller in the transport stratum takes actions for preparing those paths, attaching the needed resources from the service stratum as endpoints for those paths.

Figure 5 presents a functional mapping of the NFV architecture to the CLAS concept. According to this mapping, the management of the VNFs (via the VNF manager) resides in the service stratum. This component should also config-

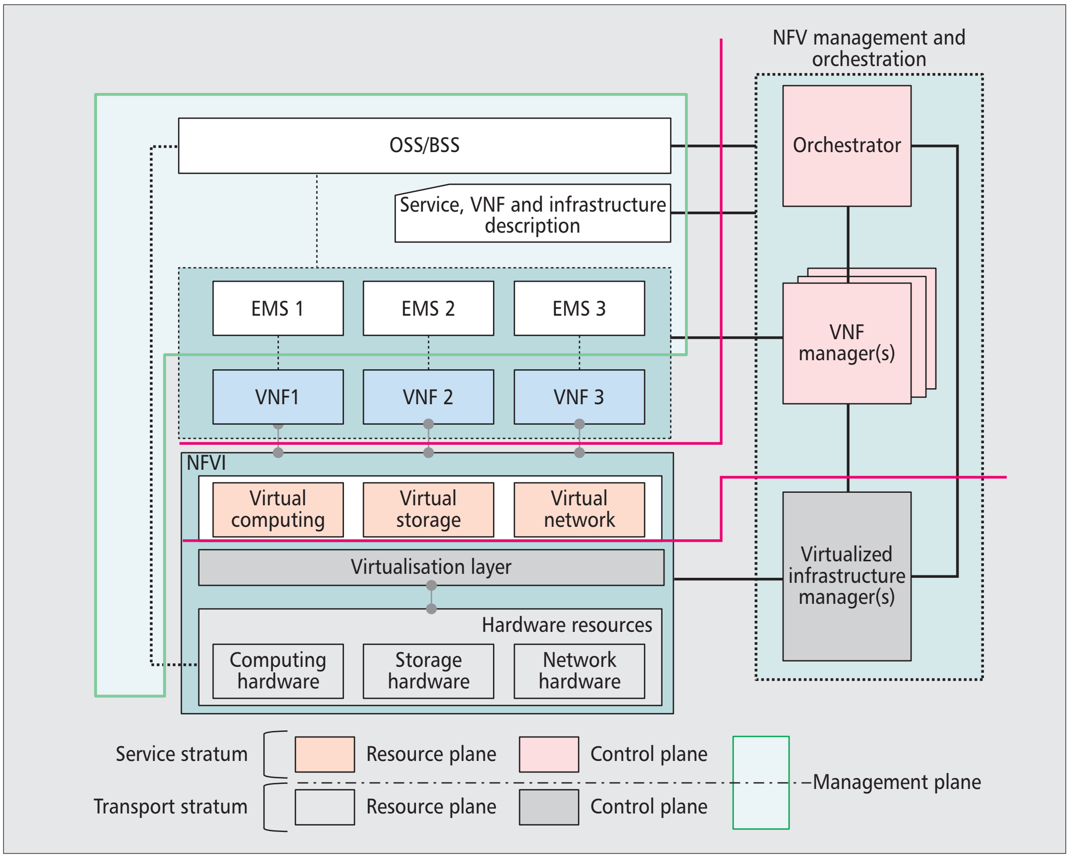


Figure 5. Layered approach to NFV control.

NFV will provide a new dimension in network flexibility by allowing a dynamic instantiation of network functions across the network. To complete the picture, coordination between the service and transport components is required to facilitate an agile composition of E2E mobile services.

ure the required resources where the VNFs are built on top, for either computing, storage, or networking. That control role in the service stratum performed by the service SDN controller in the CLAS proposal corresponds to the NFV domain SDN controller in Fig. 3.

On the other hand, the handling of the hardware resources needed to provide convenient transport of data flows between separated VNFs or VNFs and access nodes will be the responsibility of the transport SDN controller. That role is played by the E2E connectivity domain SDN controller in Fig. 3 (equivalent to the transport SDN controller in CLAS) offering the OF-mpc southbound interface.

Related Work

In recent years, the concept of SDN and NFV technologies in the mobile network has attracted significant attention in the standardization bodies and research fields. In fact, the SDN and NFV topics in the mobile networks arena have addressed several topics such as heterogeneous wireless networks [9], mobile backhauling [11], LTE/EPC [10], and prominent generic use cases in wireless networks [12]. The authors of CellSDN [13] described a generalized concept of an SDN-based cellular core network for low-cost management. This paper considered LTE technology, but lacked detailed modification requirements for LTE/EPC standards. Specifically, the LTE/EPC and their functions are not studied in this paper. This work is more closely related to a non-3GPP access network like WiFi.

On the other hand, the work in [14, 15] can be viewed as complementary to our work, emphasizing the key role that software-ization and virtualization perform in the future mobile network. In [14], the authors present a study on the evolution of cloud-based EPC, where all the control functions of the SGW, PGW, and MME are "lifted up" into the cloud. The user plane is shifted into the OpenFlow switches, and these switches are extended to support GTP. The authors of [15] propose the MobileFlow architecture for future carrier networks and provide a functional evaluation of a demonstrator. In this architecture, the user and control planes of the EPC are split into a MobileFlow forwarding engine (user plane) and MobileFlow controller (control plane). The eNB participates in this functional split, and the entire control plane (EPC and eNB) is centralized. Both aspects are different from our approach. In fact, the eNB split makes processing more complex and control messages between the eNB user and control planes more frequent. For instance, with frequent UE idle/active and handover messages, the eNB needs to communicate with the centralized eNB-C to obtain the actions, which is also another major difference compared to our approach. Indeed, any changes/split in the eNB needs to transform the existing deployed eNBs in the entire network.

Table 1 shows a comparison matrix of the solution proposed in this article with respect to the above mentioned related work. The incremental step of this work is to propose an integrated architecture considering different

Solution	Network segment	Compatibility with 3GPP	Southband interface	Integration with NFV
CROWD [9]	Mobile backhaul and radio access	No	OpenFlow (switching) and others (radio)	No
S.B.H. Said *et al.* [10]	EPC and radio access	No	OpenFlow	Possible
Costa-Requena [11]	Mobile backhaul and EPC	No	OpenFlow	Possible
Bernardos *et al.* [12]	EPC and radio access	Possible	OpenFlow, ForCES, NETCONF	Yes
SoftCell [13]	EPC and radio access	No	OpenFlow	No
Kempf *et al.* [14]	EPC and radio access	Yes (control entities)	OpenFlow	Possible
Mobileflow [15]	EPC and radio access	Yes (control entities)	MobileFlow and Open-Flow	Possible
Described in this article	EPC	Yes	OpenFlow	Yes

Table 1. Comparison to related work.

proposals in ONF, ETSI, and the Internet Engineering/Research Task Force (IETF/IRTF) for SDN control of a virtualized packet core. Compatibility with existing 3GPP networks is ensured by not modifying the interfaces dealing with entities external to the packet core. OpenFlow is the southbound protocol of choice, leveraging on the work currently done in ONF for extending such protocol for support of mobile packet core functionalities in a standard manner.

Conclusions

In this article, we have presented high-level software defined control of virtualized mobile packet core network architecture, and discussed the configuration of mobile control entities and their interfaces in the virtualization platform. The SDN- and NFV-based MPC will foster the services and rollout of new network features to the network through flexibility and programmability.

Current efforts in ONF will facilitate SDN/Open Flow architecture to abstract PDN connection segments from mobile access. However, further work is needed to enable a transport network that is able to support capabilities such as online charging or deep packet inspection. ONF is making efforts to use hybrid OpenFlow switches and specialized nodes to implement complex functionality in 3GPP. In addition, carrier grade features including scalability with multi-controllers and recovery from failure need to be addressed.

Similarly, NFV will provide a new dimension in network flexibility by allowing dynamic instantiation of network functions across the network. To complete the picture, coordination between the service and transport components is required to facilitate agile composition of E2E mobile services.

The proposed approach in this article is simple and threefold:

- Separation of control and user planes
- Separation of hardware and software
- Separation of services and transport control

In addition, in our proposal no changes are required to existing EPC interfaces specified by the 3GPP standard and interoperability toward existing approaches and deployments of EPC.

Acknowledgment

The authors would like to thank the anonymous reviewers, Serge Manning (ONF WMWG Chair), and Carlos J. Bernardos (Universidad Carlos III) for their helpful and constructive comments that greatly contributed to improving the overall quality of the article.

References

[1] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2013–2018.

[2] E. Haleplidis *et al.*, "Software-Defined Networking (SDN): Layers and Architecture Terminology," RFC 7426, Jan. 2015.

[3] Network Functions Virtualization Introductory white paper, v. 2, http://portal.etsi.org/NFV/NFV_White_Paper2.pdf.

[4] 3GPP TS 23.401, "Technical Specification Group Services and System Aspects; General Packet Radio Service (GPRS) Enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) Access," Release 12, v. 12.0.0, 2013.

[5] ETSI GS NFV 001, "Network Functions Virtualisation (NFV); Use Cases," v. 1.1.1, Oct, 2013.

[6] S. H. Yeganeh, A. Tootoonchian, and Y. Ganjali, "On the Scalability of Software-Defined Networking," *IEEE Commun. Mag.*, vol. 51, pp.136, 141, Feb. 2013.

[7] 3GPP TS 36.413, "Evolved Universal Terrestrial Radio Access Network (E-UTRAN); S1 Application Protocol (S1AP)," Release 12, v. 12.0.0, 2013.

[8] L. M. Contreras, C. J. Bernardos, and D. Lopez, "Cooperating Layered Architecture for SDN," draft-contreras-sdnrg-layered-sdn-01 (work in progress), Aug. 2014.

[9] H. Ali-Ahmad *et al.*, "CROWD: An SDN Approach for DenseNets", *2nd Euro. Wksp. Software Defined Networks*, Oct. 2013.

[10] S. Ben Hadj Said *et al.*, "New Control Plane in 3GPP LTE/EPC Architecture for On-Demand Connectivity Service", *IEEE 2nd Int'l. Conf. Cloud Networking*, Nov. 2013.

[11] J. Costa-Requena, "SDN Integration in LTE Mobile Backhaul Networks," *Int'l. Conf. Information Networking*, 10–12 Feb. 2014, pp. 264–69.

[12] C. J. Bernardos *et al.*, "An Architecture for Software Defined Wireless Networking," *IEEE Wireless Commun.*, vol. 21, no. 3, June 2014, pp. 52–61.

[13] X. Jin *et al.*, "SoftCell: Scalable and Flexible Cellular

Core Network Architecture",," *ACM CoNEXT '13*, New York, NY.
[14] J. Kempf *et al.*, "Moving the Mobile Evolved Packet Core to the Cloud," *IEEE 8th WiMob*, 8–10 Oct. 2012.
[15] K. Pentikousis, Y. Wang, and W. Hu, "Mobileflow: Toward Software-Defined Mobile Networks," *IEEE Commun. Mag.*, vol. 51, no. 7, July 2013.

Biographies

MALLA REDDY SAMA (mallareddy.sama@orange.com) is a research engineer at Orange Labs, France, where he is currently working on the research and development of mobile packet core virtualization using SDN and NFV technologies. He earned his Master's degree in telecommunication engineering from Politecnico di Torino, Italy. During his Master's work he spent a year at Telecom SudParis, France, as part of an Erasmus scholarship. His research interests lie in the field of 3GPP next generation mobile packet core and wireless networks, SDN, network virtualization, mobile cloud networking, and distributed networks. He has been on the Technical Program Committees of international conferences as well as an active contributor to ONF standardization.

LUIS M. CONTRERAS completed a six-year Telecom Engineer degree at the Universidad Politécnica of Madrid (1997), and also holds an M. Sc. in telematics from the Universidad Carlos III of Madrid (2010). In 1997 he joined Alcatel Spain, taking several positions (R&D, standardization, product development, and customer engineering) in both the wireless and fixed network fields. In 2006 he joined the Network Planning department of Orange Spain (France Télécom group) taking on responsibilities in IP backbone planning. Between 2002 and 2010 he was also an adjunct lecturer in the Telematics Department of the Universidad Carlos III, where he is currently a Ph.D. student. Since August 2011 he has been part of Telefónica I+D/Telefónica Global CTO, working on SDN, transport networks, and their interaction with cloud and distributed services. He has participated in the FP7 project GEYSERS and FI-PPP project XIFI. In July 2012 he co-organized the International Workshop on Cross-Stratum Optimization for Cloud Computing and Distributed Networked Applications, co-located with the 10th IEEE International Symposium on Parallel and Distributed Processing with Applications, Leganés, Spain. He is an active contributor to the ONF and IETF, co-authoring two RFCs.

JOHN KAIPPALLIMALIL (John.Kaippallimalil@huawei.com) is a senior standards and research manager at Huawei USA with more than 20 years of experience in IP technologies, mobile networks, and telecom product design. He is currently working on next generation architecture for 5G mobile networks, and his interests include SDN and virtualization of network functions. He is an active contributor to IETF and ONF, and currently leads the ONF MPC project. He received his Master's degree in computer science from Texas Tech University in 1993 with a thesis on real-time distributed computing and monitoring.

IPPEI AKIYOSHI (i-akiyoshi@ah.jp.nec.com) received his degree in communication engineering from Osaka University, Japan, in 2002. He has been with NEC since 2002, mainly working on the research and development of wireless and mobile networks such as 3GPP and WiMAX. He is currently focusing on SDN and network virtualization for wireless and mobile networks, and is involved with ONF standardization.

HAIYANG QIAN received his B.S. degree in communication engineering from Nanjing University of Science and Technology, China; his M.S. degree in telecommunications from the Technical University of Denmark, Lyngby; and his Ph.D. degree from the University of Missouri — Kansas City. He worked at PhoneFactor Inc. (acquired by Microsoft) and Futurewei Technologies in engineering roles. He was at China Mobile U.S. Research Center as a senior researcher. He is the co-founder of Grabtalk. He has published over 10 peer reviewed papers and filed a few patents in the NDN, SDN, mobile Internet, and cloud computing fields. He has served on the Technical Program Committees of a number of international conferences. His current research interests include cloud computing, software-defined networking, and mobile Internet systems and products.

HUI NI (hui.ni@huawei.com) received his Ph.D. degree in computer science from Tianjin University in 2007. In the same year, he joined the Wireless Research Department of Huawei as a researcher and standards delegate. From 2008 to 2011, he actively contributed to the WiMAX Forum and held leadership of the WiMAX-3GPP interworking group. Currently, he focuses on evolution of next-generation wireless networks, and his research interests include SDN, NFV, ICN, and IoT. He is an active contributor to ONF.

vConductor: An Enabler for Achieving Virtual Network Integration as a Service

Wenyu Shen, Masahiro Yoshida, Kenji Minato, and Wataru Imajuku

ABSTRACT

Network function virtualization is regarded as a promising candidate for future networks. Although many advantages are expected, there is still a lack of services that directly link to increased revenue for telecom operators. Therefore, we propose a new NFV-based service, virtual network integration as a service (VNIaaS). The service allows enterprise network administrators to construct and monitor virtual enterprise networks as necessary. As a key enabler, we further propose vConductor, an innovative technology due to its automatic network provisioning, multi-objective resource scheduling, and NFV-oriented inventory management characteristics. Finally, we show the feasibility through implementation of a prototype system. We believe that this contribution will be a catalyst to accelerate the industrialization of NFV.

INTRODUCTION

Network function virtualization (NFV) has recently attracted much attention worldwide [1]. Academia and industry have come to regard it as a promising candidate for the so-called future networks, along with software defined networking [1, 2]. Originally, by eliminating dedicated physical devices and fully utilizing the cloud infrastructure, there is the possibility of significant cost reduction. However, recently a more promising prediction has come to light in that software technologies increase network flexibility and enable rapid delivery of new functions. This yields insight into the possibility of creating a new service. After all, telecom operators are more interested in increasing revenues nowadays.

With the focus of creating a new source of income, we propose a new NFV-based service, virtual network integration as a service (VNIaaS), which is positioned to serve enterprise users. VNIaaS provides a portal site in which a network administrator constructs a virtual enterprise network on the fly by utilizing a virtualized network function (VNF) store and a graphical user interface (GUI)-based virtual network editor. In fact, VNIaaS provides one-stop integration of cloud and networking services so that all related provisioning tasks, including provisioning of network functions, application servers, and logical connections, are handled automatically. We also provide monitoring as an optional service.

However, completely software-based network construction causes unexpected complexity. In addition, a variety of business needs have to be handled for individual enterprise users when performing resource scheduling. Fault management becomes extremely difficult when the virtualized layer is mixed with the existing physical layer, and this situation becomes further complex when spanning multiple domains.

To address these problems, we propose vConductor, a key enabler for achieving VNIaaS. First, vConductor achieves full automation of virtual network provisioning by simplifying and normalizing the provisioning procedure. Second, a multi-objective resource scheduling algorithm is adopted so that individual business needs can be precisely met with limited investment. To facilitate fault management, especially fault isolation, we further enhance the existing inventory management. This is achieved in the design of an NFV-oriented data model based on the TM Forum Information Framework (SID) [3]. Third, vConductor is designed to be generic. It is positioned as a common solution to NFV management and orchestration (MANO) according to the end-to-end architecture defined in the NFV group specification. We maximize the system modularity so that vConductor can easily be applied to other use cases such as the virtualized evolved packet core. To achieve this, we employ a plug-in architecture to enhance the interoperability with various VNF managers and virtualized infrastructure managers (VIMs). This article summarizes the business goals, technical challenges, and essential design concepts on the whole toward realizing VNIaaS, while we leave the discussion of individual implementation details to our previous works [4, 5].

VNIAAS: A NEW REVENUE SOURCE FOR TELECOM OPERATORS

Figure 1 shows a service image of VNIaaS. We aim to provide a portal site that allows a network administrator to construct a virtual enterprise network as necessary.

Wenyu Shen is with NTT Communications.

Masahiro Yoshida, Kenji Minato, and Wataru Imajuku are with NTT Network Innovation Laboratories.

Wenyu Shen was with NTT Network Innovation Laboratories when the paper was submitted.

 0163-6804/15/$25.00 © 2015 IEEE

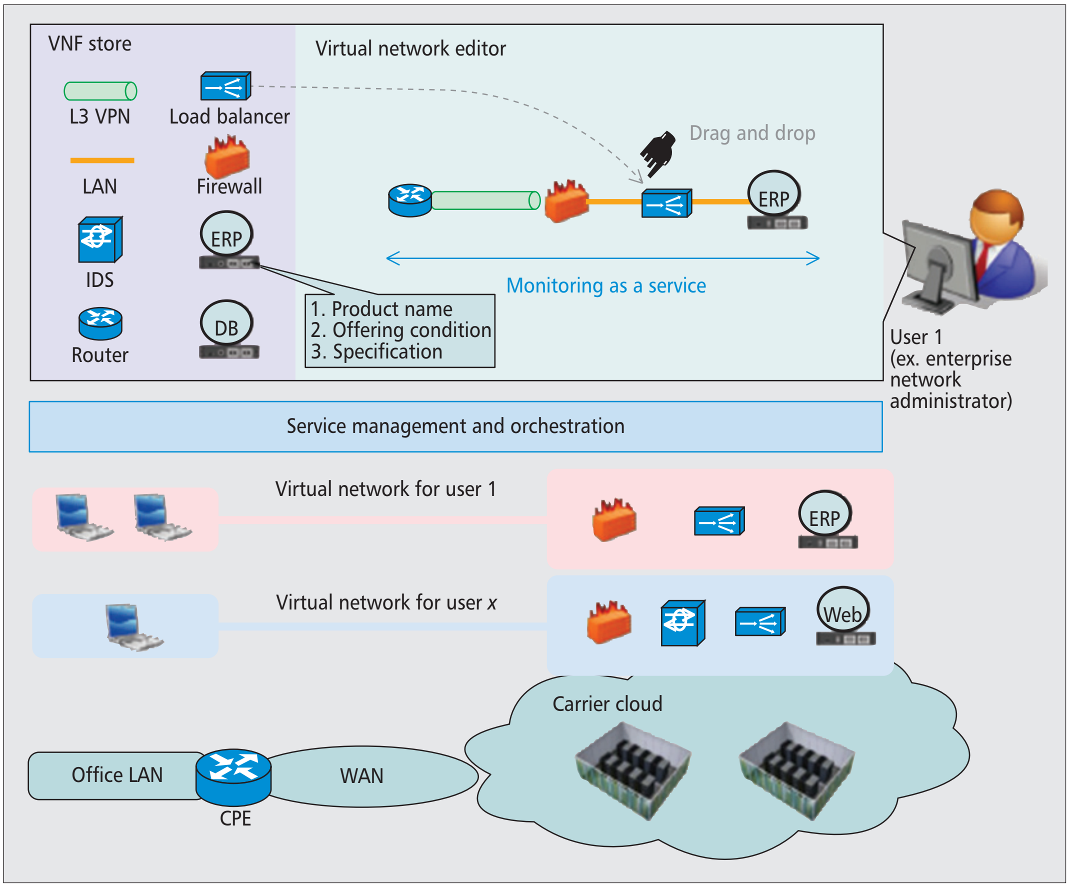


Figure 1. VNIaaS concepts.

The concepts of VNF store, virtual network editor, end-to-end connection management, and monitoring as a service differentiate VNIaaS from many of the existing cloud services.

VNIaaS Concepts

We discuss four main VNIaaS concepts below. These concepts differentiate it from many of the existing cloud services that provide users with only isolated virtualized resources instead of end-to-end solutions.

VNF store: Similar to many popular application stores nowadays, VNIaaS includes a VNF store that provides a variety of choices when users construct a virtual network. Figure 1 shows some VNFs such as an intrusion detection system (IDS), load balancer, and even application templates, such as an enterprise resource planning (ERP) system and a database (DB). To assist further in the procurement of VNFs and network design, detailed catalog information is available, which covers product names, offering conditions, and specifications.

Virtual network editor: The virtual network editor provides a GUI for users to design a virtual network comprising VNFs purchased at the VNF store. The only operations required of users are to drag and drop icons representing the VNFs, and to connect them using virtualized links (VLs). Our MANO solution enables full automation of all actual provisioning tasks.

End-to-end connection management: When an enterprise network is located in a telecom cloud, connections between the customer premises equipment and remote data centers (DCs) must be considered. This falls into the traditional area of expertise of a telecom operator. Here, the combination with a virtual private network (VPN) service makes VNIaaS a true end-to-end solution.

Monitoring as a service: We believe that VNIaaS has the potential to accelerate the transformation of communications, in which enterprise networks will be gradually moved to the telecom cloud. VNIaaS enables telecom operators to participate in the operation of enterprise networks. As an option, VNIaaS further provides a monitoring service from the viewpoint of enterprise users.

Related Work

The concept of a network as a service (NaaS), regarded as a starting point of network virtualization, has been studied for several years. We see that a variety of VPN services have become a main revenue source for many telecom operators. An overlay network is yet another form of NaaS, which is typically implemented in the application layer, although various implementations at lower layers of the network stack do exist [6]. A more challenging concept has been proposed in academia, advocating dynamic instantiation of layers [7]. NFV diversifies network virtualization; in the NFV context, VNFs can be provided on demand. Therefore, the new concept of function as a service (FaaS) appeared [8].

The proposed VNIaaS provides VNFs on demand in addition to the tunnels that are the only focus in the traditional NaaS. Besides, VNIaaS integrates functions of both networks and cloud applications to create an end-to-end ser-

> *To realize VNIaaS, we propose vConductor. The initial result of our proof of concept has shown the feasibility of achieving full automation of network provisioning, multi-objective resource scheduling, and end-to-end inventory management for virtualized environments.*

vice that spans multiple domains, while this falls outside the scope of FaaS, which considers only one function or functions within a single domain. We clarify that VNIaaS is more than a simple combination of the two concepts; as discussed below, VNIaaS has great potential to change the business model in the current ecosystem of communications industry, while new technical challenges occur in the way of actualizing the concept.

Benefits for Enterprise Users and Telecom Operators

We believe that VNIaaS will benefit enterprise users and telecom operators. For enterprises, it is no longer necessary for network administrators to purchase expensive hardware, because software is generally less expensive and even free in many cases. Actually, our VNF store offers a platform where network administrators can select the most suitable software from the perspective of price and functionality. When constructing a network, users benefit from a one-stop service, integrating wide area networks (WANs) and local area networks (LANs). After network construction, the functions, topology and even architecture can be modified easily according to dynamic requirements. In VNIaaS, we achieve these using a graphical editor. Finally, the optional monitoring service potentially releases the network administrator from routine tasks and contributes to reducing operating expenses.

As a telecom operator, we anticipate a revenue increase. Initially, we expect to create new sources of income by selling VNFs. There is also potential for product bundling (WAN and cloud service) by providing end-to-end connections. Hence, we foresee an opportunity for a telecom operator to transform from an infrastructure provider to a solution provider.

Technical Challenges for Realizing VNIaaS

In VNIaaS, hardware devices are substituted with VNFs that are implemented as software appliances running on virtual machines (VMs) while the cables are substituted with VLs that are implemented as logical tunnels. As a result, the network construction process becomes execution operations for a variety of scripts, which unexpectedly result in a complex and error-prone task. Therefore, one challenge is to achieve automatic network provisioning.

Due to the essence of virtualization, extra considerations must be taken on issues such as the amount of resources that must be reserved to guarantee VNF performance, and selecting a physical container, e.g., a DC, for deploying a VNF. In the enterprise business domain, it is critical to offer multiple service options at competitive prices to meet diverse needs. Moreover, those needs often conflict with each other. Therefore, a resource schedule oriented to multiple objectives is desirable.

Next, a network can now incorporate virtualized entities, which raises the issue that managed objects that can be dynamically modified [6]. Furthermore, the services tend to be end-to-end in most cases and are assumed to span multiple domains, e.g., DCs. These changes actually complicate the inventory management. In order to convince potential customers of the reliability of VNIaaS, it is necessary to provide a complete end-to-end network view that includes all the information regarding the physical layer, virtualized layer, and mapping between the layers. A shortfall in any of the information can lead to failed fault management resulting in a breach of the service level agreement (SLA). Therefore, achieving end-to-end inventory management for a virtualized environment becomes a must.

vConductor: Enabler Technology

To realize VNIaaS, we propose vConductor. The initial result of our proof of concept, described later, has shown the feasibility of achieving full automation of network provisioning, multi-objective resource scheduling, and end-to-end inventory management for virtualized environments. These are enabled through our work on the simplification and normalization of provisioning procedures based on a standard system structure, the design of a novel resource scheduling algorithm, and NFV-oriented data modeling.

Standard System Structure and Provisioning Procedures

The system structure, shown in Fig. 2, fully conforms to the NFV end-to-end architecture, where it has an orchestrator as the core including the virtual network life cycle manager (VNLM), resource designer, service level controller, and databases, and also embeds the VNF managers. In addition, the user portal and element managers correspond to the operation support system (OSS) and element management system, respectively. vConductor relies on existing VIMs such as OpenStack (www.openstack.org) for managing DCs and OpenDayLight (www.opendaylight.org) for managing WANs. They are assumed to be external systems.

VNLM: The VNLM controls operations of the internal components and external systems based on the received service orders (SOs) regarding network generation, modification, and deletion; they are implemented as a series of instance descriptors [9]. A SO is first input into an analyzer, where its format and feasibility are validated, after which the analyzer calculates the necessary computing and network resources by referring to the information stored in the catalog DB. It is the resource reservation component that actually performs reservations.

Resource Designer: The resource designer schedules resources and generates parameters. The former calculates how VNFs are allocated to the underlying infrastructure (e.g., DCs), while the latter designs the necessary logical resources such as IP addresses for practical resource reservation. The resource designer interacts with the resource DB to retrieve the necessary information regarding physical and logical resources. A detailed description of resource scheduling is presented in the next part.

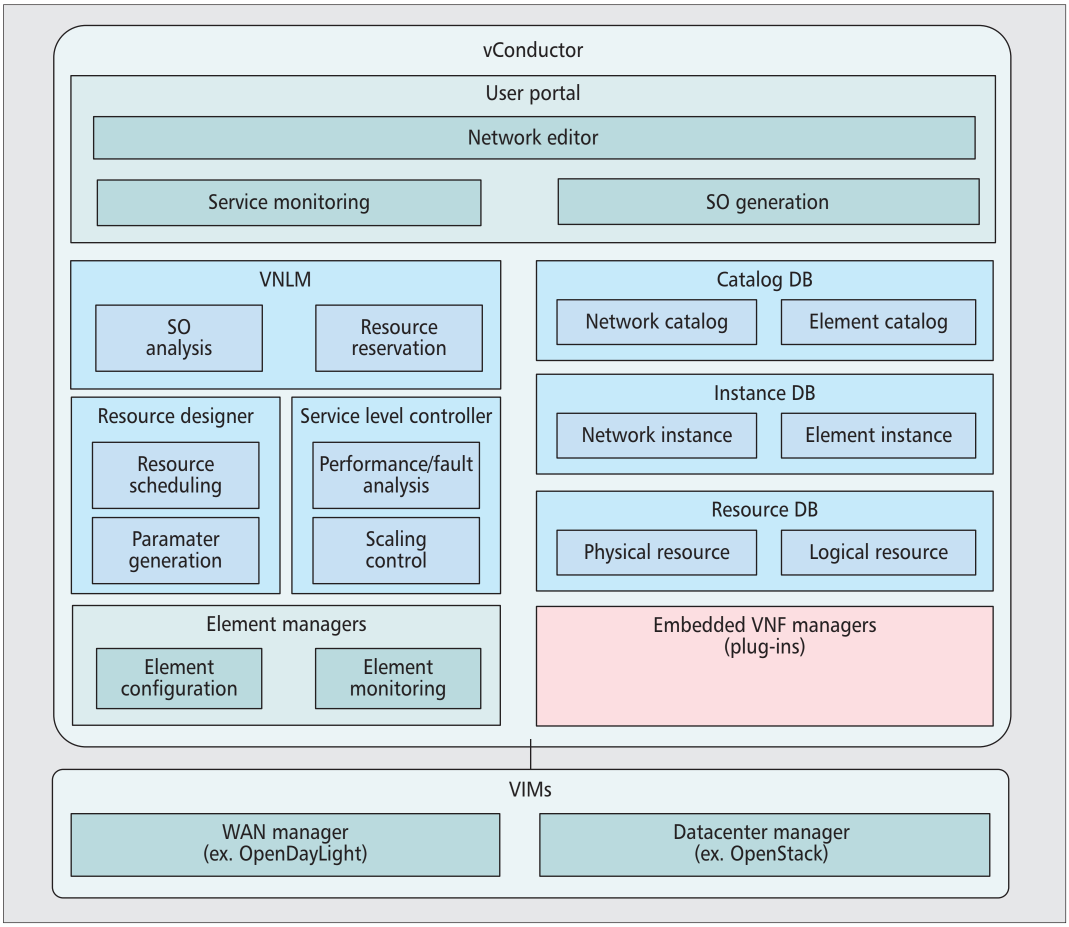


Figure 2. System structure of vConductor.

The user portal provides an editor for end users to customize their virtual networks. We designed a GUI to minimize the user's effort so that the user is only required to drag and drop some icons that represent VNFs and connect them together with VLs.

Service level controller: This component is designed to guarantee the SLA by handling hardware faults and performance degradation. To achieve this, the performance/fault analysis component analyzes the status of the VNFs and their supporting resources retrieved from the VNF managers and VIMs, based on which the root cause is located. For the recovery process, we duplicate the affected VM and reallocate it. Here, the scaling control component determines the most proper performance timing.

Catalog, instance, and resource DBs: These DBs perform information management for the whole system. The catalog DB stores network and element catalogs in the form of descriptors [9]. The element catalog stores resource-related information for implementing, for example, VNFs and VLs, while frequently used connection patterns are summarized as templates, known as network catalogs. Next, the instance DB manages the configuration information of the elements and networks that are set according to the catalogs. Let us consider VNFs as an example. Parameters such as assigned IP addresses, performance indicators, and log data are covered. To fulfill the inventory management role, the instance DB also manages a large number of correlations among the network layers and multiple domains. This is discussed in the data modeling section. Finally, the resource DB stores information regarding the physical resources (e.g., the number of CPU cores) and logical resources (e.g., IP addresses), and interacts with the VIMs to perform updates on a regular basis.

VNF manager: The VNF manager controls the life cycles of its VNFs, covering (un)installation of their hosting VMs, which are defined as virtual deployment units (VDUs) in European Telecommunications Standards Institute (ETSI) terminology, and their initialization and termination [9]. It also calls the element manager for the initial configuration. In order to perform the actions above, a manifest file is needed for each VNF package to record the operational sequence.

User portal: The user portal provides an editor for end users to customize their virtual networks. We designed a GUI to minimize the user's effort so that the user is only required to drag and drop some icons that represent VNFs and connect them together with VLs. The user input is translated into a SO before it can communicate with the VNLM. Furthermore, a user-level monitoring service can be provided, which helps the user understand the real-time condition of the services. The element managers are in charge of the real configuration and monitoring of individual elements. Currently, we are using some vendor-specific tools that are designed for the corresponding VNFs. However, we believe that more work should be performed such as interface specification and information abstraction in order to facilitate management.

We assume that a variety of business needs, also known as objective functions, coexist. Each of these factors influences the VNF allocation decision. Moreover, these factors tend to conflict with each other in many cases. A simple example is an enterprise user with a limited investment expects a minimum end-to-end latency and maximum throughput.

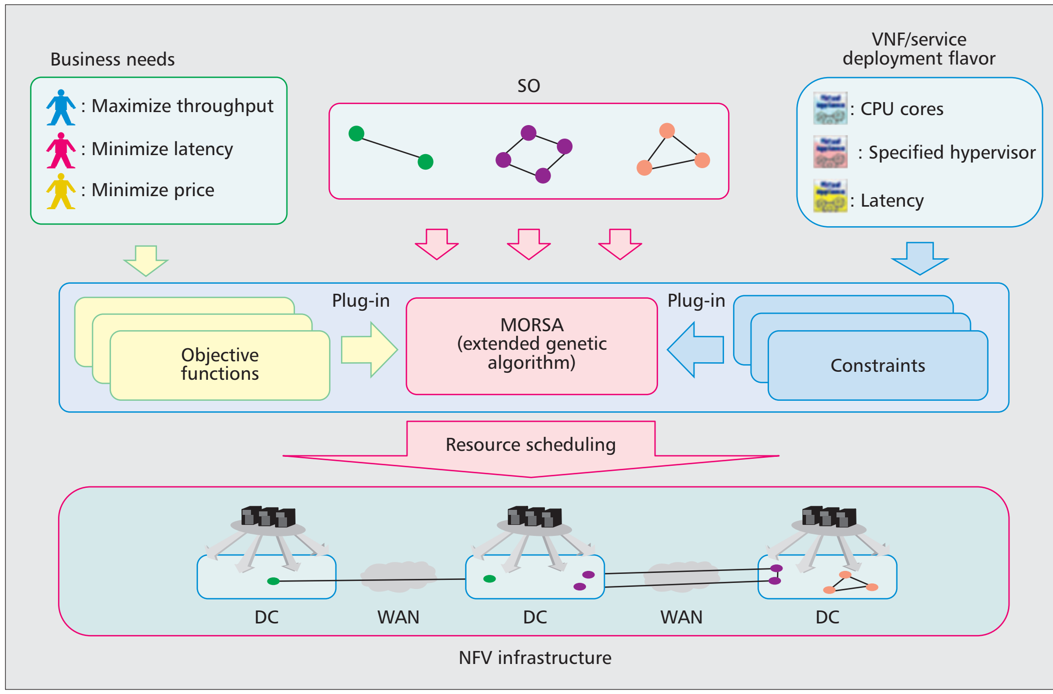


Figure 3. MORSA concept.

We further present a refined virtual network provisioning procedure, which facilitates the understanding of the function of individual components and enables automation. We also introduce a VIM driver (not illustrated in Fig. 2) in order to absorb the diversity of VIMs so that the provisioning procedure can be simplified and normalized, focusing only on the general purpose. In the following, we describe the normalized process for network provisioning that spans multiple DCs.

Step 1: The VNLM receives SOs from the user portal.

Step 2: The VNLM retrieves the necessary catalog-related information from the catalog DB.

Step 3: Based on the SOs and catalog-related information, the VNLM further calculates the resources required for running the requested service and sends the information to the resource designer.

Step 4: The resource designer polls the resource DB for the available physical resources and decides the deployment destination for the VNFs.

Step 5: The resource designer polls the resource DB for the available logical resources and designs configuration parameters for implementing VNFs and VLs.

Step 6: Based on the design results, the VNLM reserves resources via the VIM driver.

Step 7: The VIM driver establishes underlying inter-DC tunnels, interacting with the WAN manager.

Step 8: The VIM driver reserves computer resources and establishes the underlying inner-DC tunnels, interacting with the DC manager.

Step 9: The VNF manager installs and initializes the component VDUs and performs the initial VNF configuration.

Step 10: The instances of the deployed network are registered in the instance DB.

Multi-Objective Resource Scheduling

The allocation of VNFs to the underlying infrastructure, which for VNIaaS is a cloud environment spanning multiple DCs and WAN connections, is regarded as complex. First, when deciding the NFV allocation, we must consider many constraints on the resources that remain in the infrastructure in order to prevent overprovisioning, while great care should also be taken to guarantee the agreed performance of individual VNFs and end-to-end services. This refers to the constraints on the deployment flavor. Furthermore, we assume that a variety of business needs, also known as objective functions, coexist. Each of these factors influences the VNF allocation decision. Moreover, these factors tend to conflict with each other in many cases. A simple example is an enterprise user with a limited investment expecting a minimum end-to-end latency and maximum throughput. Clearly, the optimal solution that meets the needs for all the requirements is not practical. Hence, it is sometimes a must to search all the possible solutions in order to make the most reasonable decision.

There are several existing algorithms for solving the virtual network embedding problem [10, 11], but we think that none of them can meet the specific needs in the context of VNIaaS. Therefore, we propose our multi-objective resource scheduling algorithm (MORSA) for vConductor, which simultaneously considers multiple objective functions. Figure 3 shows the concept; details are given in our previous work [5]. MORSA originates from the traditional genetic algorithm (GA) [12], but we extend it to the specific needs of vConductor. In the MORSA concept, we first convert the constraints into the form of objective functions. A charac-

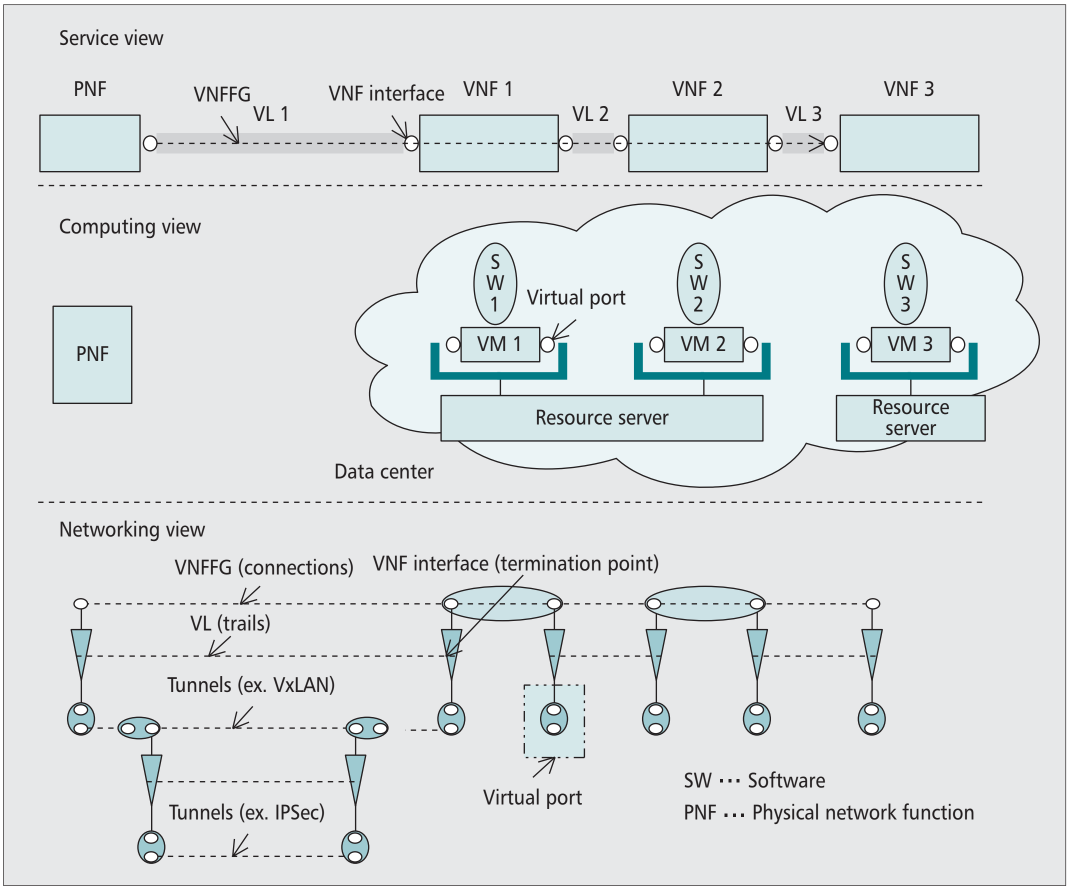


Figure 4. Function modeling for an example virtual network.

To manage the correlations behind the virtualized environment, we further propose a powerful data model. The proposed model completely conforms to the SID so that it supports newly proposed NFV concepts.

teristic of the GA is that the computation time will not drastically increase with the addition of objective functions. By utilizing this characteristic, resource scheduling in vConductor will not take a long time even if there are many constraints. Furthermore, MORSA is totally plug-in-based. Objective functions can be plugged in and out according to dynamic demands. Another advantage is that MORSA further extends the GA to perform optimization while adjusting the weights of each objective function. In this way, more solutions can be obtained instead of limited localized solutions in the case of the original GA.

The final service selection is left to the user based on a service vs. pay scale. That is to say, the user can select the final solution by arbitrarily and freely adjusting the weights among the objective functions to arrive at a desirable price point.

Data Modeling

We discuss the challenge in fulfilling inventory management for virtual networks. We advocate that the essence is to handle a variety of correlations. Thus, our discussion starts with function modeling of virtual networks. Figure 4 models an example virtual network from three different views: service, computing, and networking. The service view takes into consideration the modeling of service function chaining (SFC), as discussed in the Internet Engineering Task Force (IETF) (e.g., the attachment of a VNF interface to a VNF and its connection to a VL) and routing (e.g., a VNF forwarding graph, VNFFG) [9], but our focus is only on the management plane. From the computing view, we need to further handle, for example, how a VNF is implemented on a virtual machine (VM) and how a VM and its virtual port are allocated on a resource server. The networking view is constructed mainly based on the modeling techniques in ITU-T Recommendation G.805 [13]. By utilizing the concept of connections and trails, we envision a layering correlation of logical tunnels for implementing the VLs and VNFFGs.

To manage these correlations, we further propose a powerful data model, the core of information management in vConductor. The proposed model completely conforms to the SID model [3], so it is regarded to be general and have high affinity to many existing OSSs. We further extend SID so that it supports newly proposed NFV concepts such as VNFs and VLs. Therefore, most of the NFV-specific correlations illustrated in Fig. 4 can be handled by our proposal, the details of which are shown in Fig. 5.

First, we discuss service-related modeling. The top left of Fig. 5 shows that a virtual network service (NetworkService) is created as a subclass of ResourceFacingService. The associations such as NSRequiresVNFFGs, NSRequiresVLs, NSRequiresVNFs, and NSRequiresPNFs represent the correlations between a service and its components. We further combine a VL with a VNF by introducing

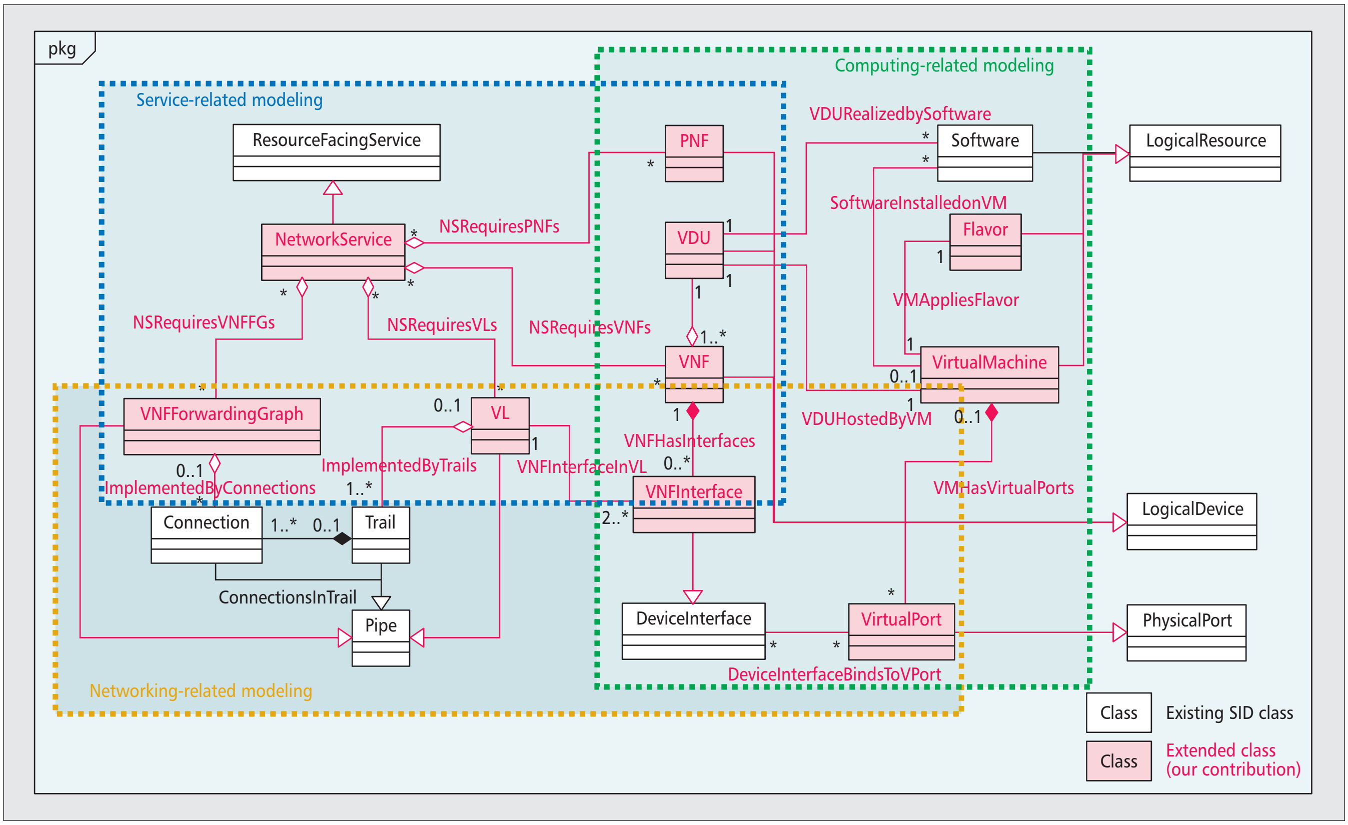


Figure 5. Data modeling in vConductor.

the VNFInterface class and two associations (VNFHasInterfaces and VNFInterfaceInVL). Inherited from DeviceInterface, VNFInterface represents a virtualized interface of a VNF.

Computing-related modeling is presented on the right of the figure, where VNF, PNF, and VDU are designed as subclasses of LogicalDevice. The fact that a VNF comprises several VDUs is represented by the association of HasVDUs. Moreover, we design a class called VirtualMachine to represent VMs and the Flavor class to represent its specification so that the concept of server virtualization can be modeled by a series of associations of SoftwareInstalledOnVM, VMAppliesFlavor, VDURealizedBySoftware, and VDUHostedByVM. These associations enable the management of correlations among a VDU, its implementing software, and the container VM and its flavor.

Finally, we discuss the networking-related modeling. The VL and VNFForwardingGraph classes are introduced and designed as subclasses of Pipe (bottom left of Fig. 5). The ImplementedByTrails association represents the fact that a VL is implemented by one or more trails. The existence of multiple trails actually means a kind of 1:*N* protection mechanism. Next, the ImplementedByConnections association represents the concept of a VNFFG, which is implemented by connecting a series of link connections and subnetwork connections. We use the VirtualPort class to represent virtual network interface cards (vNICs) while using the VPortBindsToDeviceInterface association and VMHasVirtualPorts to link a VNF interface with the vNIC and then to a VM.

Proof of Concept

We expect the emergence of new NFV-based services enabled by the vConductor proposal. vConductor enables the provisioning for constructing a virtual network to be performed automatically. In addition, user-level monitoring becomes available, so it is possible for a telecom operator to provide one-stop management services.

We implemented a prototype to show the feasibility and are planning a demonstration aimed at partially proving the concept in a future event. Figure 6 shows the detailed implementation of the prototype structure, covering the data plane and control and management plane. One physical customer premises equipment (CPE) in premises #C and two scattered DCs (DC #A, DC #B) are assumed. As DC managers, we reuse OpenStack, which comprises one cloud controller, one network node, and multiple computer nodes. Then we use simulated WAN connections to link the CPE to the DCs. Since we assume that the simulated WAN configuration is fixed, we use OpenFlow switches at the edges to control the switching of flows for each service. These switches are controlled by an overlay tunnel manager. For simplicity, we use Ansible (www.ansible.com) to control the VNF life cycle and implement a tool for configuring the VNFs selected for our example. We connect vConductor to the Ansible system, the VNF manager, the overlay tunnel manager, and the OpenStack system to yield a running system. System integration is not an easy job due to the existence of different interfaces; as a workaround, we create some wrappers to absorb the diversity. However,

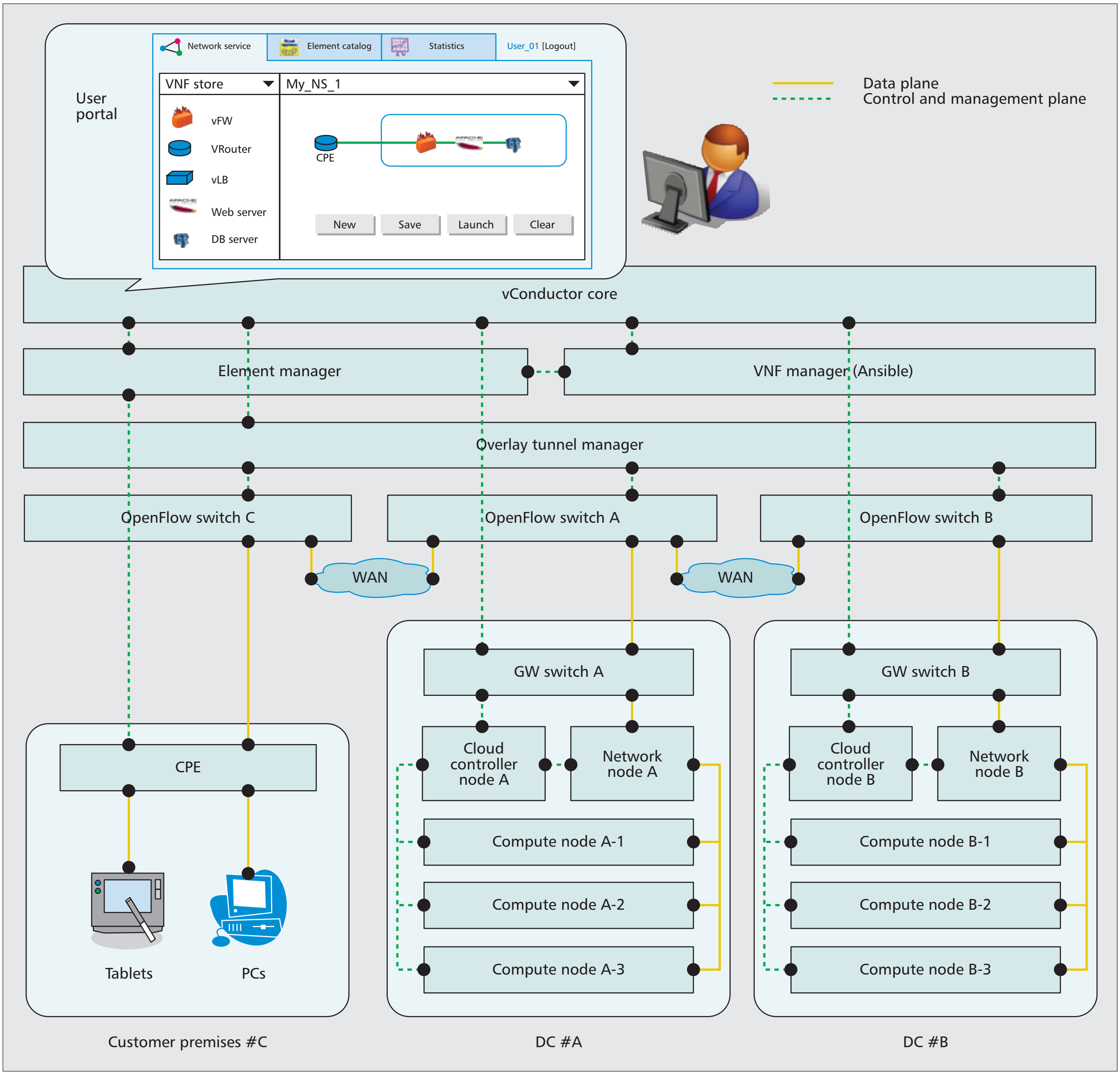


Figure 6. Implementation of prototype structure.

more effort is planned to be put into the interface unification and standardization.

Figure 6 also illustrates a GUI image where all virtualized elements are presented as icons. More details, such as the catalog information, can be viewed by double clicking an icon. The right part presents an editor that allows the user to customize his/her network. The shown service chain is only an example that covers some network and server applications. The topology can be created and changed by simply dragging and dropping the icons. Here, we hide the existence of multiple DCs. Finally, we should note that the prototype further provides extensibility as future extension to an operator interface, where inventory information of a specific service is displayed graphically, although not illustrated. Hence, when any fault occurs, the scope of its influence can be known immediately.

Conclusion and Future Work

This article proposes a new NFV-based service, VNIaaS, by leveraging vConductor. vConductor is a powerful MANO system to support VNIaaS by providing not only automatic network provisioning but also multi-objective resource scheduling and NFV-oriented inventory management capabilities. As a result, VNIaaS allows a network administrator to construct a virtual enterprise network as necessary, while we successfully achieve inter-layer virtual resource mapping to support low-cost operation to actualize rapid fault isolation. VNIaaS is regarded as an attempt to integrate cloud and networking services.

As future work, we plan to investigate the adaptation of vConductor to our commercial services such as the metro Ethernet and cloud

Although the prototype system validates the feasibility, there is still the question of how to evaluate our proposal. We are currently working closely with other telecom operators to define the necessary carrier-oriented service quality indicators in terms of scalability and performance.

services. In addition, we will try to standardize the format of resource catalogs so that new VNFs can be rapidly integrated into the system. Although the prototype system validates the feasibility, there is still the question of how to evaluate our proposal. We are currently working closely with other telecom operators to define the necessary carrier-oriented service quality indicators in terms of scalability and performance. Finally, although CPE is assumed as a physical device in the prototype, we plan to virtualize and move it to our edge cloud. We believe that this contribution will be a catalyst for accelerating the industrialization of NFV.

References

[1] M. Chiosi *et al.*, "Network Functions Virtualisation: An Introduction, Benefits, Enablers, Challenges and Call for Action," *SDN and OpenFlow World Congress*, Darmstadt, Germany, Oct. 2012.

[2] N. McKeown *et al.*, "OpenFlow: Enabling Innovation in Campus Networks," *ACM SIGCOMM Computer Commun. Review*, vol. 38, no. 2, Apr. 2008, pp. 69–74.

[3] J. Strassner, *Policy-Based Network Management: Solutions for the Next Generation*, Morgan Kaufmann, 2004.

[4] W. Shen *et al.*, "vConductor: An NFV Management Solution for Realizing End-to-End Virtual Network Services," *Proc. APNOMS '14 (TS2-2)*, Hsinchu, Taiwan, Sept. 2014.

[5] M. Yoshida *et al.*, "MORSA: A Multi-Objective Resource Scheduling Algorithm for NFV Infrastructure," *Proc. APNOMS '14 (TS4-2)*, Hsinchu, Taiwan, Sept. 2014.

[6] N. M. Mosharaf, K. Chowdhury, and R. Boutaba, "Network Virtualization: State of the Art and Research Challenges," *IEEE Commun. Mag.*, vol. 47, no. 7, July 2009, pp. 20–26.

[7] F. Esposito *et al.*, "Dynamic Layer Instantiation as a Service," demo at NSDIv '13, Lombard, IL, Apr. 2013.

[8] G. Monteleone and P. Paglierani, "Session Border Controller Virtualization Towards Service-Defined Networks Based on NFV and SDN," *Proc. IEEE SDN4FNS*, Trento, Italy, Nov. 2013.

[9] J. Quittek *et al.*, "Network Function Virtualization (NFV) Management and Orchestration (v0.8)," GS NFV-MAN 001, ESTI NFV, Nov. 2014.

[10] Y. Zhu and M. Ammar, "Algorithms for Assigning Substrate Network Resources to Virtual Network Components," *Proc. IEEE INFOCOM*, Barcelona, Spain, Apr. 2006.

[11] M. Yu *et al.*, "Rethinking Virtual Network Embedding: Substrate Support for Path Splitting and Migration," *ACM SIGGOMM Computer Commun. Rev.*, vol. 38, no. 2, Apr. 2008, pp. 17–29.

[12] M. Mitchel, *An Introduction to Genetic Algorithms*, MIT Press, 1996.

[13] ITU-T Rec. G.805, "Generic Functional Architecture of Transport Networks," Mar. 2000.

Biographies

WENYU SHEN (wenyu.shen@ntt.com) received his M.S. degree in computer science from Tongji University in 2008. In the same year, he joined NTT Network Innovation Laboratories, Yokosuka, Japan, where he worked as a research scientist. He joined NTT Communications in 2014, where he is now in charge of some SDN/NFV-based development projects for commercial services.

MASAHIRO YOSHIDA received his Ph.D. degree from the University of Tokyo in 2013. He received a research fellowship for young scientists from the Japan Society for the Promotion of Science (JSPS) during 2011–2013. He is currently a researcher at NTT Network Innovation Laboratories. His research interests include SDN, NFV, and data center networking.

KENJI MINATO [M] received his B.E. and M.E. in electrical engineering from Ehime University in 1987 and 1989, respectively. He joined NTT Laboratories in 1989. From 2005 to 2012, he conducted research and developed a Home ICT system at NTT Cyber Solution Laboratories. In 2012, he joined NTT Network Innovation Laboratories. He is a member of IEICE.

WATARU IMAJUKU, Ph.D, is a senior research engineer, supervisor in NTT Network Innovation Labs. He contributed to the innovation of NTT's optical network as well as the development of a number of new technologies which include over 65 patents, 85 international publications, and international standardizations such as IETF RFC 6163.

An Instrumentation and Analytics Framework for Optimal and Robust NFV Deployment

Paul Veitch, Michael J. McGrath, and Victor Bayon

Paul Veitch is with BT.

Michael J. McGrath and Victor Bayon are with Intel Labs Europe.

ABSTRACT

This article details a novel approach to the fine-tuning of carrier-grade virtualized network function deployments for both performance optimization and diagnostic purposes, using embedded instrumentation with an analytics framework. The work presented in this article is the output of a co-lab established between BT and Intel Labs Europe to investigate network-function-virtualization-related problems faced by traditional telecom operators. Results from comprehensive testing of virtual CDN and WAN acceleration use cases are presented. These results led to the development of a number of insights that will help network operators optimize the deployment of NFV on standard high volume servers. Significantly, it was found that the default configuration for some use cases resulted in suboptimal resource allocation and consumption. Consequently, opportunities exist for carriers to fine-tune their NFV deployments from both the technical and economic perspectives using approaches such as embedded instrumentation.

INTRODUCTION

The topic of network functions virtualization (NFV) is receiving significant attention within the telecommunications industry as a means to deliver innovative and scalable network capabilities at lower cost and with greater flexibility. Additionally, the scale of global investment in data center technology makes it increasingly attractive to deploy network functions as software that runs on standard high volume (SHV) servers [1, 2].

Network operators planning a gradual migration from legacy standalone hardware appliances to an architecture based on virtual network functions (VNFs) face a variety of technical and operational challenges. A major area of consideration is testing and diagnostics, whereby VNFs must be validated — prior to operational deployment — in terms of functionality, performance, robustness, and manageability. Traditional black box testing using externalized test points can result in a very restricted test methodology producing inadequate or, in the worst case scenario, skewed test results. It is important to recognize that a variety of factors should be considered in any testing regime. This includes both the resource consumption of the VNFs and their footprints on the host server/hypervisor.

The BT-Intel co-lab was established to combine expertise from both organizations to address NFV in a holistic fashion. This form of collaborative approach is critical to the successful rollout of VNFs into carrier networks due to the unique blend of IT and network skills required for deploying, provisioning, optimizing, and managing them. This article details research work from two use cases that focused on the fine-tuning of NFV deployments for both performance optimization and diagnostic purposes, using an embedded instrumentation and analytics framework in a carrier-grade environment. Results from comprehensive end-to-end testing are presented together with key insights that will help network operators optimize the deployment of VNFs on SHV servers.

PROBLEM STATEMENT AND RELATED WORK

The NFV architectural framework dictates that network functions reside as software-based virtual appliances running on an x86 server with a *hypervisor* providing access to the server's compute, storage, and network resources. The architecture is designed in such a manner that each VNF should be unaware of other guest VNFs running on the same server, due to the isolation and partitioning of resources conducted by the hypervisor within the virtualization layer. A high-level representation of the key components making up the overall NFV framework has been defined by the European Telecommunications Standards Institute (ETSI) Industry Specification Group [3].

The traditional approach to testing and validation of telecommunications equipment normally involves point testing of devices known as systems under test (or SUT). Additionally, devices such as network emulators can introduce

 0163-6804/15/$25.00

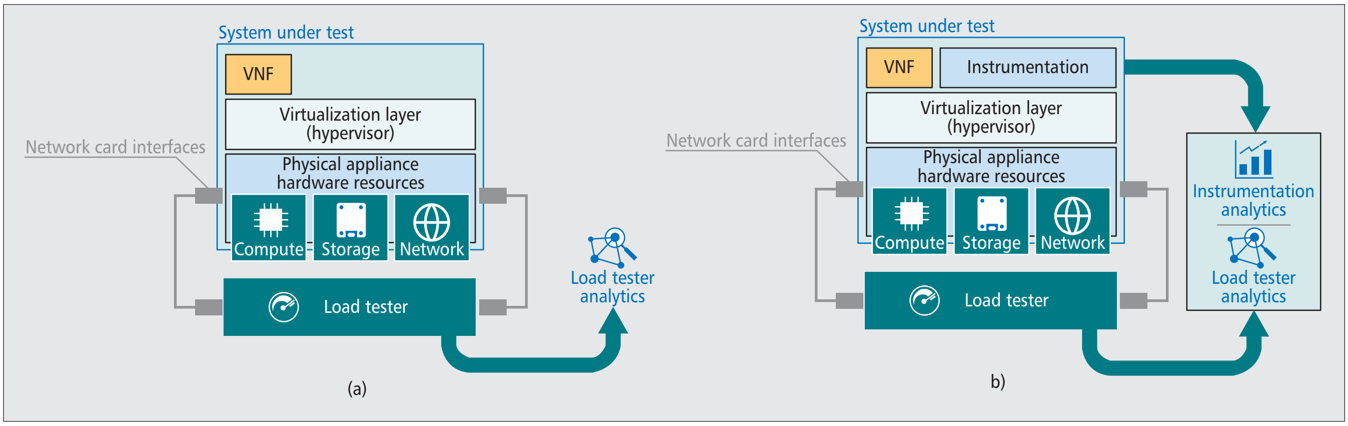


Figure 1. NFV Testing Framework: a) no instrumentation; b) embedded instrumentation.

"real-world" impairments such as latency into the test configuration. The SUTs include network functions such as routers, wide area network (WAN) accelerators, optimizers, firewalls, and load balancers. For characterization of key performance indicators such as network throughput, packets per second (PPS), and latency, a suitable load testing device must be connected to the external interfaces of the network device.

A major challenge faced by network operators planning to evolve their network estate based on the principles of NFV is the accurate characterization of VNFs in a test environment such that an optimal and robust deployment is possible in a fully operational network domain. It is therefore vital to understand any performance bottlenecks and limitations that could exist for devices running in a virtualized configuration. Since an extra layer of virtualization is introduced between the network function and the underlying hardware resources, it is critical to characterize the behavior and performance of devices in such an environment.

As the concepts of test and diagnostics within the context of NFV are relatively new, much of the published literature relevant to network operators has been channeled through standards bodies such as ETSI. A key contribution containing insights into the value and necessity of NFV test and diagnostics is the "Performance and Portability Best Practices" report [4].

The instrumentation and integrated analytics framework described in this article offers network operators an approach based on the use of instrumentation to enable performance management of VNFs in a consolidated manner. The key contribution of this article is to demonstrate the value of embedded instrumentation and integrated analytics techniques to help network operators optimize and "harden" their NFV configurations prior to an operational deployment based on a system-wide view of the VNF and its interactions with both virtual and physical resources.

Benefits and Trade-offs of Embedded Instrumentation

As shown in Fig. 1a, a traditional test methodology can be used with VNFs by connecting load generation test equipment to the network interfaces of the host x86 server. The performance characterization will be limited by the fact that the test points are 100 percent "externalized" and hence do not generate measurements from within the virtualized environment; the result is a relatively simple — but restricted — test configuration. Figure 1b illustrates an alternative approach based on embedded instrumentation that can be used within the virtualized environment to generate a more fine-grained view of performance-related test metrics. The instrumentation toolset collects and presents data for real-time processing, analysis, and visualization affording greater informational depth and detail.

Although embedded instrumentation adds additional complexity to the overall test setup, this can be offset against the following benefits for network operators:

- A richer set of metrics can be captured and analyzed due to in situ measurement within virtual machines (VMs)/VNFs resident on the NFV infrastructure (NFVI).
- Instrumentation can be deployed in a flexible fashion prior to actual NFV deployment as part of a "pre-optimization" exercise, leading to enhanced fine-tuning and customization of implementations.
- Data obtained from embedded instrumentation in a test environment can be used to determine the most useful metrics to be monitored under standard operational conditions.
- An enhanced diagnostics capability is afforded above and beyond the use of traditional externalized test points.

It should be stressed that embedded instrumentation can act as a complementary test and diagnostics resource to traditional methods, rather than as a wholesale replacement of such techniques. This fact, along with a more detailed explanation of the instrumentation and its application to real-world use cases are described in the following sections.

Instrumentation and Analysis Overview

Deploying VNFs in a performant manner on virtualized infrastructures creates significant challenges to ensure that the workload is allocated

A particular challenge for the instrumentation of virtualized network workloads is the ability to collect high rate event data without impacting the workload performance. The overhead associated with the data collection process is commonly referred to as the observer effect.

the necessary compute resources. In addition, it is necessary to ensure that the workload consumes these resources in an efficient and effective manner in order to meet the key performance indicators (KPIs) required by a network operator. Standard tools used by many operators, including hardware-based traffic load generators, can test the performance level of a VNF; however, they do not provide the necessary insights into how the VNF workload is interacting at a resource and process level within its host VM. Instrumentation of the VM environment to provide data on key system-level metrics helps to address the information gap between the *external* perspective of the test equipment and the *internal* view from within the VNF and its host VM.

Although many tools (mostly user space) already exist to assist in the gathering and monitoring of kernel counters within Linux-based systems, they provide little customization and, more importantly, offer only fragmented data sets, making a full stack view of the system and integration with analysis systems difficult [5]. This places a significant overhead on the user to process and assemble the data in a manner that can easily be exported, summarized, analyzed, interpreted, and cross-correlated with the testing equipment. The potential number of metrics that can be generated by the instrumentation easily reaches hundreds of metrics per second and millions of data points per hour.

One key feature of instrumentation is support for customization and configuration "per VNF" and the corresponding test environment. This is significant given the diversity of potential VNF workloads (multimedia, security, etc.). Typically, the configuration depends on the characteristics of the specific VNF workload (e.g., network bound, local storage bound, CPU bound). The total number of metrics available depends on:

- The number of layers instrumented and the "per layer" configuration (network layer, storage layer, performance model layer, etc.)
- The per layer granularity of the instrumentation such as aggregated per resource (e.g., overall aggregated CPU utilization)
- The sampling rate

The instrumentation framework used in the co-lab comprises two main sets of components.

Unified multi-layer/multi-host/multi service instrumentation: Focused on collecting data from the available and instrumented layers, including across hosts and service compositions. This data can be accessed and manipulated under a single unified namespace. A web-based interface uses an application program interface (API) to access and query the experimental data in the common namespace and execute analysis jobs.

Exploration, analysis, and visualization of instrumentation data. This capability allows a large number of metrics to be viewed and interacted with simultaneously. Users can employ statistical analysis capabilities to filter and identify relationships between variables across the stack in single and multiple experiment configurations.

Unified Single-Namespace Instrumentation Components and VNF Performance Modeling

A particular challenge for the instrumentation of virtualized network workloads is the ability to collect high-rate event data without impacting the workload performance. The overhead associated with the data collection process is commonly referred to as the *observer effect* [5]. The instrumentation used in the use cases presented in this article was implemented as a low overhead agent that can be embedded inside a VNF host VM. Each agent implements a series of modules, each of which typically instruments one or more layers or subsystems, such as the I/O network and CPU subsystems. The main function of the various modules developed was to read the local kernel (e.g., */proc* filesystem), service, application counters, and so on, and to transform the data collected into usable formats by parsing and normalization. The agent then collects, synchronizes, and integrates the data from the different subsystems into a single unified namespace and exports the data to other components/services such as analysis and visualization.

Each VNF requires a different set of instrumentation modules depending on the workload type. Therefore, the architecture of the instrumentation is highly extensible supporting integration of new metrics as well as the implementation of customized derived metrics that can form part of the local metric namespace. Derived metrics are based on the fusion of raw metrics using defined mathematical transformations into a single metric that is typically indicative of operational performance in a manner that cannot be measured directly. It is envisaged that as the use of embedded instrumentation in VNF deployments becomes more widespread, new categories of derived metrics will emerge on a per VNF type, offering unique opportunities for innovation.

Exploration, Analysis, and Visualization Components

The exploration, analysis, and visualization components of the instrumentation framework play an important role in helping to filter and identify the metrics of interest from the hundreds of potential metrics across different VNFs. Figure 2 illustrates the main components of the user interface (UI). Different VNF experimental runs can be selected by the user and compared across multiple test runs using different methods such as clustering, principal component analysis, and information gain, as well as summary statistics across all metrics and experiments. This form of approach allows the user to filter and find the most relevant metrics across the experiments encompassing different criteria such as which metrics changed the most, correlated differently, achieved higher values of utilization, and/or saturated across configurations.

Once the experimental runs and comparison

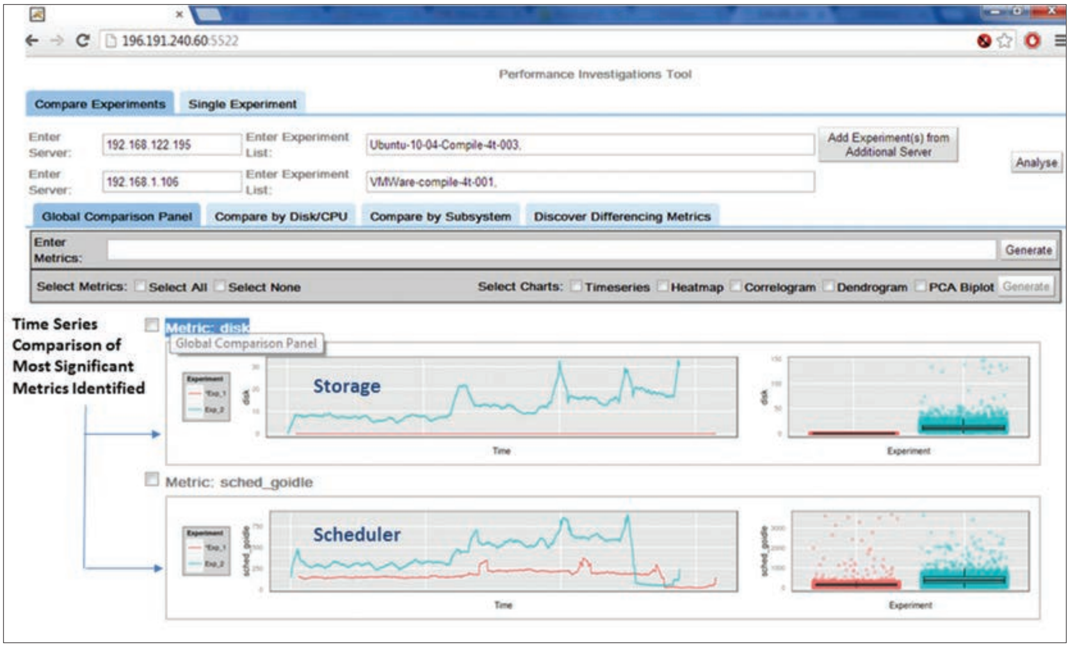


Figure 2. Web-based user interface of the instrumentation framework.

operators are selected, the framework computes the user choice and returns the result of the analysis. In the case of Fig. 2 the analysis provides a series of ranked graphs that shows a "per metric" time series and a box plot comparison of the same metric across different configurations. For this particular case, the comparisons consist of two separate experiments. The box plot allows the user to quickly evaluate the magnitude of the variation and the data distribution, while the time series graphs allow the user to see how the metric evolves on a temporal basis.

The following section provides further examples of how the tool was used in real-world test cases, including performance optimization and anomaly detection/fault finding scenarios.

Case Studies

The two case studies presented in this section are representative of real-world NFV deployments that communication service providers (CSPs) are actively progressing. Figure 3a shows the first of these, involving a content delivery network (CDN) cache as the SUT, while Fig. 3b shows the second scenario involving a WAN accelerator as the SUT. Both the CDN and WAN acceleration devices are proprietary off-the-shelf virtual instances of the network functions, and were installed and configured on Intel XEON®-based servers running a commercial type 1 hypervisor.

In the first scenario involving a virtualized CDN cache, the load tester was set up to generate multiple on-demand client requests for video content, served directly from the cache itself and returned to those clients. Although other components existed in the overall setup (including an origin server to distribute "source content" to the CDN cache, cache management software, switches, etc.), only the key components are illustrated for simplicity. The physical connectivity between the load tester and the server hosting the CDN application was 10 Gigabit Ethernet.

In the second scenario, based on a virtualized WAN acceleration device, the load tester was set up to generate multiple on-demand client requests for file content, served from designated servers on the load tester at the far end of the network link. An impairment device, which introduced a configurable and uniform network delay, was used for "WAN emulation" purposes representative of a multiprotocol label switched (MPLS) core network, and two bookended acceleration devices were deployed on either side of the impairment device. The device at the "client end" was a hardware-based appliance, while the device at the "server end" was a virtualized appliance, and was the actual SUT; together, these components form the end-to-end network service. WAN acceleration devices work in tandem to reduce protocol chattiness as well as to detect patterns of repetition in the data packets being transferred across the WAN [6]. Smaller-sized tokens are sent in place of the repetitive data packets, significantly reducing the number of packets that are sent over the WAN or viewed another way, accelerating the end-to-end data transfer. Both diagrams in Fig. 3 indicate the presence of embedded instrumentation within the VNF, as outlined in the previous section.

Results and Insights

The initial focus for both use cases was an exploratory phase to assess the stability of the instrumentation framework and to establish that it had no measureable impact on the performance of the VNF. Comparison of instrumented and uninstrumented versions of the VNF case studies confirmed no notable impact on performance, as indicated by test data presented in Table 1, collected as part of an initial baselining evaluation exercise.

The initial focus for both use cases was an exploratory phase to assess the stability of the instrumentation framework and to establish that it had no measureable impact on the performance of the VNF. Comparison of instrumented and uninstrumented versions of the VNF case studies confirmed no notable impact on performance.

	Throughput without instrumentation	Throughput with instrumentation
CDN 10G load-progressive download	10 Gb/s	10Gb/s
CDN 10G load-adaptive bit rate	10 Gb/s	9.85 Gb/s
WAN accelerator-500M load	500 Mb/s	500 Mb/s
WAN accelerator-1Gb/s load	1 Gb/s	1 Gb/s

Table 1. Baseline throughput data measurements using instrumented and uninstrumented test configurations.

Case Study 1: Virtual CDN

The virtual CDN testing included video-on-demand (VoD) content served using HTTP-based progressive download, as well as adaptive bit rate (ABR) VoD and live streaming, with a specific focus on network throughput as the KPI of principal interest.

Based on vendor recommendations, the number of virtual CPU (vCPU) cores allocated for the test of the CDN cache was initially set at 12, and this comfortably achieved the target throughput of 10 Gb/s. This was easily verifiable using the built-in load testing analysis tools. Additionally, the instrumentation indicated that the actual vCPU usage was relatively low, as indicated by significant periods of CPU idle time in both scatter plots and heat maps. This led to the consideration that similar performance could be achieved with lower numbers of vCPUs in an effort to minimize CPU idle time; hence, the tests were re-run with 4- and 6-vCPU configurations for the virtual CDN application.

Figure 4a shows the load tester view of achieved network throughput for 4, 6, and 12 vCPUs, which in all cases peaks at 10 Gb/s. Figure 4b shows the difference in the scatter plots and heat maps applicable to the 4- and 12-vCPU scenarios. It was clear from the box-and-whiskers-scatter-plots and heat maps that there is much higher idle time (where the vCPU is not doing anything useful, white/lighter areas visible on the heat maps) in the 12-vCPU case compared to 4 vCPUs. Viewed from another perspective, this is the same as saying that the 4-vCPU setup makes more efficient use of the available vCPU resources as the vCPU workloads are much higher.

This insight would not have been possible without access to the CPU idle time data gleaned from the embedded instrumentation. This leads to the compelling prospect that fewer CPUs may achieve the target performance, which could result in an economic benefit for the network provider. For example, the same server could host multiple CDN applications in a multi-tenant setup, or other virtualized applications could make use of the unallocated CPU cores.

However, a further stage of analysis may lead to the determination of the optimal setup. The graphs presented in Fig. 5 show the bytes (worked on per) millisecond ratio (BMSR) performance metric plotted for the same tests. This is a derived metric based on the sum of total network reads and writes (network throughput in bytes) vs. total CPU utilization in milliseconds. A larger value for this metric (i.e., increased throughput and less CPU required to "push" bytes around) indicates better performance which is indicative of efficient network utilization. The data shows that the 6-vCPU case had the highest

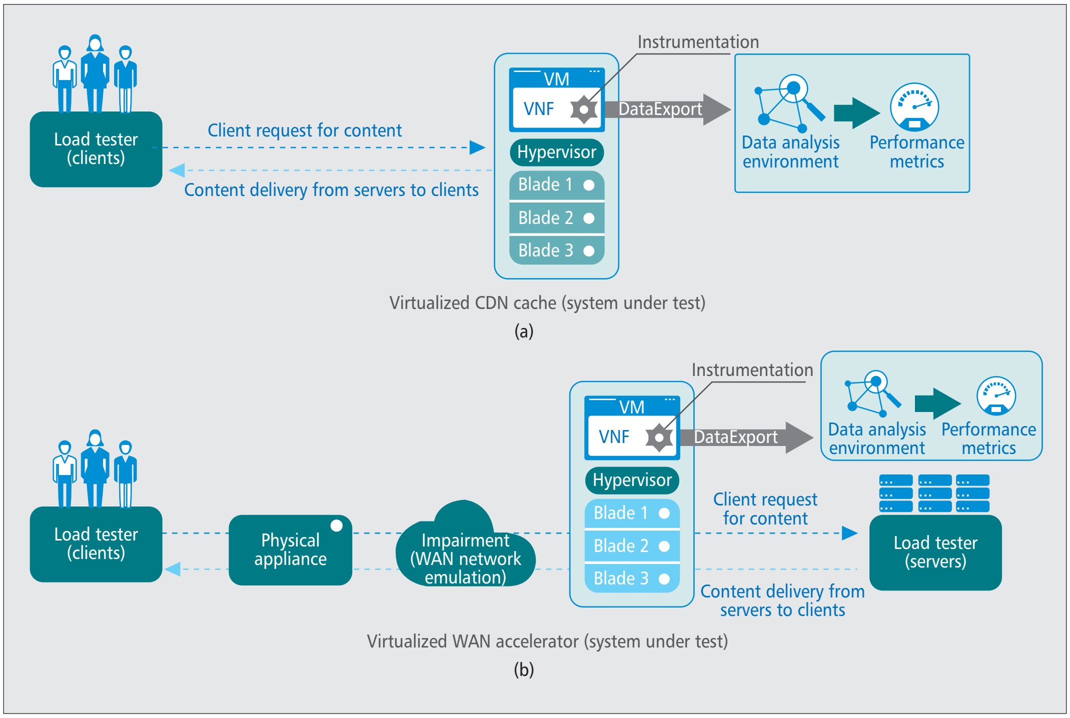


Figure 3. NFV scenarios: a) virtual CDN; b) virtual WAN acceleration.

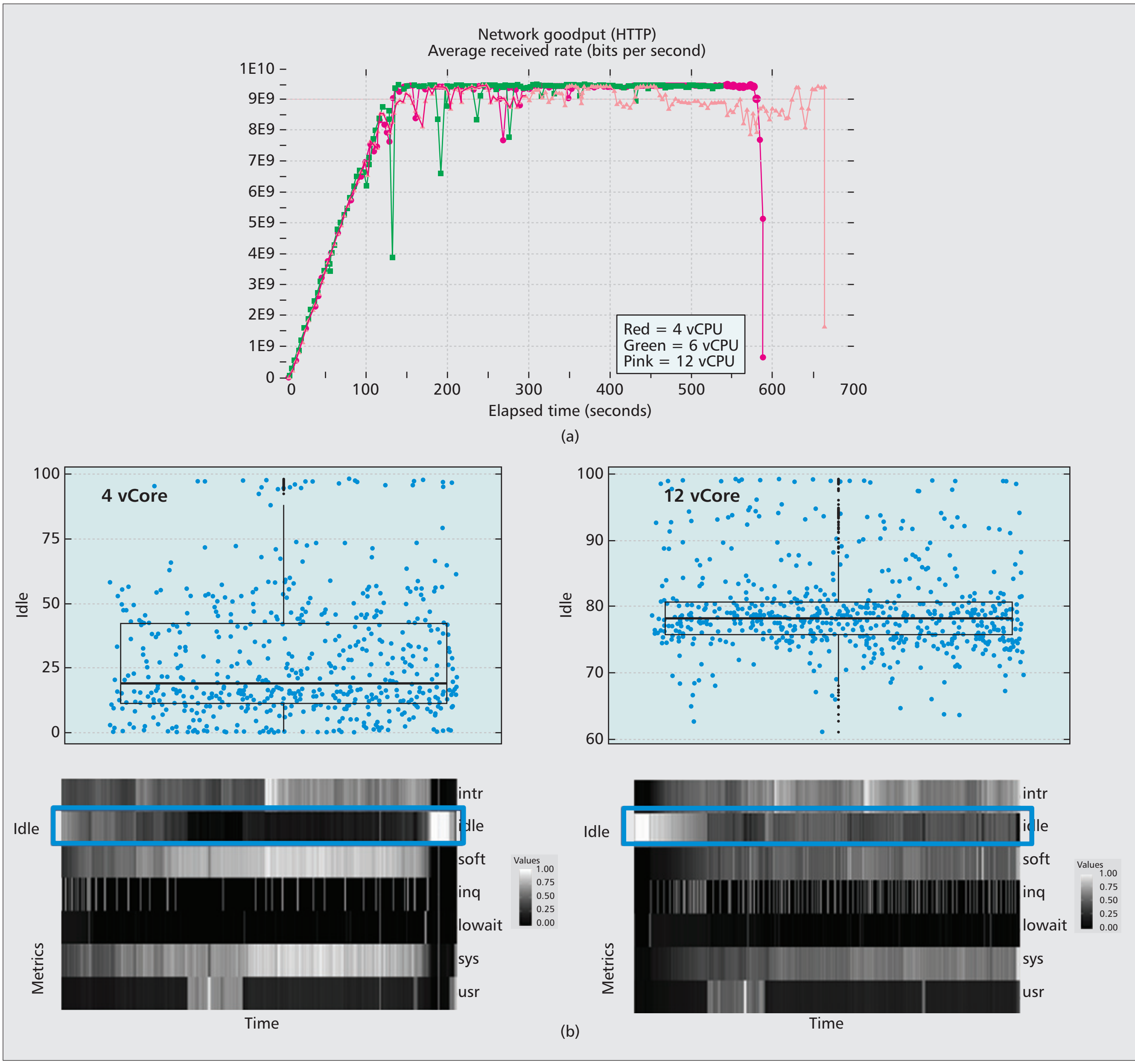


Figure 4. a) External view — 4, 6, and 12 vCPUs; b) internal view — box-and-whisker with scatter plots and corresponding heatmaps for visualization of CPU idle, 4 and 12 vCPUs.

value for this particular metric, suggesting that the "optimal" setup for this virtual CDN cache is 6 vCPU cores, which is still 50 percent lower than the original vCPU allocation and hence makes a significant saving in allocated resources. It also results in a setup that lends itself to higher vCPU utilization levels than 12 vCPUs while retaining headroom capacity. A 4-vCPU configuration — while effective from a utilization perspective — provides minimal headroom capacity and therefore may not be suitable from an operational perspective. The analytics functionality of the instrumentation framework supports the creation of derived or synthetic metrics such as BMSR, which can be implemented specifically for a VNF by the network operator.

The results for this particular scenario indicate an ability to "fine-tune" server resources that are better matched to "expected" traffic metrics such as bandwidth throughput. Any subsequent changes in traffic patterns that dictate scaling up or scaling down of resources to satisfy the demand should be accommodated more easily by leveraging the built-in elasticity of an NFV architecture compared to fixed hardware.

Case Study 2: Virtual WAN Acceleration

The virtual WAN accelerator was originally tested in the absence of any embedded instrumentation to assess performance metrics under a range of test scenarios and traffic types. This included a baseline characterization of HTTP and HTTPS (i.e., encrypted with secure socket layer, SSL) traffic flows from end user "clients" accessing remote "servers" as previously illustrated in Fig. 3b.

As the deployment of VNFs into carrier networks gathers momentum, the use of embedded instrumentation to complement existing external testing approaches and internal VNF metric capabilities is likely to grow in importance and utility. There are, however, many interesting problems that have yet to be explored in any significant depth.

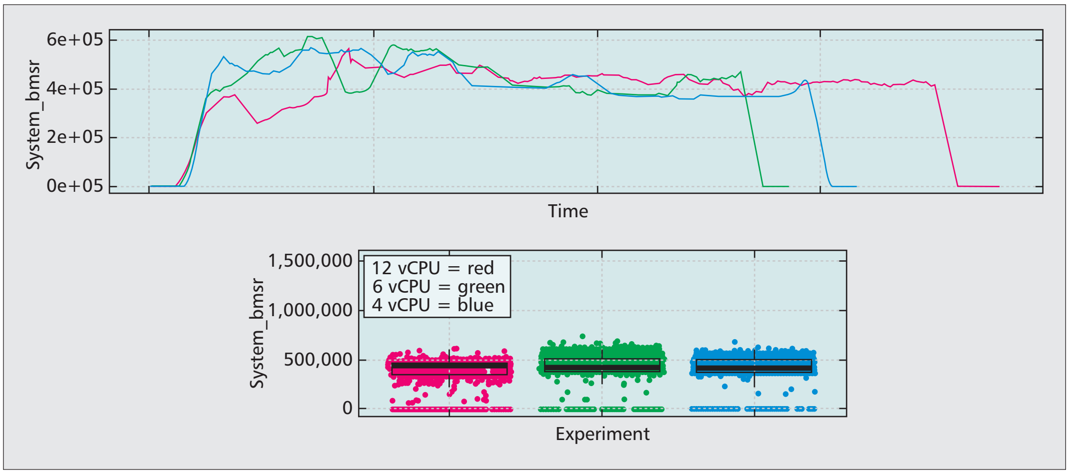


Figure 5. BMSR metric plotted for 4, 6, and 12 vCPUs.

When the instrumentation was installed, many of the tests were revisited with the initial objective of achieving performance-related "fine-tuning" of allocated server resources, similar in objective to the virtual CDN case study outlined earlier. Although not part of the original test plan, additional test validation insights were obtained from a fault diagnostic perspective. A scenario arose whereby the WAN accelerator stopped accelerating HTTPS traffic and reverted to a mode where it simply passes packets through uninspected. This is indicated on the traffic graph shown in Fig. 6, where the throughput drops from around 800 Mb/s (where traffic is being accelerated) to approximately 250 Mb/s (where traffic has stopped being accelerated), which is the default setting of the WAN emulation device.

Without instrumentation installed, thereby adhering to traditional end-to-end external test points, the traffic graph output from the load tester would act as the key evidence of the problem. Analysis of the heat map output from the instrumentation, however, indicated an I/O disk spike event just prior to the point at which optimization stops, and can actually be correlated with a traffic dip occurring prior to the major drop in traffic shown in the traffic graph of Fig. 6. This observation, garnered from the instrumentation data, resulted in the vendor focusing their attention on the virtual WAN acceleration behavior prior to the traffic drop. The eventual explanation for the behavior was attributed to the detection of a packet that could not be decrypted by the WAN accelerator and hence, as part of security best practice, reverted to a safer mode of operation where no optimization/acceleration is performed (thus passing packets through transparently). On the basis of prior experience, it was suggested by the vendor that such behavior would most likely be due to the particulars of the load tester's SSL settings, which in some cases can generate "bad packets." Review of the load tester setup identified a configuration option that deactivated specific SSL settings, which resulted in the test being successfully run with no cessation of the optimization/acceleration capability.

The slightly unexpected but no less compelling insight gleaned during this case study was that embedded instrumentation can provide an extra layer of granularity for fault diagnostic purposes during test and validation of a specific VNF. The utilization of heat maps allows us to plot different types of metrics (normalized) and to quickly observe the behavior of tens, hundreds, or even thousands of metrics simultaneously over a period of time. This feature is particularly useful for anomaly detection/fault finding scenarios, as illustrated by this example.

Conclusions and Further Work

NFV will undoubtedly present network operators with a number of significant implementation challenges, many of which are already being addressed through collaborative activities across the industry. This article has addressed the specific problem of test and diagnostics, whereby VNFs must be validated — prior to operational deployment — in terms of functionality, performance, robustness, and manageability. In some ways, thorough VNF validation is more demanding than standalone physical hardware, since there are many more possible permutations of hardware build and hypervisor, not to mention the deployment of a range of VNFs from different vendors co-resident on the same server (i.e,. the multi-tenant or noisy neighbor problem).

By devising distinct case studies using very different VNFs with unique characteristics subject to testing, the benefits of embedded instrumentation have been highlighted. The main conclusions that can be drawn from the BT/Intel tests with integrated instrumentation software are as follows:
- The integration of the instrumentation did not negatively affect the "baseline" performance of the VNFs.
- The use of analytics to capture certain performance parameters complemented the load testing capability by enabling additional layers of insight and visualization.
- The increased granularity of the performance characteristics view afforded by embedded instrumentation enabled fine-tuning of allocated server resources (vCPU cores) in a manner that provided a better match with expected traffic loads, hence

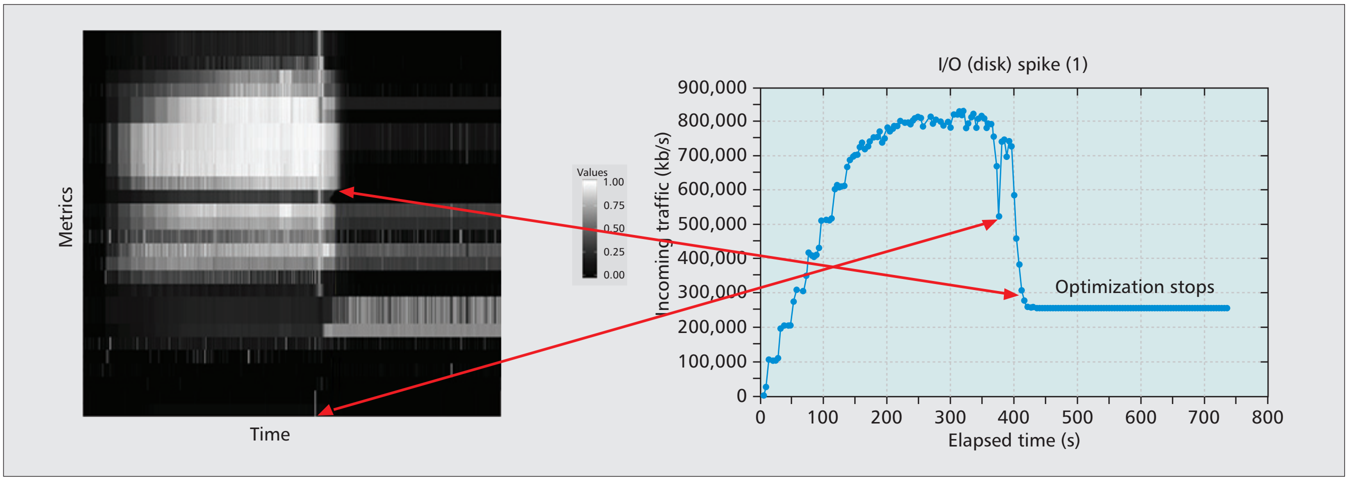


Figure 6. Virtual WAN acceleration fault diagnostics.

optimizing the set-up. This approach has significant techno-economic advantages over traditional methods, and lends itself to fully maximizing the promise of resource elasticity offered by NFV.

- The granular view of the performance characteristics afforded by the embedded instrumentation led to the diagnosis of a nonoptimal configuration of the end-to-end setup of the WAN acceleration use case. Without the instrumentation data, this issue might otherwise have been left unresolved; at best, it would have taken significantly longer to identify the root cause.

As the deployment of VNFs into carrier networks gathers momentum, the use of embedded instrumentation to complement existing external testing approaches and internal VNF metric capabilities is likely to grow in importance and utility. There are, however, many interesting problems that have yet to be explored in any significant depth. Some of these areas of interest include, but are not limited to:

- Further investigation of use cases involving multiple instances of the same VNF, as well as a range of different VNFs on the same physical server. The latter scenario represents concerns regarding noisy neighbor effects; that is, can we be certain the presence of one VNF is not detrimentally affecting the performance of another VNF?
- Investigation of more complex network services comprising multiple VNFs deployed in both a single data center or across multiple geographically dispersed data centers.
- Using results from instrumentation to devise planning rules for VNFs co-resident on servers, exposing platform-specific features to improve VNF workload placement decisions, hypervisor type, and range of VNFs planned for support on a particular server.
- Identifying how VNFs scale and perform reliably in commodity cloud environments rather than on bespoke hardware nodes that are optimized for that specific function.

It is likely that the rate at which NFV deployments occur across industry in the coming years will be influenced to some degree on how comprehensively these technical challenges are addressed.

Acknowledgments

BT and Intel acknowledge support from William Tai of Riverbed Technology for his assistance with case study 2.

References

[1] R. Jain and S. Paul, "Network Virtualization and Software Defined Networking for Cloud Computing: A Survey," *IEEE Commun. Mag.*, vol. 51, 2013, pp. 24–31.
[2] ETSI, "Network Functions Virtualisation (NFV) — Network Operator Perspectives on Industry Progress," White Paper #3, Oct. 2014.
[3] ETSI ISG, "Network Functions Virtualisation (NFV): Architectural Framework," 2013.
[4] ETSI ISG, "Group Specification-NFV Performance and Portability Best Practises," NFV-PER 001 v1.1.1, 2014.
[5] B. Gregg, *Systems Performance: Enterprise and the Cloud.*, 1st ed., Prentice Hall, 2014.
[6] S. Plamondon, "How WAN Optimization Works," *Riverbed Connections*, Aug. 2013, http://www.riverbed-news.com/2013/08/how-wan-optimization-works/.

Biographies

Paul Veitch (paul.veitch@bt.com) holds M.Eng. and Ph.D. degrees from the University of Strathclyde, Glasgow, United Kingdom. He joined BT at Martlesham Heath, Ipswich, United Kingdom, in 1996, and worked on various aspects of broadband transmission, multi-service platforms, and 3G mobile infrastructure design before joining Verizon Business (UUNET) in 2000. He returned to BT in 2003, and was infrastructure design authority for IP VPN and BT Consumer networks before joining a small team of network innovation specialists in 2012 to lead on NFV proof-of-concept validation and business development.

Michael J. McGrath (michael.j.mcgrath@intel.com) is a senior researcher at Intel Labs Europe. He holds a Ph.D. from Dublin City University. He has been with Intel for 15 years, holding a variety of operational and research roles. He is currently a researcher in the FP7 T-Nova project, which is focused on the concept of network functions as a service (NFaaS). He is also a researcher at the BT Intel Co-Lab based at Adastral Park, United Kingdom, which is focused on research relating to the deployment of VNFs in carrier grade network env(ironments. Michael has co-authored more than 25 peer reviewed publications.

Victor Bayon (victor.bayon-molino@intel.com) is a senior researcher at Intel Labs Europe and a member of the BT Intel Co-Lab based at Adastral Park. His main area of interest is the development of instrumentation, performance analysis, and performance management tools, and their application to different domains such as NVF in carrier grade environments. He holds a Ph.D. from Nottingham University, United Kingdom.

OPTICAL MULTI-SERVICE CORE AND ACCESS NETWORKS

Osman S. Gebizlioglu

Vijay Jain

The emerging trends of 2013 grew to dominate the two most notable developments in 2014: 1) the packet optical transport system (P-OTS) concept with the service provider goal to build a common network infrastructure by integrating Ethernet, IP/MPLS, and DWDM; 2) intensifying focus and attention on network functions virtualization (NFV) and software-defined networking (SDN) to give service providers the tools for more effective operation and management of communications networks in general, and optical communications networks in particular. All major global service providers have been establishing and announcing their target timeframes for transitioning to all-IP network operations, while major industry standards organizations and fora have been striving to publish their guidelines and frameworks for NFV and SDN. These important trends are expected to gain additional momentum through 2015 as the industry continues to rise to the challenges of meeting the growing global communications needs.

In this issue, we have selected six contributions that address modulation and coding in visible light communications, energy-efficient resilient optical networks, next generation elastic optical networks (three manuscripts), and 400 Gb/s Ethernet using advanced modulation formats

In the first contribution, "Modulation and Coding for Dimmable Visible Light Communication," S. H. Lee, S.-Y. Jung, and J. K. Kwon present an overview of technical considerations for enhancing the physical layer (PHY) of visible light communication (VLC), a short-range optical wireless communication technology with data rates in excess of 100 Mb/s within a coverage area of 10 m radius. Rising global interest in this technology has been driven by accelerating solid-state lighting worldwide deployments for wireless personal area networks (WPANs) based on such standards as IEEE 802.15. With VLC technology, information is conveyed by modulating the intensity of the light from an LED. Two main obstacles to VLC PHY layer design in achieving high-rate transmission are flicker and dimming. Flicker (i.e., fluctuations of the brightness of light in the course of modulating information into light sources) can have an adverse effect on human physiology. Therefore, the design of modulation techniques requires meticulous care to avoid noticeable flickering during communication. Considerable efforts have been made to produce the technical specifications in IEEE 802.15.7, which is aimed at ensuring high-data-rate compatibility and cost effectiveness based on the control of flicker and dimming. To meet regulatory requirements and enhance system performance, some prospects for modification are presented from the perspective of the PHY design for existing systems.

In the second contribution, "Energy-Efficient Resilient Optical Networks: Challenges and Trade-offs," Y. Ye, F. J. Arribas, J. Elmirghani, F. Idzikowski, J. L. Vizcano, P. Monti, F. Musumeci, A. Pattavina, and W. Van Heddeghem present the cost implications of a network-sharing scheme for different fiber to the home/passive optical network (FTTH/PON) architectures. They address the high cost of passive infrastructure and the construction work. One approach through which operators could reduce the amount of investment required to deploy fiber access networks would be network infrastructure sharing, thereby reducing the investment needed to deploy and operate an FTTH network. The authors use various metrics to understand the effect of a network sharing arrangement on costs. The results show that for the majority of cases studied, the cost per home connected and the payback period increase when employing a network sharing scheme, but the initial investment is significantly reduced. The reuse of existing passive infrastructure does not bring any cost advantage in comparison to the standalone scenario; however, it does help reduce the total cost per home connected.

In the third, fourth, and fifth contributions, project teams from the European Consortium Industry-Driven Elastic and Adaptive Lambda Infrastructure for Service and Transport Networks (IDEALIST) present their perspectives on next-generation flexible optical networks. We are pleased to acknowledge the contributions of the IDEALIST team members in the preparation of the following overview.

The three articles present results from recent investigations on optical communications with particular focus on the data plane undertaken within the IDEALIST project. The first article (third in the Series) provides the rationale for adopting the flexible network paradigm to optics, while the second and third articles (fourth and fifth in the Series, respectively) describe the actual implementation of sliceable bandwidth variable transponders (S-BVT) and the design of novel node architectures, both capable of coping with the dynamic changes of the IP traffic, respectively.

The IDEALIST Consortium was established in 2012 to seek solutions to overcome the challenges in optical networks that resulted from the continuing exponential IP traffic growth. Consequently, a substantial redesign of the deployed networks was envisioned. The current estimated compound annual growth rate (CAGR) of ~ 29 percent underlines the migration toward elastic optical networks (EONs) as the only ideal cost-effective solution to the imminent capacity crunch. In addition, the Internet of Things is soon expected to result in sky-rocketing data traffic. These considerations are part of the first article, “Next Generation Elastic Optical Networks: The Vision of the European Research Project IDEALIST.”

The IDEALIST EON architecture is built on two pillars: the S-BVT and the flexible optical node. The first pillar, described in the second article, "Next Generation Sliceable Bandwidth Variable Transponder,” represents a novel transponder architecture to introduce a high level of flexibility, for example, in terms of supported range of data rates (from 100 Gb/s up to 1 Tb/s), the number of connections, and their direction. Herein, all characteristic elements can be properly configured so that the optical flow is adaptable to the instantaneous demand. The most promising transmission techniques as well as electronic and optical components, system integration, transponder programmability, and application use cases are analyzed with the goal of reducing costs and power consumption while increasing spectral efficiency.

The second pillar, reported in the third and last of these three articles, “Next Generation Optical Nodes: The Vision of the European Research Project IDEALIST,” suggests a novel network element designed to increase the flexibility, scalability, resilience, and adaptability to today’s telecommunications systems. The proposed optical node comes as a node architecture solution able to deal with the high levels of uncertainty in communication networks following the rapid evolution of services and needs, apart from the capability to cope with the increased traffic demand. The IDEALIST team proposal represents an evolution with respect to the existing reconfigurable optical add/drop multiplexers (ROADMs) to efficiently address emerging network requirements. Within the IDEALIST Project, the drivers, architectures, and technologies were examined to enable these novel optical nodes along with multi-vendor traffic interoperability, optical defragmentation, and node cascadability.

In conclusion, with the assumption that the actual trends in data traffic growth remain stable, the IDEALIST Project Team carried out a comprehensive analysis at the data plane and network levels. Performance of novel key elements such as S-BVT and flexible optical nodes were investigated. The novel concept of a multi-flow optical module implementing a flexible subcarrier element and its performance, using different advanced transmission methods, was compared. Concerning the optical node architecture, two innovative and complementary solutions that can handle dynamic traffic and introduce a high level of scalability and flexibility were also examined.

Overall, all the innovations described are expected to lead the way to an optimized bandwidth utilization that will constitute one of the pillars of next generation optical networks. They clearly define the path for upgrading toward EONs so that the imminent capacity crunch can be significantly postponed.

In the sixth contribution to the Series, “400 Gigabit Ethernet Using Advanced Modulation Formats: Performance, Complexity, and Power Dissipation,” J. L. Wei, Q. Cheng, R. V. Penty, and I. H. White present possible architectures for 400 Gigabit Ethernet links based on advanced modulation formats. Their optical link power budget, digital complexity, and power dissipation are compared. The challenges of implementing the physical layer are discussed. Internet traffic continues to grow exponentially, fueled by applications such as high definition TV, video on demand, and cloud computing. This drives the need for massive data centers with flexible scalability. Optical interconnects play a critical role in intra- and inter-data-center communications. CPU computing speed and aggregate network traffic double every 18 months. Today’s fastest standardized Ethernet rate is 100 Gb/s, completed in 2010 (IEEE 802.3ba), with the previous two rates being 10 Gb/s and 40 Gb/s. In accordance with the 4× and 10× bit rate scaling model, IEEE 802.3 started the 400 Gigabit Ethernet (GbE) Study Group in March 2013, which successfully justified the creation of the IEEE P802.3bs 400GbE Task Force on May 2014. The 400 GbE Task Force agreed on objectives for both single-mode fiber (SMF) and multi-mode fiber (MMF) links to cover the historical reaches of 100 m over MMF, and 2 km and 10 km over SMF. In addition, a new 500 m SMF objective was agreed to support the longer intra-data-center connection required for the emergent massive data centers, which MMF cannot support at 400G.

In this first Optical Communications Series (OCS) issue of 2015, we thank all authors and reviewers for their valuable contributions to the OCS in the past and invite submissions on all aspects of optical communications technologies, with our best wishes to you in 2015 and beyond.

Modulation and Coding for Dimmable Visible Light Communication

Sang Hyun Lee, Sung-Yoon Jung, and Jae Kyun Kwon

Sang Hyun Lee is with Sejong University.

Sung-Yoon Jung and Jae Kyun Kwon are with Yeungnam University.

Abstract

The recent interest in short-range optical wireless communication technology driven by the widespread deployment of solid state lighting has led to significant efforts on the standardization of visible light communication. In such efforts, the consideration of dimming support poses fundamental challenges on the VLC system design that have not been addressed elsewhere. This article overviews the technical considerations for enhancing the VLC physical layer by summarizing the state-of-the-art advancements in modulation and coding technologies dedicated to VLC systems. In addition, the technical challenges for system enhancement under lighting restrictions are described.

Introduction

Short-range wireless technology has recently been a key driver leading the surge of deployment for wireless personal area networks based on wireless standards such as the IEEE 802.15 series of standards. As a viable candidate for enabling technologies, visibile light communication (VLC) technology using light emitting diodes (LEDs) [1–3] has captured the interest of academia and industry due to the widespread transition from incandescent/fluorescent lighting to energy-efficient solid-state lighting. The fast switching control feature of LEDs enables the use of devices for lighting purposes in data transmission through the visible light spectrum as shown in Fig. 1. VLC conveys information by modulating the intensity of light emitted from LEDs at a rate faster than human perception. VLC can offer data rates in the megabits per second range over short distances in the visible light spectrum, where low-cost optical sources are readily available, and the best performance is obtained using a cost-efficient arrangement implemented with only LEDs and photodiodes (PDs).

Two main obstacles for VLC physical (PHY) layer design in achieving high-rate transmission are flicker and dimming [4]. Flicker, fluctuation of the brightness of light in the course of modulating information into light sources, can have an adverse effect on human physiology. Therefore, the design of modulation techniques requires meticulous care to avoid noticeable flickering during communication. The support of dimming is another major feature, in which a VLC device is equipped for the main functionality of lighting. This requirement stems externally from users who can dim the light intensity over all ranges from off to full, and proper VLC operation under arbitrary dimming is strongly recommended.

Considerable efforts have been made to produce the technical specifications in IEEE 802.15.7, which realizes high-data-rate compatibility and cost effectiveness based on the consideration of flicker and dimming. To act in concert with the expansion of the technology and the demands for a higher data rate, the future enhancement of standard systems is an important target. To meet regulatory requirements and enhance system performance, some potentials for modification are presented from the perspective of the PHY design for existing systems.

This article discusses the benefits, potentials, and challenges associated with the use of various modulation and coding techniques as well as their advantages in dimming support. The essential characteristics, such as optical transmission/reception and LED physical characteristics, are also considered to ensure effective operation with dimming support, and the specific technologies of several efforts to overcome them are reviewed. In addition, the challenges ahead in terms of the technical issues and regulations are addressed.

VLC Transmission with Dimming Support

Data transmission in VLC is carried out by conveying information through temporal intensity changes in the optical pulse. The modulation and demodulation proceed with intensity modulation (IM) and direct detection (DD), respectively. In IM, the LED emits pulses at different intensity levels according to different electrical digital message symbols in the wireless link. As VLC transmitters are mostly intended for white-color lighting, they are implemented with either a single phosphor-converted (PC) white LED, generated using blue LEDs with yellow phosphor, or a combination of multicolored LEDs such as red-green-blue (RGB). If a single type of

0163-6804/15/$25.00 © 2015 IEEE

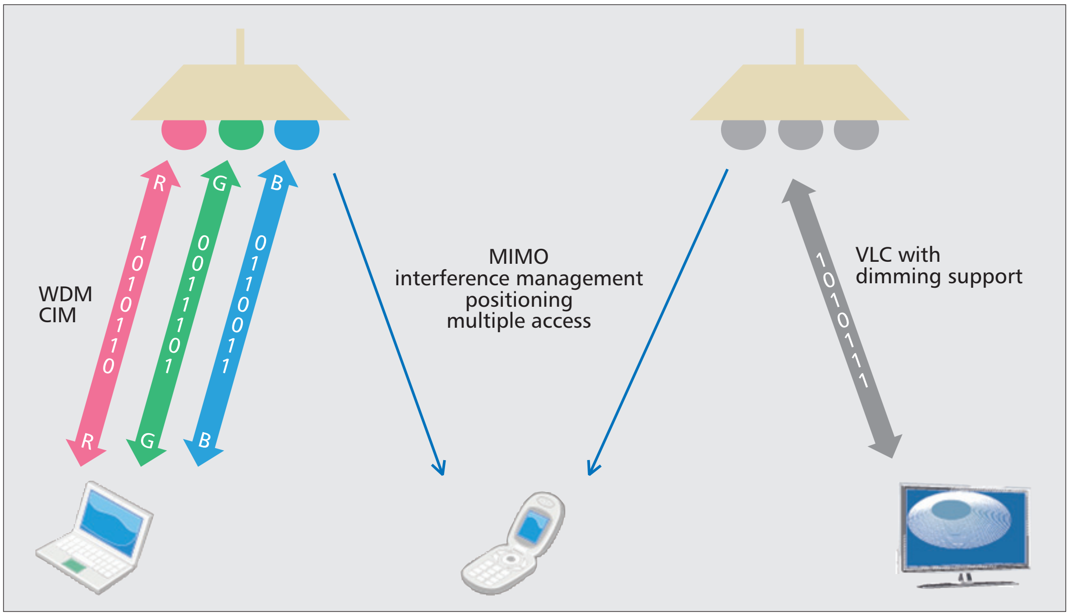


Figure 1. Configurations of VLC. VLC can be configured in various ways according to the configuration of the LED lamps: a single data stream transmission via a single PC LED, multiple data stream transmission via an array of RGB LEDs, and so on.

Some message information cannot be transmitted, or the resulting rate is reduced compared to the maximum possible rate. Both are undesirable for enhancing the data rate of VLC. Hence, dimming support challenges us to discover how to handle this inconsistency, which has not been considered in other communication systems.

LED is used to produce white-color emissions for reduced cost, only a single data stream can be transmitted, thereby exploiting a single degree of freedom in the communication link. The PD detector performs DD of the intensity signal by inducing the electrical photocurrent proportional to the received optical intensity (square law). Since VLC receivers sense the intensity of the optical signal, the demodulation is inherently non-coherent. The main impairments for DD include shot noises induced by the signal and ambient light, and the pre-amplifier noise encountered at the receiver [1]. If the ambient shot noise and pre-amplifier noise become dominant in the optical wireless links, the channel response for VLC is well approximated by the additive Gaussian noise model [5, 6].

For the dimming feature, the radiant intensity of a light source is defined as the optical power emitted per solid angle. The extent of dimming can generally be measured from the average intensity of the optical signal. The physical characteristics of the optical intensity channel impose a peak optical power constraint to prevent LED failure and to secure eye safety, equivalently viewed as constraints on peak signal intensity. The demodulation that detects the normalized power stipulates that the transmitted signal always has a non-negative signal level. In addition, physiological safety regulations along with dimming support place a limit on the average emitted signal level. Under the maximum level constraint, the capacity of the optical Gaussian noise channel is finite, and reliable communication can be achieved using two-level modulation techniques, albeit requiring exponential spectrum use [7].

The range of signal intensities can be represented in a normalized region of [0,1], where zero and one correspond to the off and full levels of intensity, respectively. The intensity of the optical signal conveying a uniform message gives rise to 50 percent (or 1/2) dimming on average. The human eye normally perceives the average illuminance instead of the instantaneous changes if the intensity changes faster than 150–200 Hz. Thus, VLC systems are designed to modulate the signal optically such that the intensity varies at a frequency higher than the inverse of the maximum flickering time period (MFTP) [4], in which the human eye cannot perceive the temporal change. Let $d \in [0,1]$ be the average ratio of the duration of the optical emission to that of the message data frame. This characterizes the dimming target of the VLC system. If the dimming target is set to $100d$ percent, the time-averaged intensity should be equal to $100d$ percent within an MFTP. The dimming target associated with $d \neq 0.5$ may require non-uniform signal levels in the data frame, resulting in non-uniform frequencies of message symbols. Therefore, some message information cannot be transmitted, or the resulting rate is reduced compared to the maximum possible rate. Both are undesirable for enhancing the data rate of VLC. Hence, dimming support challenges us to discover how to handle this inconsistency, which has not been considered in other communication systems. This article provides an overview of modulation and coding techniques developed for this purpose.

Modulation for Dimmable VLC

One-Dimensional Modulation Techniques

Modulation techniques to remedy the above challenges are classified into two categories, which are referred to as the time-domain approach and the intensity-domain approach. The time-domain approach adds compensation symbols of two levels (ON and OFF) within an MFTP to match the dimming target, whereas the intensity-domain approach changes either the intensity levels or frequencies of occurrence for

These time-domain approaches are simple to implement but do not allow high data-rate support because of bandwidth wastage by the compensation time intervals. This limitation becomes severe as the dimming target deviates from 50 percent.

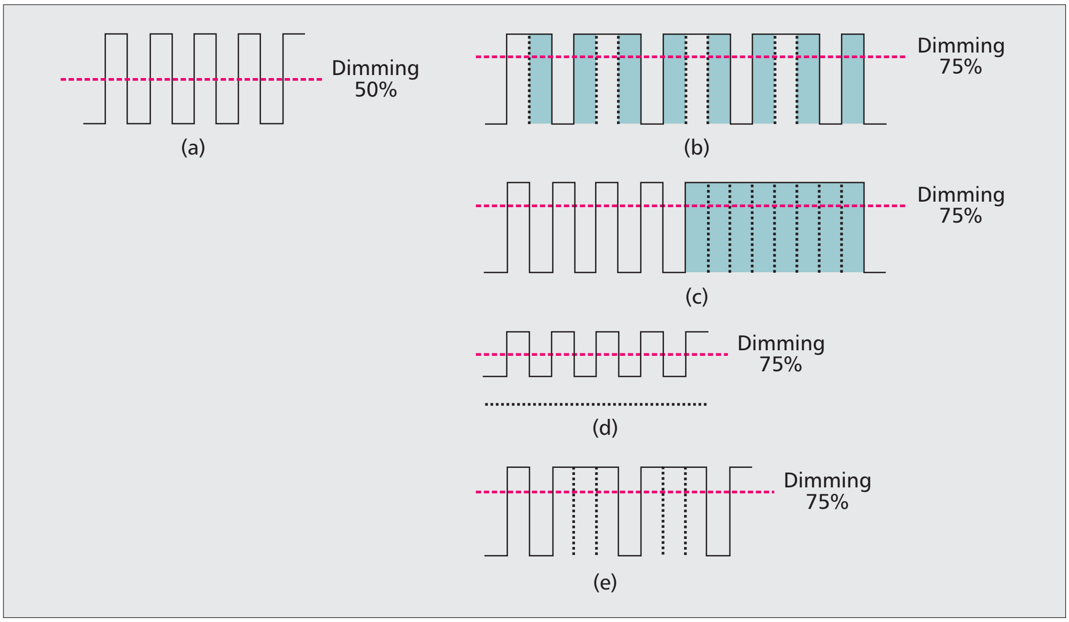


Figure 2. Examples of dimming support in binary modulations for VLC. Message symbols are normally given uniform probability. Four different types of methods can be used to offer dimming support with message symbols: a) original signal; b) intra-pulse insertion adds a compensation symbol to every message symbol; c) inter-pulse padding attaches a series of compensation symbols next to a series of message symbols; d) bias scaling changes the DC bias such that the time-averaged intensity is equal to the dimming target; e) distribution-adaptation maps a sequence of message symbols into another sequence of symbols of which the occurrence of ones and zeros adapts to the dimming target.

the message symbols. The time-domain approach is then classified into intra-pulse insertion and inter-pulse padding according to the position at which the dimming compensation symbols are placed in the data frame. Also, the intensity-domain approach includes bias scaling and distribution adaptation. The former alters the DC bias level by scaling the spacing between the intensity levels, and the latter adjusts the average frequency of the intensity levels that occur in an MFTP.

This article focuses on each of these approaches. The intra-pulse insertion (shown in Fig. 2b) places the compensation symbols into an optical pulse to match the overall dimming in an MFTP. This can be viewed as pulse-width-modulation (PWM)-based dimming control, in that the duty cycle of the optical pulse varies with the dimming target. For example, to meet a target of 75 percent dimming, the transmission of the ON and OFF binary message symbols with uniform probability is followed by ON compensation symbols, which convey no message, of the same duration. In Fig. 2b, an ON level of the same width is inserted into each pulse. This can be extended to more sophisticated techniques, such as pulse-amplitude modulation (PAM) and orthogonal frequency-division multiplexing (OFDM) [8]. A message is first modulated using those modulations, and the resulting optical pulse width is adjusted like PWM to match the dimming target. Variable pulse-position modulation (VPPM) [4], a variation of PWM-based dimming-supporting modulation, uses PPM for modulation of the message and PWM for adaptation to the dimming target, respectively. In the PWM-based dimming control, optical pulses should be emitted for a very short period to provide support for sufficient resolution of dimming targets. However, VLC systems are normally subject to a limit on the minimum pulse width. This results in a restriction of the symbol duration, which in turn limits the data rate. Inter-pulse padding (shown in Fig. 2c) appends a compensation symbol interval to a message symbol interval. The message symbols are arranged to match the 50 percent dimming, and the compensation symbols conveying no information are added with the appropriate ON/OFF symbol proportion associated with the dimming target. The OOK-based mode (time-multiplexed OOK) [4, 9], specified in the IEEE 802.15.7 standard, uses this technique. These time-domain approaches are simple to implement but do not allow high data rate support because of bandwidth wastage by the compensation time intervals. This limitation becomes severe as the dimming target deviates from 50 percent (e.g., for a 10 percent dimming target, four times longer time durations are spent for compensation than the message).

The intensity-domain approach, however, has an advantage in terms of the data rate over the time-domain approach. The bias scaling (shown in Fig. 2d) adjusts the DC bias and scales message symbols for dimming support. The DC bias associated with 50 percent dimming is initially set with the assumption of uniform symbol probability. If the dimming target is 75 percent, as shown in Fig. 2d, the DC bias is raised such that the intensity of the OFF symbol increases to the DC bias corresponding to 50 percent, whereas the ON symbol remains at the same level. On the other hand, for a 25 percent dimming target, the ON symbol decreases to the DC bias corresponding to 50 percent, and the OFF symbol remains constant. Therefore, the spacing between the ON and OFF symbols shrinks by half. This method can also be viewed as analog dimming. Although this is computationally simple, some technical difficulties emerge by nonlinear LED emission control and significant performance degradation due to the level spacing reduction. Distribution adaptation (shown in Fig. 2e) adjusts the frequency of the symbol levels such that the expectation of the resulting level distribution matches the dimming target. If the dimming target is 100d percent in the binary OOK modulation, the corresponding frequencies of the ON and OFF symbol should be d and $1 - d$, respectively. Thus, the message symbols with uniform probability need to be represented in binary intensity levels with non-uniform probability. This conversion is achieved via a decoding process related to lossless compression, such as inverse source coding (ISC) [10]. It is noticeable that the ISC yields the best achievable limit for maximally distinguishable messages and is very efficient for the extreme dimming target ranges in a low-noise environment.

Multi-level modulation can exploit several advantages of the intensity-domain approach in modulating the message on the intensity level of the symbol. Although binary modulation is a simple practical modulation technique, it allows a limited range of data rate improvement within finitely band-limited environments. Compared to high-bandwidth systems where binary techniques are found practical, low-bandwidth systems require multi-level modulations. Since multi-level modulations achieve reliable communication in band-limited VLC channels [11], a more practical choice for data rate enhancement is PAM. For PAM configurations with restrictions on the peak and average intensity, the capacity-achieving message distribution under a Gaussian noise channel is discrete with a finite number of intensity levels [12]. Although such a distribution can be determined, the design of the transmission technique that conveys the message according to that distribution is still a challenge with a lack of modulation techniques with non-uniform symbol frequencies. In an effort to handle this challenge, a rate-efficient multi-level modulation with arbitrary dimming support has been developed based on a distribution-adaptation method [13]. A convex optimization formulation is developed with the objective of maximizing the data rate with respect to the intensity level distribution, and its closed-form solution is obtained. The resulting distribution can be realized using superposition or concatenation coding schemes.

Multiple-subcarrier modulation (MSM) [14], implemented via OFDM, uses all the above methods for dimming support. OFDM conveys independent symbols with different subcarriers that can be separated orthogonally. Since the signal level of an OFDM signal can be negative, two different techniques of generating non-negative symbols suitable for IM have been proposed: DC-biased optical OFDM (DCO-OFDM) [2] and asymmetrically clipped optical OFDM (ACO-OFDM) [15]. DCO-OFDM uses the DC bias to yield a non-negative signal, whereas ACO-OFDM clips all negative levels to zero by electrical signal processing. Both techniques can accommodate all dimming methods except for distribution adaptation because an instantaneous signal level within a single OFDM symbol is hard to control by adjusting the signal level of individual subcarriers. The MSM technique realized by OFDM in general can offer a way of mitigating the fading effects caused by propagation environments (delay spread [14], atmospheric turbulence [16], etc.). However, this technique has several shortcomings for VLC systems. In IM/DD-based intensity-modulated techniques, due to the square law resulting from the power detection in optical-to-electrical conversion, an optical signal of a longer symbol duration undergoes higher noise corruption than that of a shorter symbol duration with the same amount of message information and electrical energy consumption. Since OFDM spreads time-domain symbols of short time duration along the frequency domain with a long time duration, the associated MSM may have lower power efficiency than single-carrier approaches. The signal clipping and nonlinear LED characteristics impair the orthogonality among subcarrier channels, thereby incurring inter-subcarrier interference. Furthermore, MSM does not yield an increase in transmission efficiency and has the same data rate as a single-carrier approach because only half of the subcarriers are used to guarantee non-negative real-valued optical output.

Table 1 gives a brief summary of the properties of the above methods. Some properties incur technical challenges in system design: Low-complexity design and stable lighting support are the core issues handled from the implementation's perspective. The consideration of signal shaping and transmission capacity in a noisy channel environment are also key issues for rate/energy-saving enhancement. Color temperature shift emerges as another critical implementation issue. This variation of the LED output color originates from the changes in two factors: optical intensity level and operating temperature of the LED. Different intensity levels can be observed in slightly different colors. Therefore, multiple-level modulations are more susceptible to this color change than binary modulations. On the other hand, an increase in optical power consumption causes a temperature rise, which shifts the wavelength of the LED emission. Since power consumption varies with the dimming target, all the modulation techniques may be subject to this type of color shift.

A multidimensional arrangement comprising multiple LEDs can be used for VLC. The intensity signals emitted from different LEDs are controlled to allow multiple-data-stream transmission under certain lighting requirements.

MULTIDIMENSIONAL MODULATION TECHNIQUES

A multidimensional arrangement comprising multiple LEDs can be used for VLC. The intensity signals emitted from different LEDs are controlled to allow multiple-data-stream transmission under certain lighting requirements. There are two multidimensional configurations arranged with multiple LEDs of different colors integrated in a single position or multiple LEDs located at separate positions. If M groups of LEDs with C different colors are placed at different positions, the resulting degree of freedom is CM, and as many links are ideally available.

		Noiseless channel	Noisy channel	Extreme dimming	Complexity	Channel/ line coding	Color temperature shift	Examples
Time-domain approach	Intra-pulse insertion	×	△	×	○	○	△	PWM dimming, VPPM
	Inter-pulse padding	×	△	△	○	○	△	Time-multiplexed OOK
Intensity-domain approach	Bias scaling	○	×	△	△	○	×	Analog dimming
	Distribution-adaptation	△	○	○	×	× (more research needed)	△	ISC, MPPM, CIM

Table 1. Comparison of various modulation techniques for dimming support in VLC. The qualitative comparison among four different types of modulation techniques is presented for several properties. Three markers, ○, △, and ×, represent three qualitative levels of the listed features: good, moderate, and bad, respectively. "Noiseless channel" represents the relative effectiveness of signal shaping under noiseless transmission. "Noisy channel" represents the relative effectiveness of signal shaping under a noisy channel. "Extreme dimming" indicates the performance when dimming targets are extreme (i.e., close to zero or one). "Complexity" indicates the complexity of implementation. "Channel/line coding" is related to the compatibility with error-correcting or run-length limited coding. "Color temperature shift" indicates the vulnerability of an LED emission color change.

The message signals are assumed to span in *CM*-dimensional space. The overall intensity is given by the sum of individual optical intensities and should satisfy the dimming target. Therefore, the dimming target imposes a constraint on the total intensity, which gives rise to a linear constraint on message symbols mapped in the *CM*-dimensional space. This offers a framework to facilitate the mathematical handling of multidimensional techniques. In order to fully exploit the degree of freedom, just as many PDs are needed to realize multiple-input multiple-output (MIMO) technology, which allows the corresponding number of independent links.

In a multicolored LED configuration, the color and intensity are chosen externally by the user, and their joint consideration leads to multiple color requirements. In an RGB arrangement, a message symbol can be mapped to one point in a three-dimensional color space defined by RGB intensities. Human perception of a color is easily described in this space. Since a color depends on relative RGB intensities, the normalization by intensity reduces the three-dimensional space into the two-dimensional CIE xyY chart [4]. Color shift keying (CSK) [4] is the first multicolor technique that modulates three component intensities for message transmission. Message symbols are associated with distinct colors such that the average color matches the user's desire. Figure 3 gives an example of CSK within the gamut of colors, the set of colors represented using RGB sources, in the chart. The color gamut is given by a triangle with vertices corresponding to three sources. Given the target intensity and color, symbols are mapped on the chart such that the average color is equal to the target color. However, the constant intensity forces the use of only a two-dimensional region instead of the entire three-dimensional space. Furthermore, PDs normally have different responsivity patterns from human perception. This property introduces a new three-dimensional space, referred to as a signal space, which describes message symbols detected by PDs. Since the detection performance depends on distances among symbols in the signal space rather than the color space, the data rate and color use can be enhanced by handling messages in the signal space. For this purpose, color-intensity modulation (CIM) [17] has been proposed that enables the modulation of both intensity and color. Since CIM uses the entire three-dimensional region to optimize symbol mapping in the signal space and adjust color requirements in the color space simultaneously, it outperforms CSK.

In a configuration of LEDs located in different positions [18], the line-of-sight component is dominant in the VLC signal, which does not provide sufficient scattering for PD to distinguish multiple inputs. This hinders applying spatial multiplexing techniques in VLC. For performance enhancement, the beam width and field of view can be reduced at the cost of lowering the illumination range.

Coding for Dimmable VLC

Compared to the technical progress made in the development of modulation techniques, the design of efficient forward error correction (FEC) schemes with dimming support has rarely been addressed. In the IEEE 802.15.7 standard, Reed-Solomon (RS) codes and their concatenations with convolutional codes are adopted as FEC. At transmitter, FEC encoding is followed by run-length-limited (RLL) line encoding, which is implemented using Manchester, 4B6B, and 8B10B codes, to provide DC balance and flicker mitigation. Thus, the RLL codewords — characterized by a small number of identical symbols — are encoded after the FEC based on RS codes.

To increase the data rate, the source coding aspect of the code design can be emphasized instead of the RLL line coding. This indicates that a codebook containing as many codewords as possible is constructed for a given codeword length. For this target, multiple-PPM (MPPM) has been proposed to construct a codebook of the theoretically best achievable rate [9]. However, MPPM lacks fast encoding/decoding rules that exponentially map many messages, and their implementation is impractical. On the other hand, the ISC [10] can be realized using inverse mapping of lossless compression for OOK/PAM signals. This is implemented using a reversal of practical Huffman encoding/decoding rules listed in Tables 2 and 3. Since the Huffman code allows optimal lossless compression, the resulting dimming adaptation approach is optimal with respect to the source coding gain. Therefore, the ISC establishes the maximally achievable data rate for a specific dimming target.

The line coding and ISC may have weak error correction capability. Since the encoded symbols are fed directly into the optical modulator, the decoded symbols at the receiver are vulnerable to error propagation. Therefore, the unrecovered symbols are marked as erasures for the RS decoder. This limits the length of the line code such that the number of erasures does not go beyond a correctable range, although this reduces the potential to obtain additional coding gain from the use of longer line coding. The OOK mode, instead of a specific dimming support technique, uses inter-pulse padding where Manchester line coding maintains the duty cycle of the signal to exactly 50 percent, and the compensation symbols are inserted into the data frame. However, this limits the data rate. In addition, FEC encoding followed by line coding results in the adherence of the receiver to hard-decision decoding, which incurs incompatibility with advanced iterative coding schemes [4].

To design an efficient FEC scheme that achieves arbitrary dimming targets in noisy environments, special consideration is needed for codeword symbol patterns. In principle, collection of binary sequences with non-uniform probability equal to the dimming target generates a codebook that achieves the dimming target. To offer error correcting capability, only a subset of codewords is chosen such that its elements are separated sufficiently from each other in Hamming distance. For practical considerations, encoding/decoding rules need to have useful properties, such as linearity, although the codebook exactly satisfying an arbitrary dimming target is inherently nonlinear. The codebook generated from a linear codebook such that all codewords have constant Hamming weight may be a useful choice for dimming support. This approach is applied to Reed-Muller (RM) codes to obtain a codebook that supports only $100/2^b$ percent ($b = 0, 1, \ldots$) dimming targets [19]. For an arbitrary dimming target, this is combined with compensation symbol insertion. However, this supports a limited range of data rates because the resulting data rate scales with $O((\log N)/N)$. On the other hand, a constant-weight codebook that chooses only the codewords of a Hamming weight equal to the dimming target through special algebraic theory can be used for the same objective. The existence of such a constant-weight codebook with an arbitrary weight is not well understood, and the range of available code rates is limited. In addition, very little is known about encoding/decoding rules for constant-weight codes with manageable complexity.

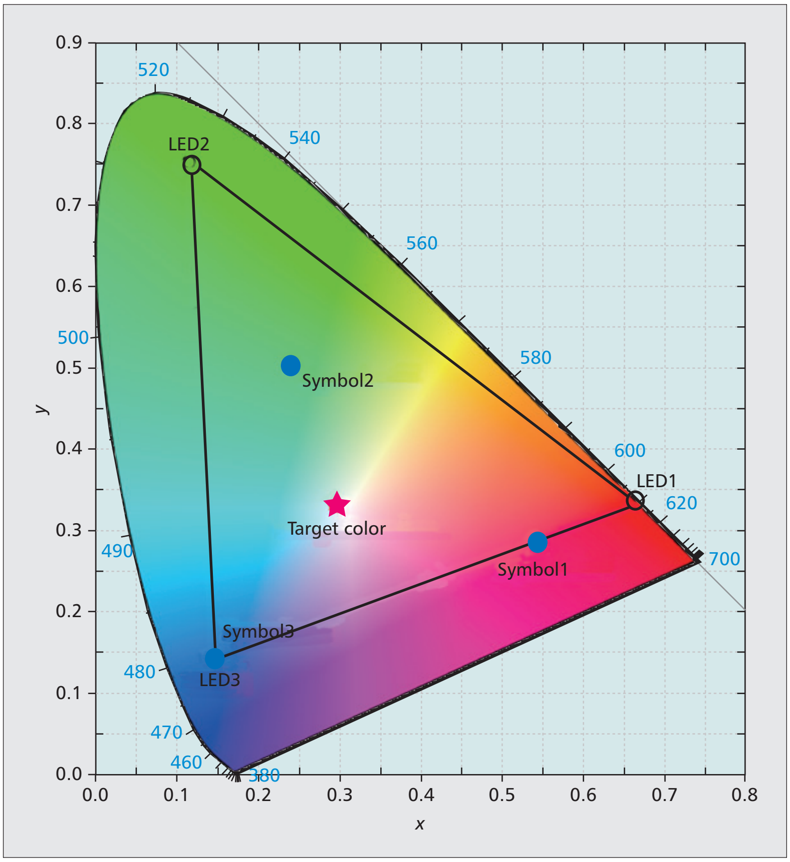


Figure 3. Example of a symbol constellation of CSK in CIE xyY color space (after [17]).

Compared to the range of achievable rates for dimmable VLC channels, the range of data rates where practically realized error correction schemes show good performance is very limited. A codebook with the exact-constant-weight property has a sub-exponential number of codewords, and the resulting code rate is strictly less than the achievable data rate. Therefore, such a constant-weight codebook is not an efficient choice for high data rate support. To extend the code rate significantly, the relaxation of the constant-weight constraint can be considered. In fact, this is the correct approach because the dimming requirement is imposed on the average value of the intensity instead of the instantaneous value. Scrambling of a linear codeword with a random sequence (i.e., masking a binary codeword with the sequence of the same length) can yield an average dimming equal to 50 percent. The weight of a codeword can be viewed as a binomial random variable, and the weight divided by the codeword length corresponds to the dimming value. Thus, the dimming value itself is a random variable with mean equal to the dimming target and variance inversely proportional to the codeword length. The adaptation of the arbitrary dimming target can be achieved by puncturing. Turbo codes are applied to obtain a large degree of data rate improvement [13, 20]. However, puncturing brings limited enhancement of the data rate, especially in a

Symbol/length	Probability	Codeword/length
0/1	0.3	00/2
10/2	0.21	01/2
11/2	0.49	1/1

Table 2. Example of Huffman encoding. This table is originally determined for lossless compression of the binary signals with non-uniform probabilities. However, this can be used in the demodulation of the optical signal at the receiver for VLC by mapping the detected symbols to the original message symbols in a sequential manner. The ISC employs reverse mapping of a lossless compression process where a certain non-uniform proportion of ones and zeros in the input signal are transformed to be as uniform as possible. This can be implemented by either table lookup or tree search. For example, a dimming target of 70 percent is given. A binary sequence with 70 percent ones and 30 percent zeros can be shortened losslessly via Huffman encoding. The resulting compression ratio is close to binary entropy $H(0.7)$, which is equal to the maximally allowed rate for a dimming target of 70 percent.

Symbol/length	Probability	Codeword/length
00/2	0.25	0/1
01/2	0.25	10/2
1/1	0.5	11/2

Table 3. Example of inverse Huffman encoding. This table is obtained by inversely mapping the input and output of Table 2. This table maps a series of message symbols normally with uniform probability of ones and zeros to a series of symbols in which the time average meets the dimming requirement. The inverse Huffman coding is carried out using the table that corresponds to the inverse mapping of Table 2. Since the decompression ratio is equal to the reciprocal of the entropy associated with the target dimming, the resulting dimming of an ISC-encoded signal comes close to the target dimming. The ISC symbols are decoded straightforwardly using the compression table of Table 2. Note that although the dimming target of 70 percent is not precisely met, it can be achieved for sufficiently long codewords.

low dimming regime. Therefore, additional techniques are expected for an extension of the range of available data rates.

Technical Challenges for Enhanced VLC Systems

This section overviews the key technical challenges faced by VLC PHY systems:

• Design of capacity-approaching coding techniques: The error correction performance of the latest coding techniques developed for dimmable VLC is far from the capacity in the low/high dimming regime. Almost all practical FEC techniques employ time-domain dimming approaches, such as intra-pulse insertion and inter-pulse padding, which degrade the data rate severely. To be precise, the data rate achieved by insertion/padding schemes has a linear relationship with the dimming target, while the capacity is given by the binary entropy of the dimming target. Therefore, the difference between the capacity and practically achieved data rate becomes larger as the dimming target deviates from 50 percent. This suggests that existing practical schemes do not fully exploit a potentially favorable property leading to data-rate enhancement. For example, the use of variable padding symbol positions can lead to an increase in data rate.

• Consideration of the LED physical characteristics: The LED is subject to chromaticity (color) shift and nonlinearity caused by the change of driving current and operating temperature by Joule heating. This change in intensity causes a slight change in light color and brightness. This poses a critical challenge for achieving multi-level transmissions implemented with LEDs. The impact is more serious in RGB LEDs than PC white LED. Furthermore, proper control technologies for color mixing are needed under changes to various parameters (e.g., temperature, driving current, aging). A nonlinear input/output relationship of an LED is an unavoidable issue that requires fine intensity control of LED emission.

• Flicker regulations: Worldwide legislative changes to regulation of flicker in LED lighting applications are ongoing. In Japan, there is a regulation for brightness change in LEDs within the range 100–500 Hz, provided by Product Safety Electrical Certification. In North America, the progress in defining flicker is made toward minimizing the relative amount of changes in light intensity by introducing some metrics for flicker measurements, such as percentage flicker and flicker index, suggested by the Illuminating Engineering Society. According to these metrics, the shape of the optical pulse can also have an impact on the flicker effect. Therefore, intensive cross-disciplinary research on flicker from physiology and engineering perspectives is required.

Conclusion

This article reviews major technical design considerations for VLC systems operating with LEDs and PD detectors. High-rate transmission over a broad visible light spectrum and dimming support are identified as the two main driving forces that motivate the creation of new enhanced specifications in VLC. To create technically sound enhancement, the system should satisfy regulatory requirements and optimize system performance. Modifications in coding and modulation are necessary to support adaptive dimming, whereas new coding schemes are essential for performance enhancement. From this perspective, the system design considerations highlighted in this article are intended to serve as guidelines to fulfill these requirements and to motivate further research in design of transmission techniques compatible with dimming support.

Acknowledgment

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2014R1A1A1006082).

References

[1] T. Komine and M. Nakagawa, "Fundamental Analysis for Visible-Light Communication System Using LED Lights," *IEEE Trans. Consumer Electronics*, vol. 50, no. 1, Feb. 2004, pp. 100–07.

[2] H. Elgala, R. Mesleh, and H. Haas, "Indoor Optical Wireless Communication: Potential and State-of-the-Art," *IEEE Commun. Mag.*, vol. 49, no. 9, Sept. 2011, pp. 56–62.

[3] K. Lee, H. Park, and J. R. Barry, "Indoor Channel Characteristics for Visible Light Communications," *IEEE Commun. Lett.*, vol. 15, no. 2, Feb. 2011, pp. 217–19.

[4] S. Rajagopal, R. D. Roberts, and S.-K. Lim, "IEEE 802.15.7 Visible Light Communication: Modulation Schemes and Dimming Support," *IEEE Commun. Mag.*, vol. 50, no. 3, Mar. 2012, pp. 72–82.

[5] S. M. Moser, "Capacity Results of an Optical Intensity Channel with Input-Dependent Gaussian Noise," *IEEE Trans. Info. Theory*, vol. 58, no. 1, Jan. 2012, pp. 207–23.

[6] S. Hranilovic and F. R. Kschischang, "Optical Intensity-Modulated Direct Detection Channels: Signal Space and Lattice Codes," *IEEE Trans. Info. Theory*, vol. 49, no. 6, June 2003, pp. 1385–99.

[7] J. G. Smith, "The Information Capacity of Amplitude-and Variance-Constrained Scalar Gaussian Channels," *Info. and Control*, vol. 18, 1971, pp. 203–19.

[8] G. Ntogari *et al.*, "Combining Illumination Dimming Based on Pulse-Width Modulation with Visible-Light Communications Based on Discrete Multitone," *J. Opt. Commun. Net.*, vol. 3, no. 1, Jan. 2011, pp. 56–65.

[9] K. Lee and H. Park, "Modulations for Visible Light Communications with Dimming Control," *IEEE Photonics Tech. Lett.*, vol. 23, no. 16, Aug. 15, 2011, pp. 1136–38.

[10] J. K. Kwon, "Inverse Source Coding for Dimming in Visible Light Communications Using NRZ-OOK on Reliable Links," *IEEE Photonics Tech. Lett.*, vol. 22, no. 19, Oct. 1, 2010, pp. 1455–57.

[11] S. Shamai, "Capacity of a Pulse Amplitude Modulated Direct Detection Photon Channel," *IEE Proc.*, vol. 137, no. 6, Dec. 1990, pp. 424–30.

[12] T. H. Chan, S. Hranilovic, and F. R. Kschischang, "Capacity-Achieving Probability Measure for Conditionally Gaussian Channels with Bounded Inputs," *IEEE Trans. Info. Theory*, vol. 51, no. 6, June 2005, pp. 2073–88.

[13] S. H. Lee, K.-I. Ahn, and J. K. Kwon, "Multilevel Transmission in Dimmable Visible Light Communication Systems," *J. Lightw. Tech.*, vol. 31, no. 20, Oct. 15, 2013, pp. 3267–76.

[14] T. Ohtsuki, "Multiple-Subcarrier Modulation in Optical Wireless Communications," *IEEE Commun. Mag.*, vol. 41, no. 3, Mar. 2003, pp. 74–79.

[15] J. Armstrong and A. J. Lowery, "Power Efficient Optical OFDM," *Electronics Lett.*, vol. 42, no. 6, Mar. 16, 2006, pp. 370–72.

[16] W. O. Popoola and Z. Ghassemlooy, "BPSK Subcarrier Intensity Modulated Free-Space Optical Communications in Atmospheric Turbulence," *J. Lightwave Tech.*, vol. 27, no. 8, Apr. 15, 2009, pp. 967–73.

[17] K.-I. Ahn and J. K Kwon, "Color Intensity Modulation for Multicolored Visible Light Communications," *IEEE Photonics Tech. Lett.*, vol. 24, no. 24, Dec. 15, 2012, pp. 2254–57.

[18] L. B. Zeng *et al.*, "High Data Rate Multiple Imput Multiple Output (MIMO) optical Wireless Communications Using White LED Lighting," *IEEE JSAC*, vol. 27, no. 9, Dec. 2009, pp. 1654–62.

[19] S. Kim and S.-Y. Jung, "Novel FEC Coding Scheme for Dimmable Visible Light Communication Based on the Modified Reed-Muller Codes," *IEEE Photonics Tech. Lett.*, vol. 23, no. 20, Oct. 15, 2011, pp. 1514–16.

[20] S. H. Lee and J. K. Kwon, "Turbo Code-Based Error Correction Scheme for Dimmable Visible Light Communication Systems," *IEEE Photonics Tech. Lett.*, vol. 24, no. 17, Sept. 1, 2012, pp. 1463–65.

Biographies

SANG HYUN LEE [M] (sanghyunlee@sejong.ac.kr) received his B.S. and M.S. degrees from Korea Advanced Institute of Science and Technology (KAIST) in 1999 and 2001, respectively, and his Ph.D. degree from the University of Texas at Austin in 2011. Since 2014, he has been with the Department of Information and Communications Engineering, Sejong University, Seoul, Korea. His research interests include coding/modulation technology, optimization and their applications to optical wireless communication, communication networks, and bio/material science.

SUNG-YOON JUNG [M] (syjung@ynu.ac.kr) received his B.S. degree from Korea University in 2000 and the M.S. and Ph.D. degrees from KAIST in 2002 and 2006, respectively. He was with Samsung Electronics from 2006 to 2009. Since 2009, he has been with the Department of Electronic Engineering, Yeungnam University. His research interest is communication signal processing and its application to optical wireless communication, UWB communications, and nanonetworks.

JAE KYUN KWON [SM] (corresponding author) (jack@yumail.ac.kr) received his B.S., M.S., and Ph.D. degrees from KAIST in 1996, 1998, and 2003, respectively. Since 2006, he has been with the Department of Electronic Engineering, Yeungnam University, Gyeongsan, Korea. His research interests include visible light communications, wireless location, and underwater communications.

Energy-Efficient Resilient Optical Networks: Challenges and Trade-offs

Yabin Ye, Felipe Jiménez Arribas, Jaafar Elmirghani, Filip Idzikowski, Jorge López Vizcaíno, Paolo Monti, Francesco Musumeci, Achille Pattavina, and Ward Van Heddeghem

Yabin Ye and Jorge López Vizcaíno are with Huawei Technologies Duesseldorf GmbH.

Felipe Jiménez Arribas is with Telefónica I+D.

Jaafar Elmirghani is with the University of Leeds.

Filip Idzikowski is with Poznan University of Technology.

Paolo Monti is with KTH Royal Institute of Technology.

Francesco Musumeci and Achille Pattavina are with Politecnico di Milano.

Ward Van Heddeghem is with Ghent University — iMinds.

ABSTRACT

Energy efficiency and resilience are two well established research topics in optical transport networks. However, their overall objectives (i.e., power minimization and resource utilization/availability maximization) conflict. In fact, provisioning schemes optimized for best resilience performance are in most cases not energy-efficient in their operations, and vice versa. However, very few works in the literature consider the interesting issues that may arise when energy efficiency and resilience are combined in the same networking solution. The objective of this article is to identify a number of research challenges and trade-offs for the design of energy-efficient and resilient optical transport networks from the perspective of long-term traffic forecasts, short-term traffic dynamics, and service level agreement requirements. We support the challenges with justifying numbers based on lessons learned from our previous work. The article also discusses suitable metrics for energy efficiency and resilience evaluation, in addition to a number of steps that need to be taken at the standardization level to incorporate energy efficiency into already existing and well established protocols.

INTRODUCTION

Traffic in core networks has been growing at a yearly rate of 45 percent since 2004 [1]. The ever increasing traffic levels have forced operators to focus on ways to increase the network capacity accordingly. However, such an increase is expected to bring higher energy consumption. The worldwide electricity consumption of communication networks was estimated to be 350 TWh in 2012 (or nearly 2 percent of the worldwide electricity consumption), showing an average annual growth rate of 10 percent since 2007 [2]. Therefore, energy efficiency is becoming one of the key design parameters for planning and operating today's telecommunication networks. In this regard, optical transport solutions might be beneficial because they reduce the number of opto/electrical/opto (O/E/O) operations, which are very costly from power consumption point of view.

A critical issue when dealing with optical transport networks is the amount of traffic that can be disrupted by a network fault, an aspect that needs to be addressed properly if certain quality of service (QoS) levels are to be guaranteed to the end user. This makes it necessary to reserve some redundant resources to ensure *network resilience* (i.e., the ability of the network to provide and maintain an acceptable level of service in the face of different faults). These redundant resources are then used to reroute traffic, bypassing the failed network element(s) over the so-called backup path(s).

In general, resilient schemes can be divided into two categories: *restoration* or *protection*. Their main difference lies in the way they compute the backup paths. In a restoration approach, the backup paths are computed on the fly (i.e., only after a failure occurs), while in a protection approach the backup paths are precomputed, with the backup resources already reserved and ready to be used in the occurrence of a failure. Each scheme has its pros and cons. Restoration allows for more efficient use of network resources, but has longer recovery time and does not provide a 100 percent recovery guarantee against failures. Protection schemes, on the other hand, have a faster recovery time and can guarantee 100 percent recovery from the failure scenario for which they were designed, but require more energy in absolute terms. Protection schemes can be further categorized in the way backup resources are used. *Dedicated schemes* do not allow sharing of backup resources among multiple backup paths. *Shared schemes*, under specific conditions, allow some backup paths to use the same wavelength resources. Dedicated protection (DP) can be implemented in a 1+1 or 1:1 fashion. In the former, the traffic is duplicated over two disjoint paths, and one of the paths is selected by the destination node, whereas in the latter, the disjoint backup path may carry traffic that can be preempted in the case of a failure happening in the working path.

Operators tend to implement the resilience concepts presented above following three high-level mechanisms. The first one is network overprovisioning, that is, from the duplication of a number of line cards in a node, up to the dupli-

 0163-6804/15/$25.00 © 2015 IEEE

cation of completely geographically separated routers, nodes, and fiber links. In the latter case, the end result will be a dual plane network at one or several layers. The second mechanism is to use 1+1 protection at the optical layer, which can provide fast switchover time (< 50 ms) when a failure occurs. On the other hand, when the backup path carries the switched-over traffic (and the failed element along the primary path is under repair), this traffic is in a vulnerable state: it is unprotected against any additional failures. For this reason, while the traffic is carried by the backup path, restoration is used to replace the old (failed) working path. This is the third option (i.e., dedicated protection plus restoration). After a new backup path is computed, the network is again protected against any single (link or node) failure scenario, with the original backup path serving as the current working path, and the repaired (old) working path now acting as the backup path. Resilient networks rely on a number of duplicated resources, which are unused most of the time. This is obviously not the most energy-efficient solution.

Core networks have always been designed focusing on minimization of the network equipment and deployment cost, that is, capital expenditures (CAPEX). Given a certain request of data traffic, the definition of a network topology, and the set of equipment to be installed, a design procedure determines the processing capability of the switching/routing nodes on one hand and the capacity of the inter-node transmission links on the other hand. However, reducing operational expenditures (OPEX) is becoming increasingly important for operators, and energy consumption is one of the key contributors. Despite numerous studies focused on resilience vs. cost or QoS, energy consumption cannot simply be interchanged with cost and QoS, so previous techniques cannot be directly applied to the problem of energy-efficient resilient optical network design. First, energy efficiency has to be traded against QoS. A simple cost minimization approach may save energy by shaping traffic over fewer routes, but on the other hand reduce network reliability due to the increased impact of a cut in a consolidated route. Second, network devices implementing sophisticated energy saving functionalities could initially be more expensive than the conventional ones. Finally, backup (and working path) routes may be designed to use time-varying renewable energy sources efficiently and/or adapting the transmission to the instantaneous traffic demand, which reduce the network carbon footprint, but may not reduce or minimize the costs directly.

Two facts should be considered for energy-efficient resilient design:
- Core networks are dimensioned assuming peak traffic levels.
- The type of protection (i.e., service guarantees) can be differentiated and adapted to the specific traffic type.

Both observations provide opportunities to reduce energy consumption. As for peak traffic dimensioning, real traffic demands vary over time (i.e., the traffic demands at night are usually much lower than those during the day). This means that unused resources (line cards in nodes or even whole nodes) can be put to sleep to save energy. As for the service guarantees, not all traffic types need the same level of protection; therefore, resource redundancy for service guarantees can be restricted to a subset of traffic types, thus leading to further energy savings. In other words, in order to improve the energy efficiency of resilient optical core networks, there is a clear need for design and provisioning strategies specifically tailored to reduce network energy consumption.

Certainly, optical network design is complicated and needs to balance a number of metrics (e.g., cost, energy, resilience, and scalability). The purpose of this article is to identify the major research challenges in the relatively new field of energy-efficient resilience in optical transport networks. In this respect, our objective, rather than presenting uneqivocal solutions, is to highlight and reason about the nature of the potential options and the challenges that they present; that is, their performance trade-offs.

Core networks are generally designed and provisioned in response to three main inputs: long-term traffic forecasts, short-term traffic dynamics, and service level agreement requirements. These factors influence the design and the choice of protection and restoration mechanisms as well as the energy efficiency performance.

Research Topics/Open Problems

Core networks are generally designed and provisioned in response to three main inputs: long-term traffic forecasts, short-term traffic dynamics, and service level agreement (SLA) requirements. These factors influence the design and choice of protection and restoration mechanisms as well as the energy efficiency performance. In this section, we provide an overview to address these long-term static network architecture design choices, the protocols and hardware functionalities needed for adaptation to the short-term traffic dynamics, and the strategies needed to meet the SLA requirements in an energy-efficient fashion. Most of the presented challenges are justified by numerical results achieved in our previous work (although in different network and traffic scenarios).

Long-Term Static Architecture Network Design Choices

Network Architecture — Current optical networks are based on wavelength-division multiplexing (WDM) technologies, which operate over fixed channel spacing defined by the International Telecommunications Union —Telecommunication sector (ITU-T). WDM network architectures may consider single line rate (SLR) transmission in all the wavelength channels, or mixed line rate (MLR) transmissions, where channels transmitted with different rates coexist on the same fiber. Elastic optical networks (EONs), which allow for flexible-bandwidth transmissions and adaptive modulation, have emerged as the future technology for the optical network. The choice of a particular network architecture over another, together with the adopted resilient scheme (i.e., restoration vs. protection, or dedicated vs. shared protection), will have clear effects on energy consumption. For instance, changing the protection scheme from DP 1+1, the most reliable and more energy-consuming scheme, to more energy-efficient ones such as DP 1:1 or shared protection (SP), will result in different energy savings depending

Core networks are typically organized as multilayer networks where each layer can be viewed as a single network. Higher layers rely on the resources and services provided by the lower layers, and resilience can be provided at any layer.

on the network architecture. For example, adopting SP or DP 1:1, the energy consumption of the optical layer can be reduced up to 49 percent with respect to DP 1+1 in EON and up to 42 percent in WDM with a SLR of 40 Gb/s [3].

Embodied Energy — The energy consumed during the whole lifetime of an installed device needs to be considered (including, e.g., manufacturing and decommissioning) in the design of optical transport networks. In fact, the embodied energy accounts for approximately 70 percent of the total energy consumption of the network (i.e., operational plus embodied energy consumption) [4]. The fiber actually has the highest impact in terms of embodied energy [4] due to the large amount of material used to protect the fibers in a cable and the hundreds of thousands of kilometers of fiber cable in a core network. Therefore, from an energy perspective, it is especially important to minimize the number of redundant fiber cables in resilient networks.

Resilience at Different Layers — Core networks are typically organized as multilayer networks where each layer can be viewed as a single network. Higher layers rely on the resources and services provided by the lower layers, and resilience can be provided at any layer. Figure 1a illustrates 1+1 protection in the IP layer, whereas Fig. 1b shows the same traffic demand protected at the optical layer. Note the subtle difference of the latter with the employment of optical bypass in combination with 1+1 protection at the IP layer shown in Fig. 1c. If protection is provided simultaneously at multiple layers and there is no coordination between the layers, parallel recovery actions may take place, which can have a significant impact on the overall network stability and lead to suboptimal resource usage. Multilayer recovery schemes have been explored in earlier research; however, it is not yet clear how these strategies perform from an energy efficiency point of view. Intuitively, the best option is to provide protection at the (most energy-efficient) bottom layer, perhaps augmented with a bottom-up escalation strategy where recovery starts in the lowest detecting layer and escalates upward. However, this strategy comes with the drawback of poorer handling of higher-layer failures, and the impact of escalation timings on the overall recovery time has to be assessed. Furthermore, network operators often duplicate their backbone networks and simultaneously operate two nearly identical networks. With 1+1 protection often used at the optical transport network (OTN) layer as shown in Fig. 1d, this means in effect that each requested demand results in four times the capacity at the optical layer. Energy-wise, this is clearly not an optimal situation.

Topologies — Looking at a single layer, the network topology should keep a relatively high number of alternative paths for the traffic demands coming from the higher layers. Although a full-mesh highly overprovisioned topology is ideal from the perspective of protection, the number of deployed links and their capacities should be traded against energy consumption. This trade-off was considered in [5] for physical topology design. The constraint on minimum nodal degree equal to 2 (securing at least one alternative path) leads to negligible increase of network power consumption (0.5 percent increase with respect to a network with minimum nodal degree equal to 1 using a multihop bypass approach and symmetric traffic demand for the NSFNET network). Furthermore, the size of the network will also play an important role when trading energy efficiency and resilience. For instance, if some links become highly occupied due to traffic growth, backup paths may become much longer and require the use of signal regeneration or a more robust modulation format, with the consequent increase in energy consumption.

Open Issues in Adaptation to Short-Term Traffic Dynamics

Novel Equipment Features — The nominal power consumption of devices/equipment considered for installation must be assessed during the network design phase. Emerging equipment features, such as sleep modes and dynamic reconfiguration of modulation format and transmission reach, could provide new opportunities to reduce the power consumption of backup links. For example, more energy-efficient protection schemes could be devised to exploit sleep modes at the expense of slower recovery. However, to the best of our knowledge, there is no current information of any commercial device implementing these innovative features, meaning all related work reporting results is based on assumptions. New protection schemes could also allow multi-rate transponders to fall back to lower transmission rates depending on the SLA and reduce energy consumption. This is related to the next section, where we discuss quality of protection (QoP) differentiation. Moreover, there are open issues which must be addressed:

1 Are complex devices more prone to failures?
2 How quickly can the devices be brought up and down to react to traffic variations?
3 How can the energy consumption information of the device and electricity cost be accurately monitored?

Trade-offs in Dynamic Adaptation to Temporal Variation of Traffic — Traffic in core networks is usually higher during the day and lower during the night. Putting idle devices into sleep or energy-saving mode (e.g., adaptation of transmission rates) can effectively reduce the energy consumption of the backup resources (e.g., adapting the backup transmission in DP 1+1 can save up to 22 percent of energy consumption at the optical layer during the low-traffic hours of the weekend [3]). Another strategy relies on concentrating backup paths into separate fibers in order to be able to put the devices on these links into sleep mode without being constrained by the presence of working paths. This is applied to DP [6] and SP [7] with power savings of up to 35 percent with respect to non-sleep-mode scenar-

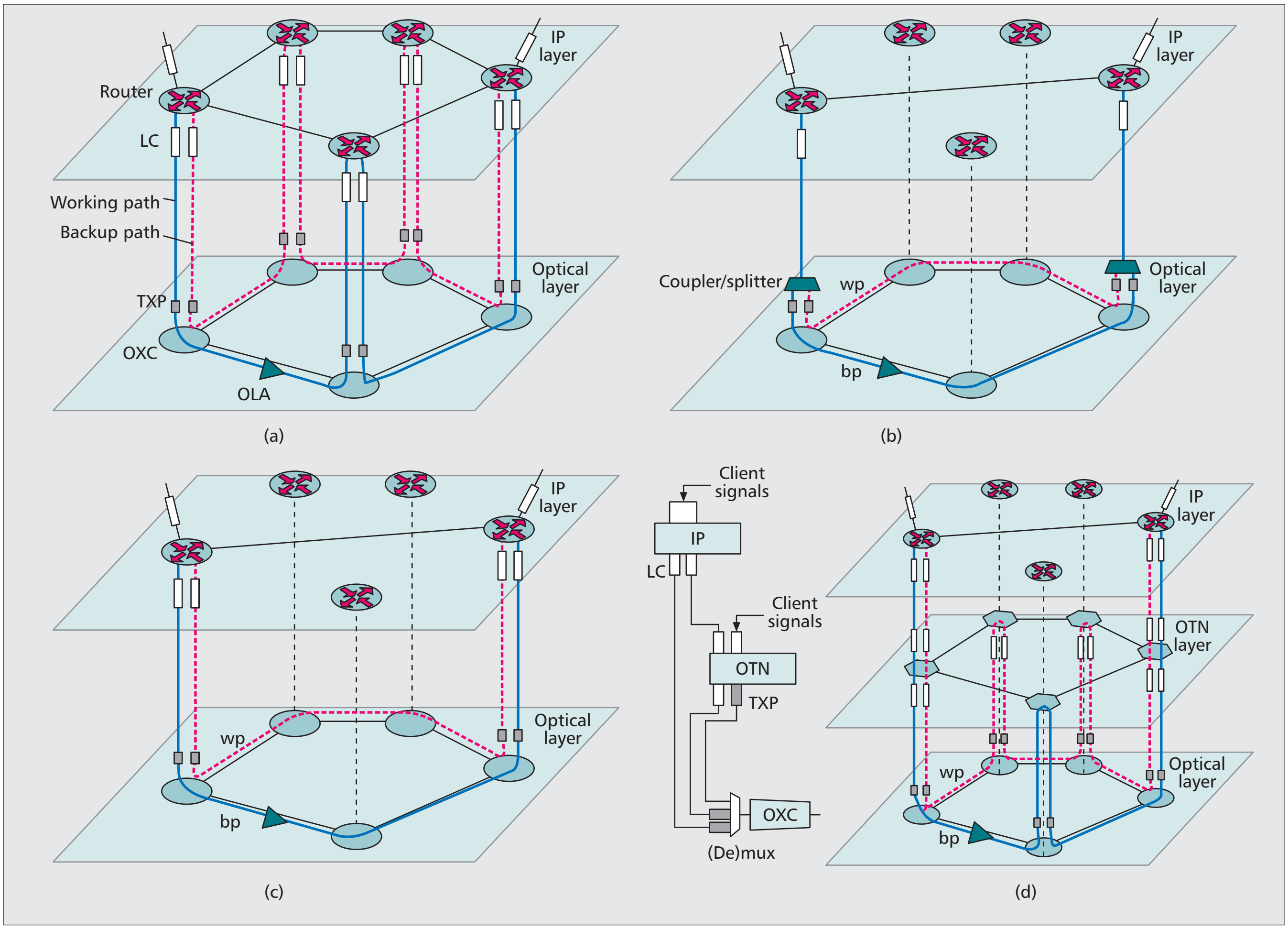


Figure 1. a) 1+1 protection at the IP layer; b) 1+1 protection at the optical layer; c) employing optical bypass at the IP layer; d) grooming at the OTN layer. LC: line card, TXP: transponder, OXC: optical cross connect, OLA: optical line amplifier, wp: working path, bp: backup path.

ios. Moreover, the sleep mode functionality can also be exploited considering grooming at the electrical layer with energy savings of up to 60 percent in both the optical and IP layers [8]. Nevertheless, for these mechanisms, it will be important to determine the best trade-off between resilience, energy efficiency, and low reconfiguration costs.

Geographical Traffic Distribution — Geographical traffic distribution in networks covering large areas may be diverse, especially in networks spanning several time zones. Is it possible to use idle resources from one time zone as protection resources in another time zone? Traffic distribution can also suffer strong variations due to external influences like natural disasters or wide audience events. Extra redundancy can be provided to face the potential traffic changes, but it will be costly. Therefore, how to handle this potential variation from the protection point of view is an open question. There are some works that differentiate traffic in different time zones [5] or use traffic data originating from measurements [9]. However, to the best of our knowledge, there is no related work focused on geographical traffic distribution and green networking.

Granularity of Traffic Demands — The size of a single traffic demand determines the possibility of grooming it into a set of lightpaths and links to utilize the available network resources in an efficient way. However, grooming of multiple traffic demands over a limited number of network resources poses a challenge from the resilience point of view, since some traffic demands need to traverse multiple hops from their origin to their target. This, in turn, increases the number of devices on their path, each being prone to failures. This leads to important questions related to the maximum granularity of traffic demands (with respect to network resources) so that traffic grooming for energy saving pays off without violation of protection constraints. While an initial exploration on this issue [9, Fig. 7] indicates that grooming can be more energy-efficient when the granularity of the traffic demands is below half of the line rate, the cited work does not look at the impact of grooming on availability (in terms of relative uptime per year).

Strategies to Meet the SLA Requirements

Physical Impairments — Network design techniques have to consider physical impairments (e.g., fiber loss, dispersion, or nonlinear effects)

Preemption is based on the intuition that low-priority services can be provisioned with the option to be discarded should a connection with a higher criticality level need to use the network resources. This could happen after a failure concatenation, or some unexpected or very infrequent event.

and their impact on the resulting quality of transmission (QoT). Some energy saving techniques can exacerbate the effect of some physical impairments (e.g., cross-phase modulation and crosstalk) [10], as they tend to concentrate most of the traffic on a few links to put unused devices into sleep mode. This problem can be overcome by introducing design techniques that are both energy- and impairment-aware. The work [10] shows that such techniques are able to achieve the same energy savings of conventional (i.e., non-impairment-aware) green strategies while providing QoT levels close to those provided by the impairment-aware design strategies (i.e., only a small percentage of the lightpaths cannot be established due to insufficient QoT). Even though the work in [10] considers an unprotected scenario, we can expect that this aspect will be even more critical when designing protected networks, where the backup paths are on average longer than their respective working paths. Power transient phenomena will also have to be considered when setting network elements into sleep mode and adequate mechanisms (e.g., disabling the transponder electronics while keeping the injected optical power) must be provided. Moreover, the continuous development of high-speed systems comes at the cost of reduced reach. Consequently, a higher number of O/E/O regenerations (highly energy consuming processes) would be needed for a generalized deployment. Employing more robust modulation formats for long backup paths can be an alternative to regeneration, but at the expense of requiring higher energy per bit [3].

Differentiated Quality of Protection — Commonly, a single policy is applied to all aggregated traffic demands as shown in Fig. 1. An alternative to this policy would be to apply differentiated QoP by assigning different resilience levels to demand "subsets" with different SLA needs [11]. These subsets could be allocated, for instance, to different residential services or to different performance classes of a corporate service (e.g., optical virtual private networks, VPNs). The application of differentiated QoP allows for a reduction of protection resources, which may result in improved energy efficiency with respect to the conventional DP 1+1 scheme (energy efficiency per gigahertz, EEPG, improvements of up to 300 percent are reported in [11]). Another possibility is to adapt the reliability performance of a given protection scheme to the reliability requirements of the provisioned traffic, that is, the differentiated reliability (DiR) concept. Thus, if for some traffic demands the backup path does not always need to be available for any possible failure scenario, it will be possible to selectively assign protection resources only to those demands that need them the most. This approach can lead to significant energy savings (i.e., up to 25 percent [12]) with respect to conventional protection strategies that are always 100 percent survivable.

EON/BVT — EON technologies provide an extra degree of freedom in assigning traffic demands to QoP levels. For example, a sliceable bandwidth variable transponder (BVT) based on a multicarrier approach could be used to assign a number of carriers to a specific QoP. This can allow the optimizing of resource utilization and providing a finer granularity on QoP mapping. Besides, a BVT can adapt the transmission rate (modifying the bandwidth and/or modulation format) to the traffic variations, and thus reduce energy consumption. EON is shown to provide much higher EEPG than traditional fixed-grid WDM technologies with any protection scheme (e.g., EEPG can be improved up to 200 percent with respect to MLR with SP [3]).

Network Virtualization — Network resources can be partitioned over different operators and/or services. In that respect, the physical network would comprise a number of virtual networks, each receiving a different degree of resilience according to the amount of resources allocated to each service/operator. The virtualization of the network components combined with the creation of segregated virtual topologies with different architectures (including different routing and resilient protocols) can help support variable QoP levels and thus reduce energy consumption by sharing the resources among different users/virtual networks. The differentiation can be achieved in one or more dimensions including time delay, availability, protection scheme/protection level, QoT, and the network layer selected (i.e., electrical and/or optical). Despite numerous works that consider realization of a virtual network over a physical one, only a few of them consider network virtualization and protection. For example, 28 percent of power can be saved by allowing virtual links to bypass physical nodes in an IP-over-WDM network (multihop bypass compared to non-bypass approach for NSFNET network at 6:00 pm [5, Fig. 5], providing protection in the physical layer by keeping minimum nodal degree equal to 2. However, protection is not the focus of [5].

Data Preemption — Preemption is based on the intuition that low-priority services can be provisioned with the option to be discarded should a connection with a higher criticality level need to use the network resources. This could happen after a failure concatenation, or some unexpected or very infrequent event. In this case, lightpath preemption levels are used to decide which signal should be allocated first, releasing the resources in use by lightpaths with lower preemption levels. We can consider this mechanism as a subset of a QoP scheme. This, of course, has to be done without compromising the required QoS level of each connection. Preemption-based strategies can improve energy efficiency, since powering-on extra resources is not always necessary. For instance, in a scenario where preemption is used to dynamically provision the service subject to different QoP requirements, redundant resources used for protection could be used to convey extra traffic if the carried traffic is able to tolerate a disruption, should one or more failures occur in the network. This will be accomplished without unnecessarily turning on extra resources in the network, with evident benefits in terms of energy efficiency. How effective the use of such energy-efficient preemption techniques is still an open question, which, to the best of our knowledge, has not been studied so far. The answer will depend on a series of factors such as the protection mechanism used and the traffic composition, just to name a couple. Higher power reduction can be expected in the presence of traffic with highly diverse survivability needs, as well as with protection techniques that are less stringent in their requirements. Finally, there is one last aspect to consider. If backup resources are used to route preemptable traffic, they need to stay in an active state, potentially nullifying the benefits of some energy-efficient protection techniques. There is obviously a trade-off that needs to be assessed here to understand under which conditions preemption will be beneficial from an energy reduction point of view.

Metrics and Standards

This section considers how issues related to energy efficiency and resilience can be evaluated and what this would mean in terms of standardisation requirements.

Metrics — The energy efficiency of a network with various resilience schemes can be assessed with common metrics such as Watts per bit per second or Joules per bit, or the inverse of both. With this approach, less protective schemes (e.g., SP) will result in more energy-efficient designs than more protective schemes (e.g., DP 1+1). While this metric is a good indicator of the actual power or energy required to transport a bit of information, it does not take into account the level of protection. To do so, it might make sense to express the energy efficiency normalized by the protection factor, where the protection factor could be 2 for 1+1 DP, 1.5 for SP, and so on. This metric provides a fairer indication of the energy required for a given level of protection. Finally, with the potential evolution toward more flexi-grid equipment, the common Watts per bit per second metric (or its derivatives) might be insufficient to capture the efficiency in utilizing the available spectrum. For example, a BVT can make use of adaptive bandwidths depending on the required transmission rate and distance. Therefore, it might make sense to take into account the spectral efficiency of transmission with an energy efficiency per gigahertz metric, as proposed in [3].

Standards — Putting unused network devices into sleep mode is one way to save power. However, current Internet protocols operate based on the assumption that network elements are always on. The application and service reinitializations when these devices wake up again would potentially result in a non-negligible amount of signaling overhead [13]. Modifications to the control plane considering link or node removal should provide the ability to choose the level of redundancy available after the network topology has been trimmed. The complete removal of nodes or links from the network topology has several impacts on the control plane that must be considered [14]. For example, it is essential to modify the network topology so that the removed links or devices are not used to forward traffic remembering that such links exist, possibly including the neighbors and destinations reachable through those links or devices. One solution to this sleep mode problem could be based on the use of a proxy [13]. Before going into sleep mode, a node delegates its functionalities to a proxy, which will then respond to routine network traffic on behalf of the sleeping node and will wake the node up when needed. The protocols and procedures for proxy operation such as discovery, selection, delegation, and wake-up have to be defined. Another example would be to require that nodes can negotiate timeouts (in protocols that make use of timeouts), so a node might be able to go into sleep mode or attempt to synchronize periodic messages across a number of protocols. Thus, all these messages fall into a certain timeframe, and in between the node can sleep. The issues described above can also be addressed by designing sleep-compatible protocols or extending existing protocols (where possible) to include the ability to distinguish sleeping elements from failed ones. Some extensions required in existing generalized multiprotocol label switching (GMPLS), Open Shortest Path First (OSPF) routing, Resource Reservation Protocol (RSVP) signaling, and Link Management Protocol (LMP) are proposed in [15] to support energy-efficient traffic engineering technology.

If backup resources are used to route preemptable traffic, they need to stay in an active state, potentially nullifying the benefits of some energy-efficient protection techniques. There is obviously a trade-off that needs to be assessed.

Conclusion

Both energy efficiency and resilience in telecommunications networks are well established topics in the research community. However, the combination of both for energy-efficient resilient optical network design is still a relatively new research field. In this article, we identify the corresponding major research challenges and performance trade-offs for core networks from three different aspects: long-term traffic predictions (including network architecture, embodied energy, and resilience at different layers and topologies), short-term traffic dynamics (including novel equipment features, trade-offs in dynamic adaption to temporal variation of traffic, geographical traffic distribution, and granularity of traffic demands), and SLA requirements (including physical impairments, differentiated QoP, EON/BVT, network virtualization, and data preemption). All these factors need to be considered as they influence not only the design and the choice of resilience mechanisms, but also the energy efficiency performance.

New metrics need to be used for energy efficiency resilience evaluation by either considering the energy efficiency with a protection factor or the utilized optical spectrum. From a standardization point of view, existing protocols need to be extended for energy-efficient resilient optical networks using sleep mode devices.

Acknowledgments

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 257740 (TREND) and grant agreement no. 318137 (DISCUS), and

New metrics need to be used for energy-efficiency resiliency evaluation by either considering the energy-efficiency with a protection factor or the utilized optical spectrum. From a standardisation point of view, existing protocols need to be extended for energy-efficient resilient optical networks using sleep mode devices.

from the Polish National Science Center (decision DEC-2014/12/S/ST7/00415).

References

[1] E. Le Rouzic, "Network Evolution and the Impact in Core Networks," *Proc. ECOC*, Turin, Italy, Sept. 2010.

[2] S. Lambert *et al.*, "Worldwide Electricity Consumption of Communication Networks," *Optics Express*, vol. 20, no. 26, Dec. 2012, pp. B513–24.

[3] J. López Vizcaíno *et al.*, "Protection in Optical Transport Networks with Fixed and Flexible Grid: Cost and Energy Efficiency Evaluation," *Opt. Switching and Networking*, vol. 11, part A, Jan. 2014, pp. 55–71.

[4] X. Dong *et al.*, "Energy-Efficient Core Networks," *Proc. ONDM*, Colchester, U.K., Apr. 2012.

[5] X. Dong *et al.*, "On the Energy Efficiency of Physical Topology Design for IP Over WDM Networks," *IEEE/OSA J. Lightwave Technology*, vol. 30, no. 12, June 2012, pp. 1931–42.

[6] A. Jirattigalachote *et al.*, "Dynamic Provisioning Strategies for Energy Efficient WDM Networks with Dedicated Path Protection," *Optical Switching and Networking*, vol. 8, no. 3, July 2011, pp. 201–13.

[7] A. Coiro *et al.*, "Reducing Power Consumption in Wavelength Routed Networks by Selective Switch Off of Optical Links," *IEEE J. Selected Topics in Quantum Electronics*, vol. 17, no. 2, Mar. 2011, pp. 428–36.

[8] F. Musumeci *et al.*, "Energy-Efficiency of Protected IP-over-WDM Networks with Sleep-Mode Devices," *J. High Speed Networks*, vol. 19, no. 1, 2013, pp. 19–32.

[9] W. Van Heddeghem *et al.*, "A Power Consumption Sensitivity of Circuit-Switched versus Packet-Switched Backbone Networks," *Computer Networks*, 2014; http://dx.doi.org/10.1016/j.comnet.2014.09.015.

[10] C. Cavdar *et al.*, "Design of Green Optical Networks with Signal Quality Guarantee," *Proc. IEEE ICC*, Ottawa, Canada, June 2012.

[11] J. López *et al.*, "Energy- and Cost-Efficient Protection in Core Networks by a Differentiated Quality of Protection Scheme," *Proc. ECOC*, London, U.K., Sept. 2013.

[12] A. Muhammad *et al.*, "Reliability Differentiation in Energy Efficient Optical Networks with Shared Path Protection," *Proc. IEEE GreenComm*, Oct. 2013.

[13] R. Winter, S. Jeong, and J. H. Choi, "Towards an Energy-Efficient Internet," http://tools.ietf.org/id/draft-winter-energy-efficient-internet-01.txt, Oct. 2012.

[14] A. Retana, R. White, and M. Paul, "A Framework and Requirements for Energy Aware Control Planes," http://www.ietf.org/id/draft-retana-rtgwg-eacp-00.txt, July 2012.

[15] S. Okamoto, "Requirements of GMPLS Extensions for Energy Efficient Traffic Engineering," http://tools.ietf.org/id/draft-okamoto-ccamp-midori-gmpls-extension-reqs-01.txt, Feb. 2012.

Biographies

YABIN YE (yeyabin@huawei.com) received his B.E. and Ph.D. degrees from Tsinghua University, Beijing, China, in 1997 and 2002, respectively. From 2002 to 2004 he was a research scientist with the Institute for Infocomm Research (I2R), Singapore. From 2004 to 2008 he was a senior researcher with Create-Net, Italy. Currently, he is a senior researcher with Huawei Technologies European Research Center, Munich, Germany. His research interests include optical networking, high-speed transmission technologies, and hybrid optical/wireless access technologies.

FELIPE JIMÉNEZ ARRIBAS (felipej@tid.es) holds an electronic physics degree from UCM. He joined Telefónica I+D in 1994 on a Ph.D. student scholarship, becoming part of the staff in 1995. He started working on the development of the Telefónica Spain nationwide IP network, and since then has been involved in many projects related to access and core architectures as well as in field trials of advanced fiber technologies. He has also worked on relevant EU projects, like the FP6 MUSE, FP7 Accordance, TREND, Discus, and Idealist projects.

JAAFAR ELMIRGHANI [F] (J.M.H.Elmirghani@leeds.ac.uk), professor at University of Leeds, is a Fellow of the IET and Fellow of the Institute of Physics. He received the IEEE Comsoc 2005 Hal Sobol award, the IEEE Comsoc Signal Processing and Communication Electronics outstanding service award, 2009, and an IEEE ICC 2013 Best Paper award. He has been awarded £22 million in grants, published over 400 papers, and has research interests in communication networks. He is an IEEE Comsoc Distinguished Lecturer.

FILIP IDZIKOWSKI (filip.idzikowski@put.poznan.pl) is an assistant professor at Poznan University of Technology (PUT), Poland. He received his Ph.D. (Dr.-Ing.) from Technical University of Berlin (TUB), Germany, in 2014. His main research interests include power consumption of multi-layer core networks, traffic modeling, and routing. Part of this work was carried out when he was with TUB and Sapienza University of Rome, Italy.

JORGE LÓPEZ VIZCAÍNO (jorge.vizcaino@huawei.com) received his Bachelor's degree in telecommunications engineering from the Technical University of Madrid in 2008 and his M.Sc. degree in telecommunications from the Technical University of Denmark in 2011. Since 2012, he has been a researcher at the European Research Center of Huawei Technologies in Munich, Germany, and a Ph.D. candidate at the Technical University of Dortmund, Germany. He has been involved in the EU project TREND related to energy efficiency. His research interests are in network planning, network protocols, energy efficiency, and optical networking.

PAOLO MONTI (pmonti@kth.se) received a Ph.D. (2005) from the University of Texas at Dallas and is now an assistant professor at KTH Royal Institute of Technology. His main research interests are within the networking aspects of all-optical networks, while part of his work has recently focused on the energy efficiency of telecommunication networks with particular emphasis on current and future converged (wired + wireless) broadband access solutions.

FRANCESCO MUSUMECI (francesco.musumeci@polimi.it) received his Ph.D. degree in information engineering from Politecnico di Milano, Italy, in 2013. He is currently a research assistant at the Department of Electronics, Information and Bioengineering of Politecnico di Milano. He is an author of more than 20 papers published in relevant international journals and conference proceedings in the field of optical communication networks. His research interests focus on energy-saving switching and networking, optical core network architectures, content delivery networks, and network resilience.

ACHILLE PATTAVINA (achille.pattavina@polimi.it) received his Dr. Eng. degree in electronic engineering in 1977 and since 1995 has been a full professor at Politecnico di Milano," (Italy). He has been an author of more than 200 papers and two books in the area of communications networks. He has been a coordinator of national and international research activities, including European Union funded projects. His current research interests are in the areas of green ICT, optical networks, switching theory, traffic modeling, and broadband convergent access/metro networks.

WARD VAN HEDDEGHEM (ward.vanheddeghem@intec.ugent.be) received his Master's degree in applied engineering electromechanics from Hogeschool Gent, Belgium (1999), and his M.Sc. degree in computer science engineering from Vrije Universiteit Brussel, Belgium (2009). In 2014 he obtained a Ph.D. degree at the IBCN research group at INTEC UGent/iMinds. His research interests are in the field of environmental impact of ICT and energy-efficient network architectures.

Call for Papers
IEEE Communications Magazine
Communications Standards Supplement

Background

Communications standards enable the global marketplace to offer interoperable products and services at affordable cost. Standards development organizations (SDOs) bring together stakeholders to develop consensus standards for use by a global industry. The importance of standards to the work and careers of communications practitioners has motivated the creation of a new publication on standards that meets the needs of a broad range of individuals including: industrial researchers, industry practitioners, business entrepreneurs, marketing managers, compliance/interoperability specialists, social scientists, regulators, intellectual property managers, and end users. This new publication will be incubated as a Communications Standards Supplement in *IEEE Communications Magazine*, which, if successful, will transition into a full-fledged new magazine. It is a platform for presenting and discussing standards-related topics in the areas of communications, networking, and related disciplines. Contributions are also encouraged from relevant disciplines of computer science, information systems, management, business studies, social sciences, economics, engineering, political science, public policy, sociology, and human factors/usability.

Scope of Contributions

Submissions are solicited on topics related to the areas of communications and networking standards and standardization research, in at least the following topical areas:

Analysis of new topic areas for standardization, either enhancements to existing standards, or of a new area. The standards activity may be just starting or nearing completion. For example, current topics of interest include:

- 5G radio access
- Wireless LAN
- SDN
- Ethernet
- Media codecs
- Cloud computing

Tutorials on, analysis of, and comparisons of IEEE and non-IEEE standards. For example, possible topics of interest include:

- Optical transport
- Radio access
- Power line carrier

The relationship between innovation and standardization, including, but not limited to:

- Patent policies, intellectual property rights, and antitrust law
- Examples and case studies of different kinds of innovation processes, analytical models of innovation, and new innovation methods

Technology governance aspects of standards focusing on both the socio-economic impact as well as the policies that guide it. This would include, but are not limited to:

- The national, regional, and global impacts of standards on industry, society, and economies
- The processes and organizations for creation and diffusion of standards, including the roles of organizations such as IEEE and IEEE-SA
- National and international policies and regulation for standards
- Standards and developing countries

The history of standardization, including, but not limited to:

- The cultures of different SDOs
- Standards education and its impact
- Corporate standards strategies
- The impact of Open Source on standards
- The impact of technology development and convergence on standards

Research-to-Standards, including standards-oriented research, standards-related research, research on standards

Compatibility and interoperability, including testing methodologies and certification to standards

Tools and services related to any or all aspects of the standardization lifecycle

Proposals are also solicited for Feature Topic issues of the Communications Standards Supplement.

Articles should be submitted to the *IEEE Communications Magazine* submissions site at

http://mc.manuscriptcentral.com/commag-ieee

Select “Standards Supplement” from the dropdown menu of submission options.

Next Generation Elastic Optical Networks: The Vision of the European Research Project IDEALIST

Antonio Napoli, Marc Bohn, Danish Rafique, Alexandros Stavdas, Nicola Sambo, Luca Potì, Markus Nölle, Johannes Karl Fischer, Emilio Riccardi, Annachiara Pagano, Andrea Di Giglio, Michela Svaluto Moreolo, Josep Maria Fabrega, Emilio Hugues-Salas, Georgios Zervas, Dimitra Simeonidou, Patricia Layec, Antonio D'Errico, Talha Rahman, and Juan Pedro Fernández-Palacios Giménez

Antonio Napoli, Marc Bohn, and Danish Rafique are with Coriant R&D GmbH.

Alexandros Stavdas is with University of Peloponnese.

Nicola Sambo and Luca Potì are with CNIT.

Markus Nölle and Johannes Karl Fischer are with Fraunhofer Institute for Telecommunications Heinrich Hertz Institute.

Emilio Riccardi, Annachiara Pagano, and Andrea Di Giglio, are with Telecom Italia.

Michela Svaluto Moreolo and Josep Maria Fabrega are with Centre Tecnològic de Telecomunicacions de Catalunya (CTTC).

Emilio Hugues-Salas, Georgios Zervas, and Dimitra Simeonidou are with University of Bristol.

Patricia Layec is with Alcatel-Lucent Bell Labs France.

Antonio D'Errico is with Ericsson Telecomunicazioni SpA.

Talha Rahman is with Eindhoven University of Technology.

Juan Pedro Fernández-Palacios Giménez is with Telefonica Innovación y Desarrollo.

Abstract

In this work we detail the strategies adopted in the European research project IDEALIST to overcome the predicted data plane capacity crunch in optical networks. In order for core and metropolitan telecommunication systems to be able to catch up with Internet traffic, which keeps growing exponentially, we exploit the elastic optical networks paradigm for its astounding characteristics: flexible bandwidth allocation and reach tailoring through adaptive line rate, modulation formats, and spectral efficiency. We emphasize the novelties stemming from the flexgrid concept and report on the corresponding proposed target network scenarios. Fundamental building blocks, like the bandwidth-variable transponder and complementary node architectures ushering those systems, are detailed focusing on physical layer, monitoring aspects, and node architecture design.

Introduction

In the last decade we have witnessed significantly increased utilization of web-based applications such as streaming, sharing, and downloading of multimedia contents that are saturating the capacity of transnational backbones. Globally, consumer mobile devices and network connections grew to 7 billion in 2013, up from 6.5 billion in 2012. Simultaneously, the Internet of Things is envisioned to be practical in the short term. Moreover, the dynamic nature of Internet traffic sets further requirements on the flexibility of next generation optical communication systems.

Based on data obtained from backbone reference networks, operators within the Industry-Driven Elastic and Adaptive Lambda Infrastructure for Service and Transport Networks (IDEALIST) project estimate a compound annual growth rate (CAGR) of Internet traffic around 35 percent; therefore, a substantial redesign of current networks is required. At the moment, three strategies are possible to postpone the imminent capacity crunch:

1. The development of ultra-low-loss and low-nonlinear optical fibers, in conjunction with hybrid amplification schemes (i.e., lumped and distributed)
2. The utilization of spatial division multiplexing transmission
3. The implementation of elastic optical networks (EONs)

All combinations of the three options are possible. However, the two first alternatives necessitate massive infrastructure rollout, leading to substantial network cost overheads. Conversely, IDEALIST expressly focuses on short- to medium-term industrial outcomes; consequently, we have opted for the latter alternative.

The realization of EONs [1] is considered a realistic perspective because:

1. It can be seamlessly deployed.
2. It optimizes the required optical spectrum.
3. It provides flexibility with minimum disruption to the existing infrastructures since they can support the innovative concepts of sliceable bandwidth-variable transponder (S-BVT) and flexible optical cross-connect (Flex-OXC) at competitive costs.

The concept of bandwidth on demand stems from mobile communications and recently has been proposed for optics to efficiently utilize the available spectrum, thus increasing the spectral efficiency (SE), defined as the channel information rate over the utilized frequency slot. Within optical networks, flexibility is obtained by redesigning and adding features to the aforementioned blocks. The S-BVT, for example, must be able to transparently transmit signals with selectable modulation formats at adaptive rates [2] and allocating the capacity into one or several independent optical flows that are transmitted toward one or multiple destinations [3]. The Flex-OXC implements dynamic optical functions in a modular and flexible manner by employing the architecture-on-

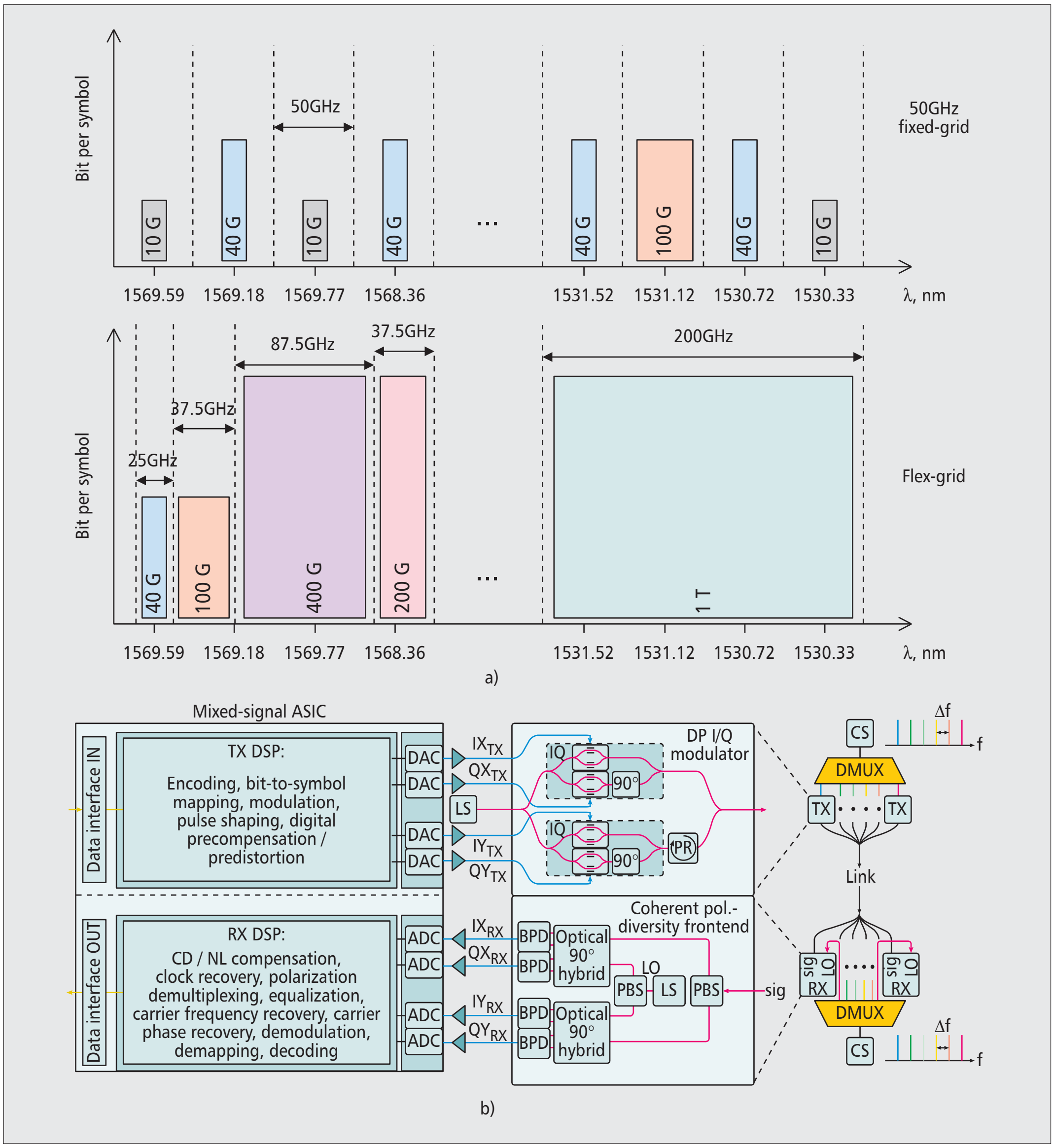


Figure 1. a) Fixed (50-GHz granularity) vs. Flex-grid optical systems (12.5-GHz granularity); b) block diagram of a Flex subcarrier module. ASIC: application-specific integrated circuit (ASIC); TX: transmitter; DP: dual-polarization; DMUX: demultiplexer; RX: receiver; LS: light amplication by stimulated emission of radiation (LASER) source; LO: local oscillator; BPD: balanced photo-detectors; CS: comb source; NL: nonlinearity.

demand (AoD) paradigm [4] and utilizing the switchless elastic rate node (SERANO) [5] as one of the key functional elements for signal processing and routing. These innovations must be delivered by simultaneously decreasing the transmission costs and reducing the energy consumption, paying particular attention to the operation, administration, and maintenance (OAM) issues to enable flexibility and awareness of the dynamical traffic during the whole migration process.

The article is structured as follows. We introduce the migration from state-of-the-art systems toward EONs. Next, we describe the most relevant characteristics of the S-BVT and Flex-OXC. We compare EONs and commercial fixed-grid networks, providing the motivations for the envisioned migration. Finally, the last section draws the conclusions.

From Fixed-Grid-Based Optical Networks toward EONs

In the category of emerging telecommunication technologies, EONs have gained momentum in relation to optical transport networks (OTNs). The EON architecture efficiently utilizes the optical spectrum through a network infrastructure that implements flexible channel allocation using frequency slots with reduced sizes (e.g., from current fixed-grid 50 GHz wavelength-division multiplexing, WDM, systems to configurable slots of 12.5 GHz granularity). In addition, EONs present the benefit of providing customized spectral grids whenever new lightpaths are established. The allocation of several channels to form super-channel configurations is performed according to user requests in a highly spectrum-efficient and scalable manner

EONs present the benefit of providing customized spectral grids whenever new lightpaths are established. The allocation of several channels to form super-channel configurations is performed according to user requests in a highly spectrum-efficient and scalable manner.

Feature	2018	Goal (2025)
Maximum optical switching capacity	60 Tb/s	500 Tb/s
Maximum electrical switching capacity	18 Tb/s	150 Tb/s
Type and number of client interface (typical node)	10G-100G-1T, IEEE 803.1/3, number: 150	100G-1T, IEEE 803.1/3, number: 150
Number of fiber pairs attached to the node (largest node)	About 4–6 where ND = 6	About 30 fibers pairs, ≫ topological ND (maximum ND = 6). This implies several parallel WDM links.
Minimum super-channel rate	10 Gb/s	100 Gb/s
Maximum super-channel rate	1 Tb/s	10 Tb/s
Typical super-channel rate	100 Gb/s	1 Tb/s
Minimum transparency reach	700 km	700 km

Table 1. Node and network target requirements.

[1]. Figure 1a illustrates an example of allocation of different data rates from 40 Gb/s up to 1 Tb/s for both fixed- and flex-grid.

With this in mind, the International Telecommunication Union Telecommunication Standardization Sector (ITU-T) adapted recommendations G.694.1 and G.872 to include flexibility within ITU standards. A novel WDM concept was defined by ITU-T Study Group 15, starting with Recommendation G.694.1, with the formalization of the nominal central frequencies (6.25 GHz granularity), the channel width (multiples of 12.5 GHz), and the concept of frequency slots. In such a scheme, a data plane connection is switched, based on allocated and variable size frequency ranges within the optical spectrum (media channels), providing a first brick for EONs. This technology can be exploited by multicarrier WDM transmission through a super-channel approach as well as by optical-orthogonal frequency-division multiplexing (O-OFDM). Once the ITU grid has been adapted to the flexible scheme, current transponders and reconfigurable optical add-drop multiplexers (ROADMs) will follow the same path. The transponder has to move from fixed-grid architecture toward flexible transmission with the possibility of adjusting the spectrum by expanding or contracting the bandwidth on demand. This is possible by varying the number of sub-carriers, the symbol rate, or the employed modulation format based on a dynamic trade-off between reach and capacity. However, when BVTs need to transmit at low bit rates, part of its capacity remains unused. To address this issue, the concept of S-BVT was introduced [3], which further increases the level of elasticity and efficiency inside the network. S-BVTs enable transmission from one point to multiple destinations, changing the traffic rate to each destination and the number of destinations on demand (e.g., multi-flow transmission). Concerning Flex-OXC, the latest spectral selective switch (SSS) technology enables channel routing with the aforementioned finer granularity, and the embedded flexibility in the node architecture [4] helps follow dynamic traffic variation.

Finally, the two aforementioned main blocks of the EON data plane architecture proposed in IDEALIST are supposed to be fully programmable so that they can be dynamically controlled by external software agents implementing online service network optimization mechanisms. In that respect, IDEALIST follows a software defined network (SDN) approach based on the application based network operation (ABNO) architecture proposed in [6].

Target Networks

Within IDEALIST the target reference application scenario is a backbone transport network with a CAGR of ~ 35 percent and target reach of ~ 1500 km (the typical size of a large national European network). System design parameters are estimated on deployed topologies, realistic traffic matrices, and evolution predicted until 2018 when the first rollout should begin. Scalability, flexibility, end-to-end performance, low cost, and limited energy consumption are the most critical issues to be addressed. A possible solution is to integrate OTN or IP with the optical layer to share flexibility of the network, thus increasing node agility and scalability. Concerning energy consumption and costs, photonic integration jointly with S-BVTs represents a key technology to simultaneously reduce both cost and energy consumption.

From a functional point of view, IDEALIST identifies two types of network nodes: border and core. The first is suitable for metro regional traffic grooming, while the second is used in long-haul networks for direction switching. Both nodes need enhanced modularity and programmability to enable smooth upgrades and variable traffic conditions. It is generally agreed that the differences in terms of size among nodes located in different parts of the same network will probably not imply the complete redesign of the architecture. Border and core nodes are essentially different flavors of the same structure, varying only in size and equipment. Nevertheless, a modular design is envisaged, following the *pay-as-you-grow* approach. In particular, node design requirements should take into account that a node has a 10-year lifetime, and its maximum capacity should refer to the estimated traffic conditions at the end of life (a factor of 10 from 2018 is assumed with a yearly traffic growth of ~35 percent). In Table 1 we summarize the most relevant specifications for the identified target network elements. These values are extrapolated from a set of backbone reference networks, provided by the operators within our project, with the aim of roughly identifying the size and needs of a future network element. The maximum size of the optical switching fabric by 2025 is estimated to be around 500 Tb/s, scalable in a hitless way from 25 Tb/s minimum capacity. Similar figures apply to the electrical crossconnect with a maximum size of 150 Tb/s derived from a typical 30 percent add/drop capacity.

Operators' forecasts identify 100 Gb/s (optionally 40 Gb/s) as the most common bit rate for client interfaces. 10 Gb/s should still be considered for 2018 deployment. The add/drop capacity, divided by the *most common bit rate interface*, roughly defines the number of client interfaces to be equipped in a single node. The maximum topological nodal degree (ND) provides the information about the maximum number of directions (at the line side) exiting from a site. This information alone does not identify the correct number of WDM systems linked to that particular network element, since parallel WDM systems might sometimes be required. Nevertheless, it is still possible to derive the number of fiber pairs directly connected to the network element. By 2025 we conjecture that 30 fiber pairs over C-band could be a reasonable value.

The super-channel bandwidth can be calculated using the minimum, typical, and maximum traffic. Considering that the figures of *minimum traffic per demand* range from 2 to 40 Gb/s by 2018 and the minimum frequency slot is 12.5 GHz as reported earlier, it is realistic to specify a minimum super-channel rate of 10 Gb/s. The maximum super channel rate will be 1 Tb/s by 2018 and possibly up to 10 Tb/s by 2025.

The Sliceable Bandwidth-Variable Transponder

From the optical perspective, the key component inside the S-BVT is the optical front-end, which is called the multi-flow optical module. The module distributes different traffic demands over several media channels, which are then grouped into super-channels. It contains a set of Flex subcarrier modules (where each subcarrier is modulated by a specific traffic portion) and a number of subcarrier generation modules, each generating non-modulated subcarriers.

The most important elements within the multi-flow optical module based on spectrally efficient super-channels are: flexible comb sources, coherent receiver, rate-adaptive coded modulation, and modulation format transparent digital signal processing (DSP). At the transmitter, a flexible comb source, as the one described later, generates an optical frequency comb with a variable carrier spacing Δf. After demultiplexing, each optical carrier is individually modulated by a Flex subcarrier module. The modulated wavelength channels are then combined and transmitted over a link. The local oscillators for the coherent receivers are also derived from a comb source, creating a scenario in which joint DSP of all subcarriers (e.g., for intra super-channel fiber nonlinearity compensation) is possible.

General Concept

Figure 1b illustrates the block diagram of a possible Flex subcarrier module realization, which comprises a pool of coherent front-ends (light sources, I/Q modulators, and drivers) associated with the processing electronics (digital-to-analog converter [DAC], analog-to-digital converter [ADC], transmitter, receiver, DSP at the transmitter and receiver, and forward error correction [FEC] coding and decoding). The transmitter DSP consists of encoding and mapping blocks supporting various modulation formats and optionally variable FEC code rates for rate-adaptive coded modulation. An optional precompensation stage at the transmitter allows for chromatic dispersion (CD) and/or component limitations and/or fiber nonlinearity precompensation [2, 7]. The DSP-based receiver digitally corrects the optical front-end characteristics, CD and/or fiber nonlinearity compensation, clock and carrier recovery, as well as polarization demultiplexing and linear equalization [2, 7]. The requested bandwidth can be tuned by changing the modulation format, symbol rate, spectral shaping, and coding. For cost-efficient realization of Flex subcarrier modules, modulation format transparent solutions for all mentioned functions are required [2]. Ideally, Flex subcarrier module functionalities, and more generally S-BVT parameters, can be controlled, for example, by SDN that has access to the physical layer parameters and can adjust S-BVT configurations, as discussed later.

Subcarrier Generation Module

The comb generator is a fundamental block of the multi-flow optical module. Two solutions can be adopted for subcarrier generation: an array of LASER sources or a multi-wavelength source able to generate several subcarriers from a single LASER. In the first case particular attention has to be paid to frequency locking and stabilization, because they are independently generated, and as a consequence, the modulated spectra may overlap.

On the contrary, a frequency comb generator produces a spectrum that consists of a series of equally spaced frequency- and phase-locked sharp lines. In order to generate a comb signal, one may employ super-continuum LASER source techniques, such as mode-locked LASERs (MLLs). Based on MLLs, simultaneous tones are generated due to the LASER emission of a train of periodic ultra-short optical pulses establishing a fixed phase relationship across a broad spectrum of frequencies. The periodic pulses naturally generate a comb of discrete, regularly spaced series of sharp lines. In addition, by pass-

Node design requirements should take into account that a node has a 10-year lifetime, and its maximum capacity should refer to the estimated traffic conditions at the end of life (a factor of 10 from 2018 is assumed with a yearly traffic growth of ~35 percent).

Transmission scheme	Pros	Cons	Application scenarios
NWDM	• Suitable for long distances • Cost-effective • DSP-enabled adaptive capabilities • Self-performance	• DAC/ADC bandwidth limitation • Nonlinear limitations	• National, long-haul, and ultra long-haul
O-OFDM	• Sub-wavelength granularity • DSP-enabled adaptive capabilities • Electrical subcarrier control/manipulation (BL/PL) • Self-performance monitoring • 1-tap equalizer • Ease of bandwidth/bit rate scalability	• DSP complexity • DAC/ADC bandwidth limitation • Electrical/optoelectronic components linearity • High peak-to-average power ratio • Nonlinear limitations	• Metro, regional, national, and long-haul
TFP	• Highest SE • Self-performance monitoring • No DAC/ADC limitation	• High DSP complexity at the Rx (sequence detector instead of a less complex symbol-by-symbol detector) • Nonlinear limitations	• Data center, national, and long-haul

Table 2. Comparison of transmission schemes.

ing the ultra-short pulses from the MLL to a highly nonlinear fiber (HNLF), a self-phase modulation effect is created, affecting the signal by widening its spectrum and shortening the pulse duration. Within IDEALIST a novel comb generator, based on an MLL and an HNLF, was undertaken, where optical components were optimized. The proposed approach can generate, with a relatively simple setup, a 555 Gb/s super-channel consisting of 52 subcarriers carrying 10.675 Gb/s each, from a single 10 GHz MLL and an HNLF, as displayed in Fig. 2 [8].

Transmission Schemes

In IDEALIST we considered optical transmission schemes addressing the network specifications listed earlier. These also include spectrum allocation techniques, DSP algorithms, and power consumption considerations.

We examined three different transmission methods: Nyquist WDM (NWDM) [2], O-OFDM [9], and time-frequency packing (TFP) [10]. Table 2 reports their main differences.

NWDM uses digital spectral pulse shaping to reduce the spectral width of the signals, allowing for channel spacing equal to the symbol rate, providing higher SE. As mentioned earlier, one of the most innovative features of S-BVTs is to provide data rate on request by varying, for example, the symbol rate. Nevertheless, such a transponder should be engineered for a limited ensemble of modulation schemes, thus keeping costs acceptable. Consequently, starting from the criteria mentioned earlier, we determined the minimum ensemble of needed modulation schemes for an S-BVT to transmit from 100 Gb/s to 1 Tb/s. The selected modulation formats are: pulse modulated quadrature phase shift keying (PM-QPSK), pulse modulated 8-quadrature amplitude modulation (PM-8QAM), and PM-16-QAM [2].

O-OFDM offers unique spectral domain manipulation with super- and sub-wavelength granularity. The former is considered for building super-channels, while the latter includes the electrical subcarrier level, which is ideal for accurately adapting the spectrum to instantaneous needs at the expense of increased DSP complexity. The bit and power per subcarrier can be finely adapted and reconfigured at the S-BVT DSP according to the bandwidth demand and channel profile for a flexible rate/distance adaptive transmission. Furthermore, the S-BVT based on OFDM technology can be sliceable in both time and frequency, thus supporting multiple variable capacity data flows. A key issue for an actual implementation of the OFDM-based S-BVT is DSP optimization for reducing its complexity (e.g., employing alternative fast transforms [9]).

The first fundamental difference, with respect to NDWM or OFDM, is that TFP gives up orthogonality. In TFP, the subcarriers partially overlap in time or frequency or both, an approach that significantly increases SE at the expense of induced linear cross-talk penalty as intersymbol interference (ISI) and intercarrier-interference (ICI). Low-density parity check (LDPC) coding and detection are properly designed to account for the introduced interference. Because of the increased ISI, TFP requires a sequence detector instead of a symbol-by-symbol one. Moreover, at the transmitter a simple electrical filter is needed (DAC is not mandatory, because reach adaptation is achieved by employing variable code rate with PM-QPSK), thus reducing complexity and power consumption. The receiver negligible additional complexity (mainly to reduce the resulting ICI) is hence justified by the transmitter benefit. With particular respect to NWDM, TFP has fewer requirements in terms of LASER wavelength stability (e.g., important in the case of a super-channel) and flexibility, which is enhanced through code rate adaptation [10].

Besides the above transmission techniques, we also investigated mixed, coded, and multi-dimension modulation formats that achieve higher levels of optimization providing finer granularity [1]. Moreover, we considered the power consumption and demonstrated that (energy) sustainable development of high-speed

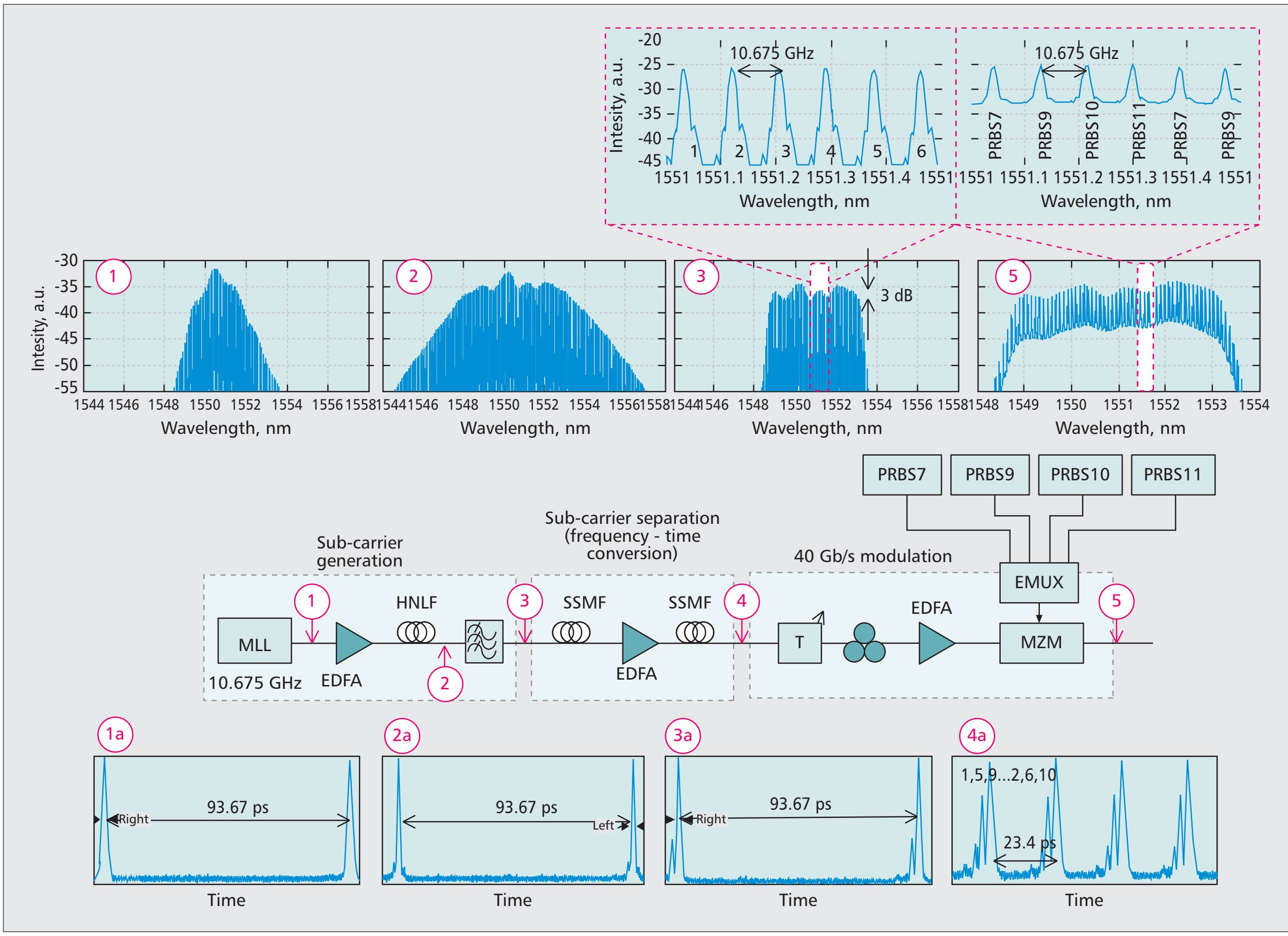


Figure 2. Experimental setup of a 555 Gb/s optical comb generator. PRBS: pseudo-random binary sequence; EDFA: Erbium-doped fiber amplifier; EMUX: electrical multiplexing; MZM: Mach-Zehnder modulator.

transponders is leveraged by SDN. In particular, flexible transponders can considerably reduce energy consumption when adapting, for example, the complexity of the CD compensation algorithm inside the DSP [1], and by evaluating the interplay of FEC and performance vs. power consumption [11]. Finally, we evaluate strategies on spectrum allocation to limit fiber nonlinear propagation impairments in dynamic flexible optical networks [12].

Optical Monitoring Techniques

Optical layer monitoring is exploited by network OAM to accomplish various network functionalities:
- Fault detection
- Commissioning and provisioning
- Performance degradation monitoring
- Verification of service level agreements (SLAs)

Several parameters such as optical power (for fault detection), frequency deviation (to monitor LASER performance degradation), and delay (as verification of the SLAs) may be monitored in view of the aforementioned functionalities. OAM relies on standard monitoring techniques, both at the single subcarrier and super-channel level. Additionally, DSP creates new possibilities for monitoring, especially for the case of subcarrier frequency deviation, dispersion, and mean square error (MSE). In coherent systems, it is possible to measure the offset between the frequencies of the local oscillator and the received subcarrier, exploiting automatic frequency control inside the DSP [13]. Frequency deviation can reveal LASER instability. Differential group delay (DGD) can be monitored through equalizer parameters within the DSP [14]. MSE can be directly monitored through DSP as a subcarrier quality parameter before FEC so that effects of possible signal degradations can be observed. This information can be directly correlated with possible faults or function degradations so that in case of alarm, specific procedures can be adopted.

For example, a multi-function amplifier may generate an MSE increase detected by the receiver with a threshold comparison. In the case of SDN control, the SDN agent sends an alarm message to the SDN controller. At that point, the SDN agent can decide to reroute the connection along a disjoint path. Alternatively, provided that fiber impairments are limited, the SDN controller can increase the redundancy of the adaptive code (i.e., increase robustness to physical impairments) of the transmitted channel so that no rerouting is required.

If data-aided transmission is implemented,

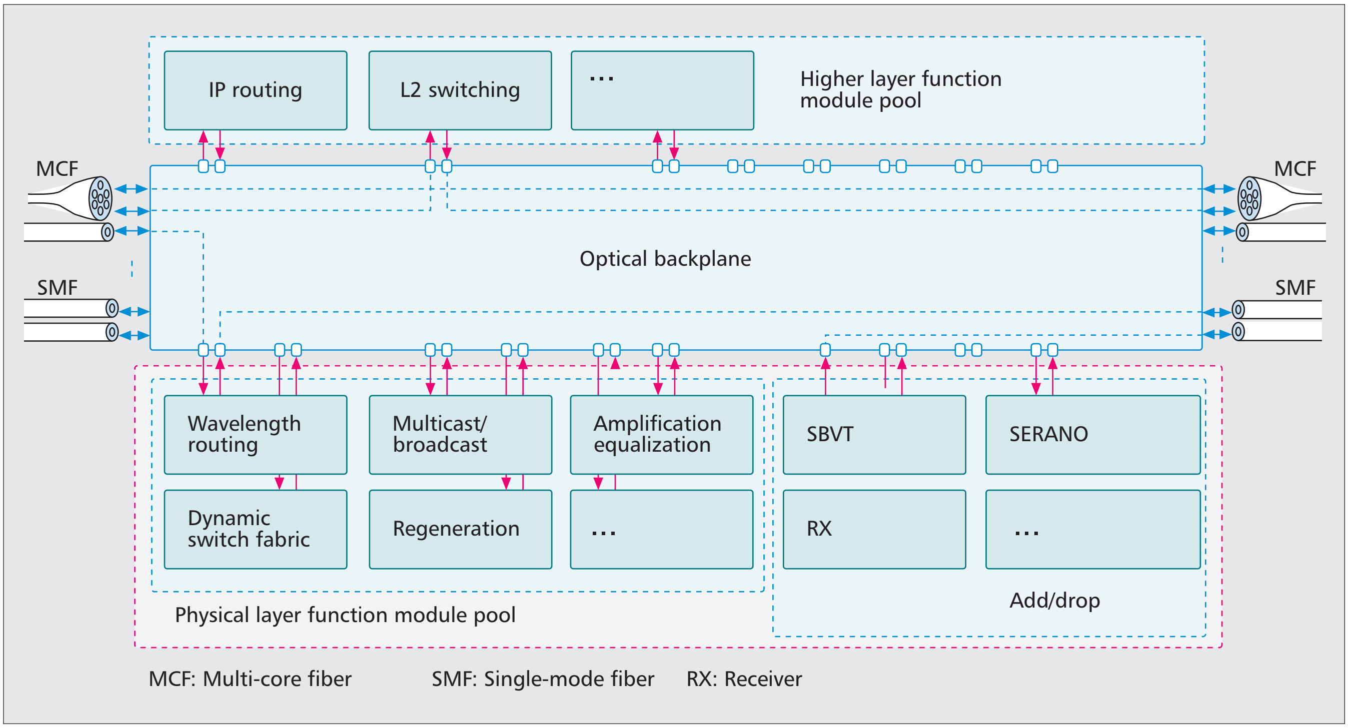


Figure 3. AoD-based NFP node.

the DSP at the S-BVT receiver enables self-performance monitoring in the electrical domain [15], as a set of system parameters required for channel estimation/equalization are acquired thanks to the overhead of information transmitted for a correct detection. Furthermore, for the case of O-OFDM, by allocating guard-bands, we are able to measure the noise by means of an optical spectrum analysis, allowing in-band non-intrusive optical signal-to-noise ratio (OSNR) monitoring that can be placed anywhere along the network [15].

Next Generation Flex-OXC

Optical transport networks continue to introduce new requirements to next generation optical nodes such as *node flexibility* in different domains (e.g., switching, bandwidth and transmission rate), node scalability to higher transmission capacities, *node reliability* and *survivability* (considering redundant components), and simple node adaptability to emerging network applications. Based on this *node adaptability*, various node subsystem architectures with different internal component distribution approaches can be depicted and supported by the proposed new Flex-OXC.

AoD-Based Network Function Programmable Node

As EONs are envisaged as the future optical network infrastructure, it is mandatory to adopt nodes with advanced features. These optical nodes consider the programmability of functions at higher layers (e.g., routing, switching) and the physical layer (e.g. amplification, regeneration) at the request of the network user, who is unaware of the available optical hardware resources (Fig. 3). This novel network function programmable (NFP) node paradigm brings to the optical network a new perspective, since the abstraction of optical (and electronic) components for network functionality is practically achieved, and node slices are enabled associated with arbitrary traffic types.

The NFP node is supported by the AoD concept [4], consisting of an optical backplane connected to the node inputs/outputs and several plug-in modules, providing the required signal processing functions, including bandwidth-variable WSS (BV-WSS), fast optical switches, power splitters, fixed-grid demultiplexers, EDFAs, SERANO, and so on. With AoD, different arrangements of inputs, modules, and outputs can be constructed by setting up appropriate cross-connections in the optical backplane. Therefore, the AoD-based NFP provides increased flexibility since the components used for optical processing are not hard-wired as in a static architecture, enabling module interconnection in an arbitrary manner. Flexibility analysis was undertaken in [4] demonstrating that the NFP node provides routing, switching, and architectural flexibility. In addition, the NFP node performance was investigated through experimental demonstration, showing dynamic composition and reconfiguration of synthetic architectures to support different sets of signals and requirements. The functional and architectural flexibility of the AoD node can provide on-demand elastic time-spectrum switching, subwavelength channel aggregation, and spectrum defragmentation. This experimental demonstration involved signals with a variety of bit rates, modulation formats, bandwidth, and signal processing requirements. Satisfactory per-

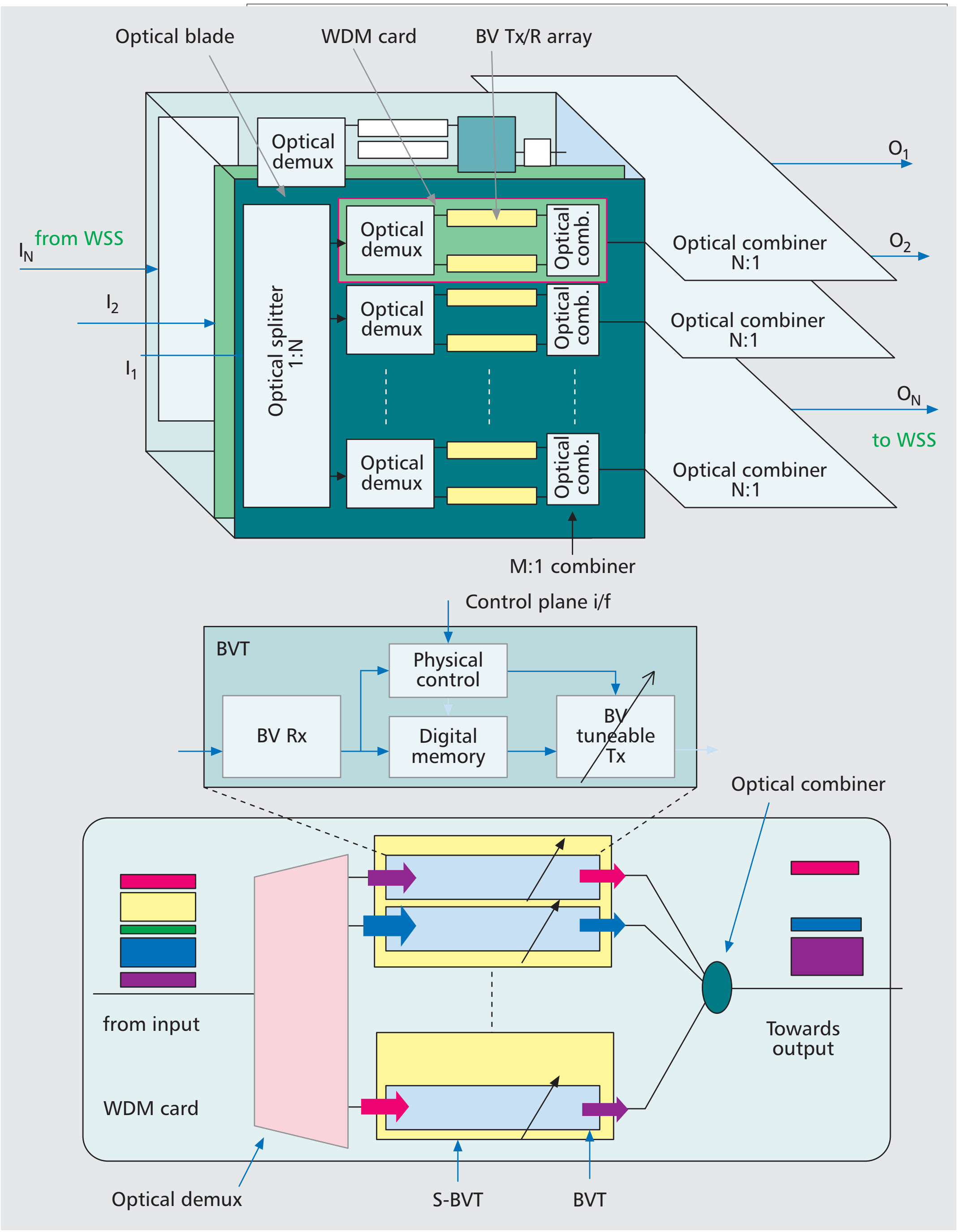


Figure 4. a) Overview of the SERANO architecture; b) details of the WDM card and the S-BVTs.

Optical transport networks continue to introduce new requirements to next generation optical nodes such as node flexibility in different domains, node scalability to higher transmission capacities, node reliability and survivability (considering redundant components), and simple node adaptability to emerging network applications.

formance is presented with a maximum power penalty of 1.9 dB.

With respect to node synthesis, to cope with the demanding traffic requirements of the optical network, NFP nodes have proven to be highly beneficial since an algorithm is used to calculate a synthetic node design providing the required functionality with available modules. Furthermore, if traffic requirements change, an alternative synthetic node design is calculated and implemented to fulfill the new requirements hitlessly (no traffic loss). Based on performance analysis, the NFP node has demonstrated hardware reductions (i.e., node cost) of up to 40 percent compared to static ROADMs through proper network design. This advantage also leads to 25 percent savings in the overall power consumption by diminishing the number of SSSs. Furthermore, NFP nodes also support node scalability. This scalability in the NFP node has also been studied, proving the possibility to reduce at least by half the number of hardware modules used, compared to other conventional architectures, by applying fiber switching, which involves simple cross-connections when several traffic demands require the same output port destination. This feature is only available in the AoD node, since the cross-connection can be configured depending on the traffic scenario.

A reduction in terms of capital expenditure is achieved by equipment and device integration. As far as operational expenditures are concerned, the improved energy efficiency of EONs has an impact both from a network architecture perspective and at the individual component level.

The SERANO Architecture

A SERANO architecture is an integral building block of a Flex-OXC; it makes use of the unique features S-BVTs offer to reduce overall switching complexity and cost [5]. In particular, a fraction of the traffic through the Flex-OXC is added/dropped from/to the SERANO block by means of *N* I/O ports/fibers as illustrated in Fig. 4a. Each input fiber is terminated in an optical blade which consists of an 1:*N* optical splitter where each of the *N* outputs are sent to *N* WDM cards in parallel. As illustrated in Fig. 4b, a WDM card incorporates an optical demultiplexer followed by a multi-flow optical module consisting of fixed-receiver tunable-transmitter array pairs connected back to back with no client interfaces between them. Therefore, the functionality of this block is to terminate the signals of the entire comb of spectral slots but only full 3R regeneration to a selected number of flows (i.e., those directed to the selected output fiber at the desired spectral slot/modulation format).

The final stage of the WDM card consists of a combiner with *M* input ports, where *M* is the upper number of flows a particular input fiber may forward to a particular output fiber. In the final stage of SERANO, the outputs of each WDM card are passively recombined by means of an additional *N*:1 optical combiner. The SERANO block, apart from 3R and modulation format adaptation to tailor line rate to the subsequent transparent section length, provides for vendor interoperability, since the receiver-transmitter block arrays could potentially be purchased separately, and for spectrum fragmentation. The aforementioned SERANO architecture is flow and link modular since only the couplers need to be provisioned; the building blocks of optical blades and WDM cards can be added progressively following a pay-as-you-grow principle. It is also worth noting that no other switching technology, optical or electronic, is introduced since the main active building block is multi-flow optical module arrays. This integrates transmission and data forwarding, allowing the industry to concentrate on optimizing a single element for both functions. A preliminary physical layer performance/scalability of the proposed architecture, considering fixed-grid data center applications, is reported in [5].

Advantages of Flexible Optical Communication Systems

Migration from fixed-grid networks to EONs provides several benefits. Hereafter, we summarize the most relevant.

Increased spectral efficiency: By enabling optimization of the spectral allocation on request, the flexible ITU channel grid increases SE and network capacity significantly. Deployed optical links can be used more efficiently, thus prolonging their lifespan.

Accommodation of 1 Tb/s and beyond client signals: Client signal bit rates of 1 Tb/s and beyond will inevitably exceed the limits set by the fixed 50 GHz channel grid. However, flex-grid architectures can accommodate such demands.

Dynamic reconfiguration: EONs enable dynamic reconfiguration of the network by using S-BVTs as basic building blocks. Such potential holds great promise for increased energy efficiency of the network and improved multi-layer protection applications.

Better economics: A reduction in terms of capital expenditure is achieved by equipment and device integration. As far as operational expenditures are concerned, the improved energy efficiency of EONs has an impact both from a network architecture perspective and at the individual component level. Another issue of fixed-grid networks is the number of transponders required, which may significantly increase when large bandwidth demands have to be accommodated by inverse multiplexing.

Conclusions

We have investigated the challenges deploying EONs for terrestrial European networks under the assumption that the currently estimated CAGR will be maintained until 2018. A comprehensive analysis of technical requirements and performance of novel key elements such as S-BVT and Flex-OXC has been performed according to the network capacity needs.

The novel concept of a multi-flow optical module, implementing a flexible subcarrier module, has been investigated and its performance using different transmission methods compared. Moreover, advanced DSP techniques and energy saving modulation formats have also been deeply studied.

Concerning the Flex-OXC architecture, two innovative and complementary solutions that can handle dynamic traffic and present a high level of scalability and flexibility, AoD and SERANO, have been investigated. These solutions enable a clear upgrade path towards EONs.

To conclude, we firmly believe that EONs will significantly postpone the imminent capacity crunch, and that an efficient bandwidth allocation will build the foundation of next generation optical networks.

Acknowledgment

This work was supported by FP-7 IDEALIST project, grant agreement number 317999.

References

[1] P. Layec *et al.*, "Elastic Optical Networks: The Global Evolution to Software Configurable Optical Networks," *Bell Labs Tech. J.*, vol. 18, no. 3, 2013, pp. 133–51.
[2] J. K. Fischer *et al.*, "Bandwidth-Variable Transceivers Based on Four-Dimensional Modulation Formats," *J. Lightwave Tech.*, vol. 32, no. 16, Aug. 2014, pp. 2886–95.
[3] M. Jinno *et al.*, "Multiflow Optical Transponder for Efficient Multilayer Optical Networking," *IEEE Commun. Mag.*, vol. 50, no. 5, May 2012, pp. 56–65.
[4] N. Amaya, G. Zervas, and D. Simeonidou, "Introducing Node Architecture Flexibility for Elastic Optical Networks," *J. Opt. Commun. and Net.*, vol. 5, no. 6. June 2013.
[5] A. Stavdas *et al.*, "A Novel Architecture for Highly Virtualised Software-Defined Optical Clouds," *ECOC*, Paper: Mo.3.E.4, 22–26, London, U.K., Sept. 2013.
[6] A. Aguado *et al.*, "ABNO: A Feasible SDN Approach for Multi-Vendor IP and Optical Networks," *Proc. OFC*, Th3I.5, Mar. 2014.
[7] A. Napoli *et al.*, "Reduced Complexity Digital Back-Propagation Methods for Optical Communication Systems," *J. Lightwave Tech.*, vol. 32, no. 7, Apr. 2014, pp. 1351–62.

[8] N. Amaya *et al.*, "On-Demand Spectrum and Space Defragmentation in an Elastic SDM/FDM/TDM Network with Mixed Multi- and Single-core Fiber Links," *OFC*, Paper: OTh4B.3, Anaheim, CA, Mar. 2013.
[9] M. Svaluto Moreolo *et al.*, "Experimental Demonstration of a Cost-Effective Bit Rate Variable IM/DD Optical OFDM with Reduced Guard Band," *Optics Express*, vol. 20, no. 26, Dec. 2012, pp. B159–B164.
[10] L. Potì *et al.*, "Casting 1 Tb/s DP-QPSK Communication into 200 GHz Bandwidth," *Proc. ECOC*, Paper: P4.19, 2012.
[11] Danish Rafique *et al.*, "Technology Options for 400 Gb/s PM-16QAM Flex-Grid Network Upgrades," *IEEE Photonics Tech. Lett.*, vol. 26, 2014, pp. 773–76.
[12] Danish Rafique *et al.*, "Flex-Grid Optical Networks: Spectrum Allocation and Nonlinear Dynamics of Super-Channels," *Optics Express 21*, no. 26, 2013, 32184–191.
[13] G. Berrettini *et al.*, "Frequency Conversion and Source-Independent Push-Pull for Super-Channels in Flex-Grid Optical Networks," *IEEE Commun. Lett.*, vol. 18, no. 2, Feb. 2014, pp. 364–67.
[14] E. Matarazzo *et al.*, "DGD Monitoring Issues in High-Speed Polarisation Multiplexed Coherent QPSK Systems," *IEE Elect. Lett.*, vol. 48, no. 8, Apr. 2012, pp. 446–48.
[15] J. M. Fàbrega *et al.*, "All-Optical In-Band OSNR Measurement in Intensity-Modulated Direct-Detection Optical OFDM Systems," *Proc. Int'l. Conf. Transparent Optical Networks*, Cartagena, Spain, 2013

Biographies

ANTONIO NAPOLI (antonio.napoli@coriant.com) received M.Sc. and Ph.D. degrees in electrical engineering from Politecnico di Torino in 2002 and 2006, respectively. In 2004 he was a visiting student at Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, and in 2005 at University College London, United Kingdom. In 2006, he joined Siemens COM and in 2007 NSN, working on DSP for high-speed optical networks. Since 2013 he has been with Coriant. His main interests are DSP and advanced modulation formats for long-haul optical systems. He has (co)-authored 80+ peer-reviewed articles and seven patents.

MARC BOHN received his Dipl.-Ing. degree in electrical engineering in 1999 and his Ph.D. degree in 2005 from the University of Kiel, Germany. He was with Siemens, Nokia Siemens Networks, and Coriant working on transmission aspects in optical communications in different research and development and management functions. Since 2013 he has been heading the Research and Technology Department inside Coriant.

DANISH RAFIQUE holds a Ph.D. in electrical engineering (2012). He joined Nokia Siemens Networks in 2012, and is currently pursuing a career with Coriant R&D GmbH, Germany, where he is working on product development and research programs on next-generation transponder solutions involving flex-grid networks. He has published more than 55 articles, and acts as a reviewer for *IEEE PTL*, *IEEE JLT*, *OSA Optics Express*, and others.

ALEXANDROS STAVDAS holds a B.Sc. in physics from the University of Athens, Greece, an M.Sc. in optoelectronics and LASER devices from Heriot-Watt /St-Andrews University, United Kingdom, and a Ph.D. from University College London in WDM networks. Currently he is a professor of optical systems and networking in the Department of Informatics and Telecommunications of the University of Parma (UoP), Italy. He is an author of over 140 journal publications and conference articles. He has also served as a TPC member for many international conferences. His current research interests include future Internet architectures, heterogeneous access networks, physical layer modeling of optical networks, OXC architectures, and optical packet/burst switching.

NICOLA SAMBO (n.sambo@sssup.it) received his M.Sc. cum laude in telecommunication engineering from the University of Pisa, Italy, in 2004, and his Ph.D. from Scuola Superiore Sant'Anna (SSSUP), Pisa, in 2009. In 2007, he was a visiting student at France Télécom, Lannion, France. Currently, he is an assistant professor at SSSUP. He is also collaborating with Inter-University National Consortium for Telecommunications (CNIT) and Ericsson Italy. His research interests include optical network architecture, transmission performance modeling, and the control plane. He is an author of papers published in more than 80 publications including international journals, conference proceedings, and patents.

LUCA POTÌ received his M.S. degree in electronics engineering from UoP in 1997. Since 2010 he has been head of the research area in the Photonic Networks National Laboratory in Pisa, Italy. He has published more than 350 international journal papers, conference papers, and patents.

MARKUS NÖLLE received his Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from the Technische Universität Berlin in 2007 and 2013, respectively. Since 2008, he has been with the Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute (FhG-HHI) in Berlin, Germany. Currently he is a researcher and project manager in the Submarine and Core Systems group within the Department of Photonics Networks and Systems. His research interests are spectrally efficient higher-order optical modulation formats and digital distortion mitigation in coherent optical systems.

JOHANNES KARL FISCHER [S'04, M'09] received his Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from the Technische Universität Berlin in 2003 and 2009. He is with FhG-HHI, where he heads a research group on digital signal processing. His research interests are advanced modulation formats, digital signal processing for coherent systems.

EMILIO RICCARDI, after graduating in physics, joined Telecom Italia Lab (formerly CSELT) in 1992. His current interest is dedicated to elastic optical networks. He has been involved in several EU projects and presently leads the data plane Working Package of the IDEALIST EU project. He has coauthored more than 40 journal and conference publications.

ANNACHIARA PAGANO received her degree in physics from the University of Turin in 1994. In the same year, she joined Telecom Italia (formerly CSELT) to work on optical fiber optics and DWDM metro and long distance communication systems. Her current fields of interest are next-generation optical networks, focusing both on high-capacity network design and testing, together with research activities on optical transmission and scouting of new technologies. She has co-authored more than 50 technical publications covering design and testing aspects of optical components and systems. She has been involved in several European projects.

ANDREA DI GIGLIO received a Dr. Ing. degree in electronic engineering from the University of Pisa and his Engineering License degree from SSSUP. He joined Telecom Italia Lab (formerly CSELT), which is Telecom Italia Group's company for study, research, experimentation, and qualification in the field of telecommunications and information technology. The fields of his researches are addressed toward Internet security, storage area networks, and optical and packet network architecture. He has been involved in several EURESCOM and European projects on IP over optical networks, and coordinated ICT STRONGEST

MICHELA SVALUTO MOREOLO [S'04, A'08, SM'13] received her M.Sc. degree in electronics engineering and Ph.D. degree in telecommunications engineering from University Roma Tre, Italy, in 2003 and 2007, respectively. She is currently a senior researcher in the Communication Networks Division and coordinator of the Optical Transmission and Subsystems research line at CTTC, Spain. Her research interests include optical/digital signal processing and advanced transmission technologies for future optical networks. She has authored/co-authored over 70 publications.

JOSEP MARIA FABREGA [S'05, M'10] received his Ph.D. degree in signal theory and communications from UPC in 2010, for which he received the EuroFOS student research award. Since 2010 he has been with the Optical Networks and Systems Department at the Centre Tecnologic de Telecomunicacions de Catalunya, Spain. He has published +60 papers. His research interests include optical advanced modulation formats.

EMILIO HUGUES-SALAS (e.huguessalas@bristol.ac.uk) received his M.Sc. and Ph.D. degrees from the University of Essex, United Kingdom. He is a research fellow in the High Performance Networks Group at the University of Bristol, United Kingdom. His research interests include the development of high-capacity, reconfigurable, flexible, and multidimensional optical nodes and networks. He is an author and a co-

We firmly believe that EONs will significantly postpone the imminent capacity crunch, and that an efficient bandwidth allocation will build the foundation of next generation optical networks.

author of over 60 international journals and conference papers.

GEORGIOS ZERVAS holds M.Eng. and Ph.D. degrees from the University of Essex. He is a lecturer in optical and high performance networks at the University of Bristol and a member of HPN group. He leads activities in optical transport and switching, as well as software/hardware programmable systems and networks for backbone and DCN/HPC interconnects. He has published over 150 journal and conference papers, including post-deadline papers, and has received best paper awards. He has been a TPC member of numerous leading conferences and a contributor to standards.

DIMITRA SIMEONIDOU is a professor in the High Performance Networks (HPN) group at the University of Bristol. She is a leading academic in the field of optical networks, a pioneer of future Internet research and experimentation (FIRE), an influential researcher in the fields of grid/cloud computing, and a founder of Transport SDN. She has chaired several international conferences (ECOC '13, OFC '15) and committees across these technical fields. She is author/co-author of over 400 publications, many of which have received best paper awards, 12 patents, and several major contributions to standards

PATRICIA LAYEC is a Bell Labs researcher in France. She received her Ph.D. degree from University of Paris-Sud XI, France, in 2009. In 2005, she worked at Orange on MIMO transceiver design and multi-user information theory. In 2009, she joined Alcatel-Lucent, first working on wireless physical layer dimensioning (HSPA+/LTE), then joining Bell Labs. Her research interests include information theory, DSP, and elastic optical networks. She led the CELTIC European project EO-NET, which received the Celtic-Plus gold excellence award in 2014.

ANTONIO D'ERRICO received his Ph.D. degree cum laude in telecommunication engineering from the SSSUP in 2008. He is currently with Ericsson Research as an experienced researcher. His research interests include advanced technological solutions for optical networks. He is an author of more than 75 papers published in international journals, conference digests, and patents.

TALHA RAHMAN received his M.Sc. (Hons.) degree in advanced optical technologies with specialization in optical communication systems from FAU Erlangen-Nürnberg, Germany, in 2012. He conducted his internship and Master's thesis research at Alcatel-Lucent Bell Labs Stuttgart. He is currently employed by Eindhoven University of Technology as a Ph.D. candidate under EU-FP7 project IDEALIST and conducts his research at Coriant R&D Labs in Munich, Germany. His area of research includes flexible optical transceivers focusing on 400 Gb/s and 1 Tb/s superchannels

JUAN PEDRO FERNÁNDEZ-PALACIOS GIMÉNEZ received his M.S. in telecommunications engineering from Polytechnic University of Valencia in 2000. In September 2000 he joined TIDSA, where he currently leads the Metro and Core Transport Networks Department. He has been involved in several European projects such as Eurescom P1014 TWIN and IST projects DAVID, NOBEL, NOBEL-2, STRONGEST, BONE, and MAINS, as well as other internal projects related to the development of optical networks in the Telefónica Group. Currently he is coordinating the FP7 project IDEALIST as well as the standardization activities within the cluster, where he is acting as co-chair.

Next Generation Sliceable Bandwidth Variable Transponders

Nicola Sambo, Piero Castoldi, Antonio D'Errico, Emilio Riccardi, Annachiara Pagano, Michela Svaluto Moreolo, Josep M. Fàbrega, Danish Rafique, Antonio Napoli, Silvano Frigerio, Emilio Hugues Salas, Georgios Zervas, Markus Nölle, Johannes K. Fischer, Andrew Lord, and Juan P.F.-P Gimenez

Nicola Sambo and Piero Castoldi are with CNIT.

Antonio D'Errico is with Ericsson.

Emilio Riccardi and Annachiara Pagano are with Telecom Italia.

Michela Svaluto Moreolo and Josep M. Fàbrega are with CTTC.

Danish Rafique and Antonio Napoli are with Coriant GmbH.

Silvano Frigerio is with Alcatel-Lucent Italia.

Emilio Hugues Salas and Georgios Zervas are with the University of Bristol.

Markus Nölle and Johannes K. Fischer are with Fraunhofer HHI.

Andrew Lord is with British Telecom.

Juan P.F.-P Gimenez is with Telefonica.

ABSTRACT

This article reports the work on next generation transponders for optical networks carried out within the last few years. A general architecture supporting super-channels (i.e., optical connections composed of several adjacent subcarriers) and sliceability (i.e., subcarriers grouped in a number of independent super-channels with different destinations) is presented. Several transponder implementations supporting different transmission techniques are considered, highlighting advantages, economics, and complexity. Discussions include electronics, optical components, integration, and programmability. Application use cases are reported.

INTRODUCTION

Operators have estimated continued growth of around 35 percent per year in backbone networks, stemming from drivers including mobile Internet, high-resolution video, and increased use of cloud-based services. This seems set to continue with demand coming from the Internet of Things, ubiquitous sensor deployment, and machine-to-machine communications. Together with high bandwidth growth, these services are anticipated to generate increasingly dynamic traffic.

Installing more of the same technology, where costs scale linearly with bandwidth, will quench these growth trends, as the bandwidth price will be too high to support the new services. One approach to this problem is to efficiently use the optical spectrum and consequently significant research is now being directed toward flexgrid or elastic optical networks (EONs) [1].

One key aspect of an EON is a flexible transponder. These transponders will have a range of functions, including support of multiple bit rates (e.g., from 10 Gb/s to 1 Tb/s) and dynamically changeable modulation formats and baud rates. This flexibility enables the operator to optimize network capacity by configuring the transponder to have the most spectrally efficient format for the distance required.

However, further flexibility is under intense study to design transponders capable of generating multiple optical flows, routed to different destinations by optical switches and filters in the network. A single transponder with this capability is known as a multi-flow transponder or a sliceable bandwidth variable transponder (S-BVT) [2]. S-BVTs enable the generation of multiple optical flows that can be routed into different *media channels* (a media channel is a specific portion of the optical spectrum and an optical path through the EON between two endpoints) and flexibly directed toward different destinations.

This article reports investigations on S-BVT technology. The article begins by providing a design architecture for the S-BVT, including client and line side as well as the required internal switching. The OTN layer is key, having a broadened scope, which now facilitates the flexible mapping of clients onto media channels. Then we consider the most promising transmission techniques to build such a transponder: Nyquist wavelength-division multiplexing (NWDM), orthogonal frequency-division multiplexing (OFDM), and time frequency packing (TFP), comparing their capabilities. We then discuss practical issues around manufacturing, integration, and programmability. We complete the review with a discussion of some of the key operator future applications for S-BVTs.

S-BVT ARCHITECTURE

S-BVT is an evolution of the BVT, which is a class of transponders able to dynamically tune the required optical bandwidth and transmission reach by adjusting parameters such as gross bit rate, forward error correction (FEC) coding, modulation format, and shaping of optical spectrum. BVTs enable a trade-off between spectral efficiency and transmission reach, using spectral efficient modulation formats (e.g., polarization multiplexing 16-quadrature amplitude modulation, PM-16-QAM, PM-64-QAM) for short-reach connections, and more robust but less efficient modulation schemes (e.g., PM-quadrature phase shift keying, PM-QPSK) for long-haul links. The transmission bandwidth (up to a maximum value) is exclusively dedicated to a single traffic demand and adjusted in fixed steps under

S-BVT is an evolution of the Bandwidth Variable Transponder (BVT), which is a class of transponders able to dynamically tune the required optical bandwidth and transmission reach by adjusting parameters such as gross bit rate, FEC coding, modulation format and shaping of optical spectrum.

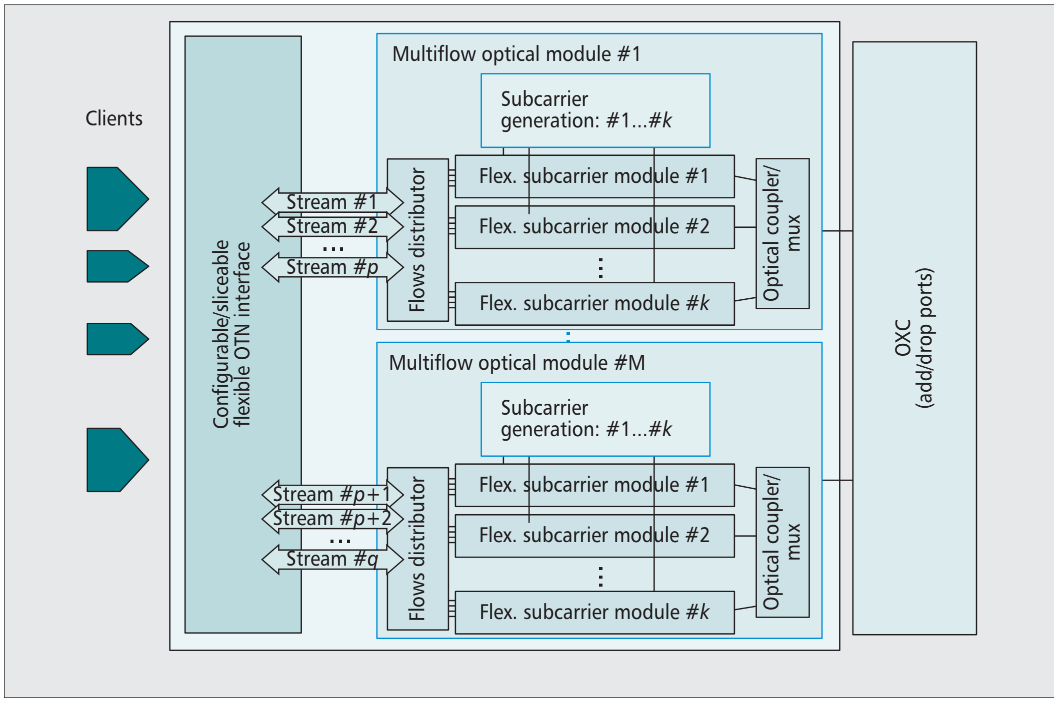


Figure 1. S-BVT architecture.

software control in order to be dynamically adapted to the actual traffic demand and reach. A single optical carrier is generated, feeding a single media channel.

S-BVTs enhance BVT functionalities by being able to allocate their capacity into one or several independent optical flows that are transmitted toward one or multiple destinations. Consequently, unlike BVTs, their net transmission bandwidth (up to a maximum value) may be spread (or sliced) to serve several independent traffic demands simultaneously. All S-BVT electronic and optical resources can be flexibly partitioned into a number of groups equal to the number of simultaneous independent optical flows. As a result, the optical output of an S-BVT is a group of super-channels with different destinations and modulation formats employing different portions of the optical spectrum and media channels, efficiently feeding a single add/drop port. Thus, an S-BVT should be considered as a collection of "virtual" lower-capacity BVTs (one for each elemental subcarrier), logically associated in groups to generate independent super-channels. A general S-BVT modular architecture supporting several transponder implementations (e.g., supporting NWDM or OFDM transmission) is reported in Fig. 1 and will be described hereafter.

In order to fully exploit sliceability, specific functions are suitably distributed among electrical (optical transport network — OTN — layer) and optical layers, as shown in Figs. 1 and 2. The first role is played by the electronic layer, which adapts service information content (e.g., coming from the IP layer) to the photonic layer. This feature is in charge of the optical transport hierarchy (OTH) of the OTN, described in International Telecommunication Union — Telecommunication Standardization Sector (ITU-T) G.872 and G.709, which is located in the **configurable/sliceable flexible OTN interface** module. The adaptation of service information for the photonic layer consists of creating a data structure including data information (from the IP layer clients) and redundancy to accomplish forward error correction (FEC) and monitoring. Moreover, besides FEC and monitoring features, the configurable/sliceable flexible OTN interface module can flexibly fragment high data rates (e.g., 1 Tb/s) into lower data rates (e.g., 600 Gb/s plus 400 Gb/s) if it is impossible to fit the total client traffic into a single super-channel (e.g., due to unavailable contiguous spectrum along the path). An example is shown in Fig. 2: the traffic from a client at 1 Tb/s is divided by the OTN layer (a Flex OTU frame in the figure) into two different streams, each feeding two separate media channels (#1 and #2). Similarly, the traffic from an interface at 200 Gb/s is divided into two different streams, one complementing media channel #2 and the other one feeding media channel #3. Traffic could also be not fragmented, as in the case of the other 200 Gb/s interface and the 100 Gb/s interface in Fig. 2. A rate-flexible OTU (Flex OTU) line interface, complementing the flexibility currently provided by ODUflex in the LO-ODU service layer (ITU-T G.709), is expected to be one of the distinguishing attributes of the OTN hierarchy advancement "beyond 100G" (B100G OTN). For more details related to the ODUflex concept, the reader can refer to ITU-T Recommendation G.709, and to [3] for an introduction to OTUflex. The current predominant thinking in the OTN evolution debate is in favor of defining an $n \times 100$ Gb/s (with $n \geq 2$) iterative structure achieved by interleaving n subframes, resembling a standard OTU4 (i.e., 100 Gb/s information rate plus overhead). The index n suggests the

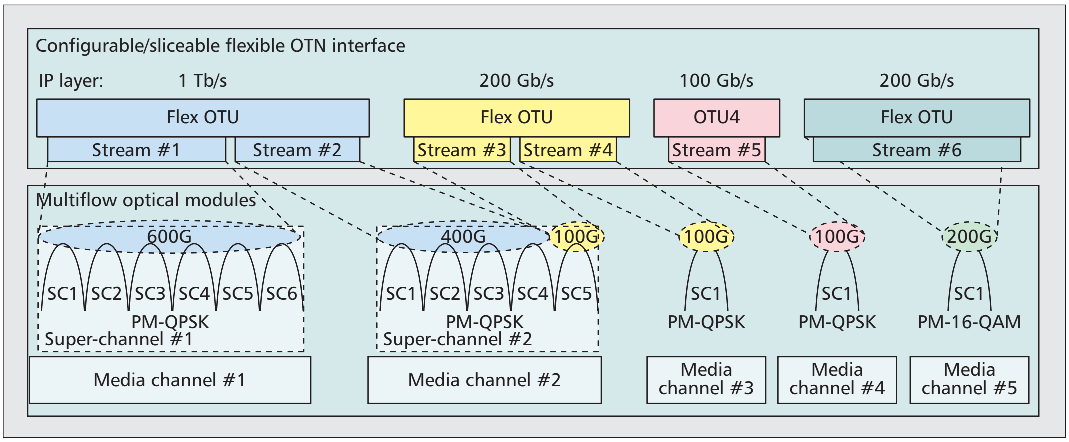


Figure 2. OTN frames and flexible association with media channels.

The Flex Sub-carrier module has been implemented focusing on three main transmission techniques: NWDM, OFDM, and TFP. These techniques present different levels of spectral efficiency (measured in b/s/Hz) and complexity.

possibility of variable bit rates in 100 Gb/s steps, as shown in Fig. 2, with the potential of segmenting high-rate data (e.g., 1 Tb/s) into multiple thinner streams to be transported over multiple subcarriers (e.g., as in the case of media channel #1 composed of six sub-carriers, each one supporting 100 Gb/s information rate). Lower service granularities can be accommodated in the electrical domain by sub-wavelength multiplexing, as specified by G.709 for current OTN systems.

Then the generation of super-channels is in charge of the **multiflow optical** modules (Fig. 1). Each multiflow optical module generates multiple subcarriers that can be distributed among different super-channels. The p independent streams from the OTN layer (in the form of electrical multi-lane aggregates) are used to modulate the optical subcarriers within the multiflow optical module. A **flow distributor** module, a **subcarrier generation** module generating up to k non-modulated subcarriers, a number $k \geq p$ of **flex subcarrier** modules (where each subcarrier is modulated by a specific traffic portion), and a **multiplexer** module (or a coupler) compose the multiflow optical module. First, the flow distributor (e.g., an electronic switching matrix) directs the OTN streams to specific flex subcarrier modules. The flow distributor enables the traffic to modulate a single subcarrier for transmission in the optical layer. The flex subcarrier is shown in Fig. 3. At the transmitter side, data can be encoded (e.g., for TFP, see next section), shaped (filtering), and predistorted to compensate for possible distortions introduced by the optical layer (e.g., NWDM). Digital-to-analog converters (DACs) are exploited to adapt the modulation format (e.g., providing multi-level signals for 16-QAM). IQ modulators are exploited to modulate a generic subcarrier at frequency f_i, while a polarization rotator (PR) has orthogonal polarizations, needed for polarization multiplexing, which doubles the information rate. The bit rate and bandwidth can be tuned by changing the modulation format, the baud rate, spectral shaping, and coding. All generated subcarriers are first logically grouped into super-channels feeding different media channels and optically coupled together, through the multiplexer or the coupler, into a single add/drop port of the source optical cross-connect (Flex-OXC) node. Depending on the number of subcarriers, a coupler can be used instead of a multiplexer: for a high number of subcarriers (e.g., eight), the losses introduced by the coupler can be too high, so a more expensive multiplexer (introducing less loss) should be adopted. Then, with proper node configuration, media channels are routed independently and transparently across the photonic network up to specific (and different) destination nodes. At the receiver side (Fig. 3), coherent detection can be exploited. A polarization beam splitter (PBS) divides the signal into two arbitrary but orthogonal polarizations, which are superimposed with the output of a local source (LS) (e.g., a laser with a narrow line width and higher power compared to the detected signal, i.e., the receiver sensitivity is consequently enhanced). The polarization components are fed into a 90° hybrid and afterward mixed with the LS to separate the in-phase and quadrature components, which are then detected by balanced photo-diodes (BPDs). This is followed by a radio frequency amplifier (RFA), and four high-speed analog-to-digital converters (ADCs). Standard DSP blocks such as clock recovery, equalizer, and phase recovery follow and terminate the receiver. Finally, decisions on the received data and decoding are performed.

Several solutions have been proposed and considered for implementing the modules shown in Fig. 1, as detailed in the next sections.

Transmission Techniques for the Flex Subcarrier Module

The flex subcarrier module has been implemented focusing on three main transmission techniques: NWDM, OFDM, and TFP. These techniques present different levels of spectral efficiency (measured in bits per second per Hertz) and complexity. The choice of transmission technique may depend on the particular scenario, and can be influenced by the links and affordable or required costs. Figure 4 shows typical examples of spectrum achieved with such transmission techniques, while Table 1 summarizes their main characteristics.

OFDM: Optical OFDM (O-OFDM) is based on the transmission of multiple orthogonal sub-carriers, which can be independently modulated by different formats. The orthogonal sub-carriers are overlapped in the frequency domain, providing a high spectral efficiency.

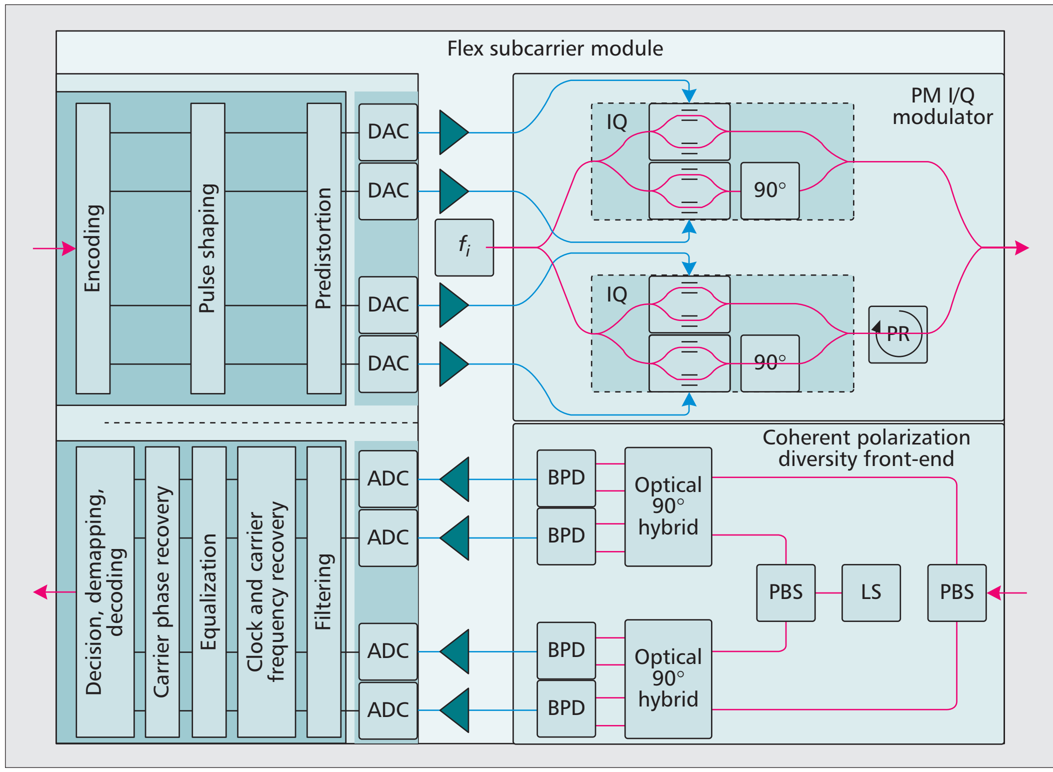


Figure 3. Flex subcarrier module building blocks.

• **NWDM** is a straightforward transmission implemented in the S-BVT flex subcarrier module. The main idea is to apply a pulse-shaping filter at the transmitter (see the **pulse shaping** module in Fig. 3), confining the bandwidth within the Nyquist frequency of the signal, which is half the symbol rate. This allows dense spectral allocation of NWDM subcarriers in a media channel as shown in Fig. 4c, creating a super-channel. Root raised cosine (RRC) is a popular choice for the pulse-shaping filter. Matched RRC filters at transmitter and receiver (**filtering** module in Fig. 3) reduce the intersymbol interference (ISI) due to the narrow filtering. A sharp rolloff (i.e., ≤ 0.2) is essential to reduce linear crosstalk with neighboring channels. Predistortion may be required to compensate for the electrical bandwidth limitation (e.g., of the DAC), and for linear and nonlinear optical layer impairments (e.g., optical filtering effects). A certain number of NWDM subcarriers are multiplexed to construct a super-channel according to current traffic demands. The bit rate carried by each NWDM subcarrier can be programmed individually (e.g., by varying the modulation format [4]). Thus, the control plane, possibly implemented by a software defined network (SDN), can adjust key parameters such as the bit rate per subcarrier, spectral efficiency, and number of NWDM subcarriers in a super-channel to the actual traffic demands. Furthermore, FEC can be varied to satisfy current traffic requests and all-optical reach. Such a software-defined super-channel traverses links as a single channel entity (e.g., it will be routed as one channel) embedded in a single media channel. NWDM achieves high spectral efficiency and is suitable for long haul transmission.

• **OFDM**: Optical OFDM (O-OFDM) is based on the transmission of multiple orthogonal subcarriers, which can be independently modulated by different formats. The orthogonal subcarriers are overlapped in the frequency domain, providing a high spectral efficiency (SE). In all-optical OFDM, the frequency multiplexing/demultiplexing is performed in the optical domain (by specialized devices), and the number of subcarriers is limited to minimize cost and complexity. Conventional optical OFDM systems are based on DSP, and the OFDM subcarriers are generated in the electrical domain before optical modulation, exploiting an additional module with respect to Fig. 3. Thus, also compared to all-optical OFDM, a finer granularity (on the order of hundreds or even tens of megahertz) and narrower subcarrier spacing can be achieved, yielding unique bit rate/bandwidth scalability and spectral domain manipulation capability, with sub- and super-wavelength granularity, and thus suitable for EONs. As shown in Fig. 4 and as demonstrated in [5], where OFDM and NWDM are compared, using a number of electrical subcarriers equal or greater than 64, a more flexible OFDM spectrum with squared profile is obtained, and the same SE as NWDM is achieved. The BVT based on OFDM technology can also be designed to be sliceable in time and frequency, supporting a wide range of granularities. A key element for enabling the transponder programmability and reconfigurability is the DSP, which allows adaptive modulation format selection at each subcarrier, and variable code rate and reach, according to the traffic requirements and channel profile (as further specified later). Both non-coherent and coherent optical front-ends are considered for the programmable flex subcarrier module. Either intensity modulation or linear field modulation can be implemented to use simple direct detection (DD), for a cost-effective BVT design suitable for metro/regional networks [6]. In this case the coherent receiver stage of Fig. 3 is replaced with a single photodiode. On the other hand, coherent O-OFDM makes full use of the optoelectronic front-ends described in Fig. 3. By combining coherent detection and DSP, the tolerance to transmission impairments can be significantly enhanced, achieving ultimate performance.

• TFP is another transmission technique that can be implemented in the **flex subcarrier** module. TFP consists of sending pulses that strongly overlap in time or frequency or both to maximize spectral efficiency, while introducing ISI and/or inter-carrier interference (ICI) [7]. Coding and detection are properly designed to account for the introduced interference. A low-density parity check (LDPC) code can be used to approach the maximum information rate achievable with the given modulation (typically PM-QPSK), accounting for the presence of noise, ISI, ICI, and so on. Such code is introduced by the Encoding stage in the flex subcarrier module. Code rate, and thus spectrum efficiency, may vary with the optical signal-to-noise ratio (OSNR) of the subcarrier (the lower the OSNR, the larger the redundancy). The **pulse shape** module in Fig. 3 is used to filter data beyond Nyquist. The receiver of each subcarrier exploits coherent detection with DSP. After filtering, and clock and carrier recovery, the **equalizer** module can be implemented with a two-dimensional adaptive feed-forward equalizer compensating for linear propagation impairments (e.g., dispersion). Given the introduced ISI, TFP requires a receiver based on sequence detection, such as the well-known Bahl-Cocke-Jelinek-Raviv (BCJR) detector [8], which exchanges information with an LDPC decoder in the **decision**, **demapping**, **and decoding module**. Regarding BCJR complexity, a four-stage BCJR detector has been demonstrated with TFP (e.g., on a 5000 km path) in [7]. Its complexity is comparable with that of the two dimensional adaptive feed-forward equalizer. As an example, a multiflow optical module supporting 1.12 Tb/s may be composed by seven flex subcarrier modules, each generating a PM-QPSK subcarrier at 160 Gb/s. TFP is exploited to reduce the bandwidth of each generated subcarrier and their frequency separation below the Nyquist limit, finally having 160 Gb/s PM-QPSK in 28 GHz [7]. TFP provides high spectral efficiency (e.g., 5.16 b/s/Hz with a code rate of 8/9, meaning 8 bits of information each 9 bits, along a 3000 km path). The required all-optical reach can be achieved through the selection of proper code rate, and in this case, a flex subcarrier module supporting multiple modulation formats could be avoided.

A Discussion on Components and Integration

To cope with future networks' flexibility requirements (e.g., rate and reach), S-BVTs must transmit advanced modulation formats at high symbol rates meeting high spectral efficiency, and low costs and power consumption, by adaptively looking for a trade-off [9]. The key components/parameters driving costs and power consumption mainly constitute the multiflow optical module: ADC, DAC, I/Q modulators, the electronic processing (e.g., for filtering or equalization) in the flex subcarrier module shown in Fig. 3, and the laser sources of the subcarrier generator module.

A related component technology challenge is integration, which provides cost reduction as well as smaller physical dimensions. The S-BVT design can be enhanced if all the above modules (e.g., laser sources of the subcarrier generator and IQ modulators of the flex subcarrier) could be integrated on a single platform. High integration levels also permit better monitoring, management, and control of system performance. The enabling integration technologies could be based on silicon photonics, which allows for a well adapted matching between the electronic part (e.g., DAC, driver amplifier, ADC, DSP, and FEC) and the optical section (e.g., modulators, photodiode, laser source). Energy efficiency of an integrated system is a further benefit. Silicon photonics integration can help, for example, share thermal control and power dissipation functions among a subset of subchannels, but with the current hybrid approach the contribution to total power saving would be limited to about 10 percent. To achieve significant power reduction, a more promising technology involves complementary metal oxide semiconductor (CMOS) and photonics integration, which will evolve in the next five years with a level of optoelectronic integration lower than a 130 nm silicon-on-insulator (SOI) CMOS process (e.g., a single lithographic process can integrate hundreds of photonic components with millions of transistors with significant power reduction benefits) [10]. For instance, based on CMOS pho-

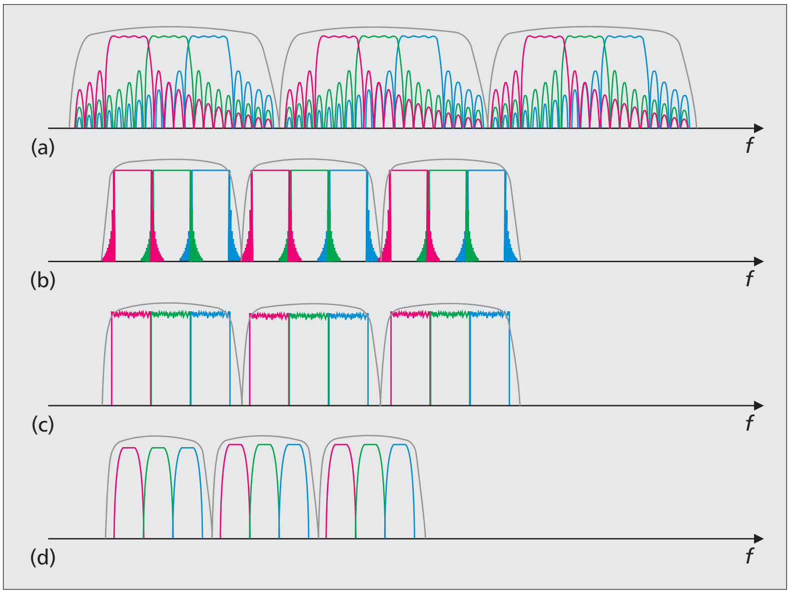


Figure 4. Illustrative examples of spectrum for OFDM with a) 4 subcarriers and b) 256 subcarriers; c) NWDM; d) TFP.

	Maximum spectral efficiency[1]	All-optical reach	Cost and complexity		Use cases
			TX	RX	
Nyquist WDM	Dependent on the modulation format; channel spacing ≥ symbol rate: e.g., 4 b/s/Hz for PM-QPSK and 8 b/s/Hz for PM-16-QAM	Dependent on the modulation format (several thousands km for PM-QPSK and PM-16-QAM, less for higher formats)	Mainly driven by DAC (e.g., electronic bandwidth ≥ 0.5 × symbol rate)	Mainly driven by ADC and DSP (e.g., electronic bandwidth ≥ 0.5 × symbol rate). Sampling rate ≥ symbol rate	Core/long-haul Metro/regional
Time frequency packing (with PM-QPSK)	Channel spacing can be smaller than the symbol rate: e.g., 8 b/s/Hz [14]	Thousands of km with variable spectrum efficiency: e.g. 3000 km (with 5.16b/s/Hz) 5500 km (with 4.2 bit/s/Hz) [15]	DAC is not required by assuming PM-QPSK and avoiding digital pre-emphasis or pre-distortion as in [15]	Mainly driven by ADC and DSP (e.g., electronic bandwidth < 0.5 × symbol rate). Moreover, a sequence detector is required instead of a symbol-by-symbol detector. Sample rate: one sample per symbol	Core/long-haul Metro/regional
O-OFDM	Dependent on the modulation format (for > 64 subcarriers); channel spacing is at least equal to the symbol rate: e.g., 4 b/s/Hz in case of PM-4-QAM and 8 b/s/Hz in case of PM-16-QAM	Dependent on the modulation format and detection scheme (several thousands km for PM-4-QAM and PM-16-QAM and CO-OFDM, less for higher formats or cost-effective metro solutions)	Mainly driven by DAC and DSP [5] (e.g., inverse Fourier transform processing, oversampling-sampling rate > symbol rate)	Mainly driven by ADC and DSP (e.g., electronic bandwidth ≈ 0.5 × symbol rate). Sample rate: one sample per symbol	Core/long-haul Metro/regional

[1] Spectral efficiency (SE) is defined as the information rate over a given bandwidth. SE can vary depending on BL, FEC, and front-ends adopted.

Table 1. Transmission techniques and characteristics.

tonics, a multiflow optical module can be redesigned using a single chipset, equipped with C-band tuneable lasers (e.g., based on a distributed Bragg reflector), optical Mach Zehnder modulators (e.g., based on Indium Phosphide), driven by an integrated electrical driver amplifier and photodiode, followed by a trans-impedance amplifier, and so on. All parts can be integrated into a thermally and power efficient package (i.e., the transceiver is intended as a pluggable module) and electrical interface to connect the transceiver on the S-BVT. In this way an S-BVT could transmit a 1 Tb/s super-channel, consuming the same system power as current 100 Gb/s transponders.

Finally, we discuss the subcarrier generation module. Two solutions can be adopted for subcarrier generation: an array of lasers or a multiwavelength source (MW) [11] (i.e., a source able to generate several subcarriers from a single laser). With the former, if tuneable lasers in the C-band are used, the spectrum allocation of subcarriers does not present any limitation (i.e., any available portion of the spectrum can be used). With the latter, other advantages may be achieved (implementation based on a tuneable laser and modulator, as in [11], is considered). First, costs and power consumption can be reduced given that on N generated subcarriers, $N - 1$ lasers are saved. Moreover, subcarriers generated by an MW are intrinsically spectrally locked together. In contrast, subcarriers generated by an array of lasers may present instability (inducing degradations from subcarrier overlap). As a drawback, MW may present limitations when allocating subcarriers in the spectrum. Indeed, MW generates subcarriers that are symmetrically spaced. Thus, further research is needed in the field of subcarrier generation.

S-BVT Programmability Perspectives

This section discusses the programmability of the proposed S-BVT, with particular reference to the modules/devices that need to be controlled in order to set/program specific transmission characteristics. Indeed, an S-BVT finds application whenever transmission characteristics can be set based on the actual traffic demands: by expanding or contracting the bandwidth of an optical path (e.g., varying the number of subcarriers), adapting the optical reach, and directing the generated super-channels toward specific destinations. To achieve these functionalities, several subcarriers can be connected or disconnected, and the modulation format or code-rate can be modified based on the required optical reach. The following S-BVT transmission characteristics are considered:

- Association of an OTN stream with a specific subcarrier (or flex subcarrier module)
- Line rate (e.g., number of subcarriers)
- Optical reach adaptation
- Modulation format
- Code rate
- Optical carrier frequency
- Specific functionalities at the DSP (e.g., equalization)

As shown in Fig. 2, OTN streams have to be associated with a media channel, which is directed to a specific destination. The media channel is achieved by several sub-carriers, each one obtained by a flex subcarrier module. The association of OTN streams with a set of optical subcarriers is obtained by an electronic switch implementing the flow distributor shown in Fig. 1, enabling remote control of its cross-connections. Line rate can be set by activating or not a subcarrier, thus controlling the activation of a flex subcarrier module. Optical reach adaptation can be obtained by relying on a proper modulation format achieved by the flex subcarrier module. By controlling the DAC in Fig. 3, it is possible to set the modulation format: for example, multi-level data from the DAC is used to achieve PM-16-QAM, while single-level data is used for PM-QPSK. Optical reach adaptation can also be achieved with code adaptation as in [7]. The code rate can be electronically programmed via software by the encoder in the flex subcarrier module (Fig. 3). Optical carrier frequency can be set by configuring the array of lasers or the MW source in the subcarrier generation module. Specific functionalities at the DSP can be configured via software (e.g., pulse shape, filtering). Finally, the node has to be configured to direct optical flows to specific paths. This is done by programming the node's spectrum selective switches (SSSs) to switch optical flows toward the correct node outputs. In addition, if the S-BVT is integrated within an architecture-on-demand (AoD)-based optical node [12], the entire node system requires further programmability [13] and ability to modify its configuration (e.g., for scalability, resiliency, multidimensional adaptability, and node/network synthesis).

The sub-wavelength granularity of an OFDM-based S-BVT allows the programmability via DSP to be extended to the electrical subcarrier level, including arbitrary subcarrier suppression, adaptive bit loading (BL), and power loading (PL). Bit loading consists of independently loading the subcarriers with symbols mapped with different modulation formats. Thus, fine bit rate selection and efficient use of the spectrum can be achieved. Additionally, in order to optimize system performance, subcarriers with lower signal-to-noise ratio can be loaded with data mapped with the most robust modulation format. Power loading can also be implemented at the S-BVT DSP, in which each subcarrier or a set of subcarriers is multiplied by a gain coefficient to adaptively vary the subcarrier power. This introduces additional flexibility to the system, enhancing overall performance. Furthermore, due to the overhead from pilot tones, training sequences, and cyclic prefix, the programmable OFDM-based S-BVT intrinsically provides self-performance monitoring.

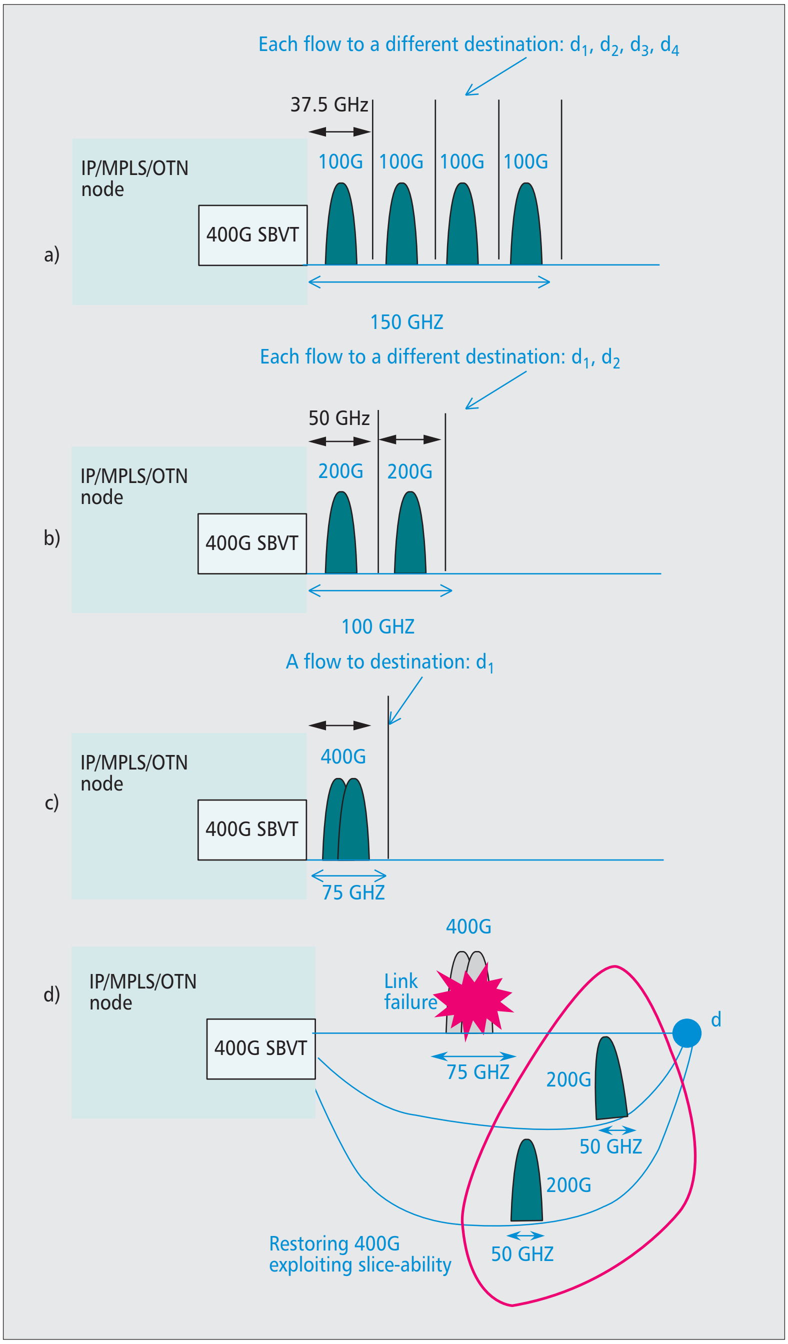


Figure 5. a)–c) Flexible utilization of an S-BVT by varying the bit rate per destination; d) sliceability used in restoration.

Based on these new paradigms, S-BVT plays an important role in the operation of next-generation optical networks by supporting the on-demand configuration of programmable network functions such as rate, switching, and sliceability.

Scenarios of Application

Some ways to make use of S-BVT functionality include:

- Dynamicity of traffic
- Migration toward higher traffic rates
- Restoration

Directionless ROADMs can redirect circuits, allowing restoration whilst minimising additional transponders. S-BVTs can improve this further as their output can be sliced into multiple restoration circuits, enabling restoration bandwidth to multiple nodes affected by the failure.

If, as well as bandwidth growth, dynamics and direction change over hours or days (an example is the traffic generated by data centers), an optical layer that can respond to these variations will be more cost effective. S-BVTs are well suited for this application as they permit super-channel circuits to be maintained but with total bandwidth dynamically changed (e.g., by activating a variable number of flex subcarrier modules or varying the baud rate). Moreover, the sliceability of S-BVT allows serving a requested rate on a given destination on demand.

Figure 5a–c shows the utilization of S-BVT as the network migrates toward high rates. In year one, Fig. 5a, a single 400 Gb/s S-BVT serves four different destinations (d_1, d_2, d_3, and d_4), each with 100 Gb/s connectivity; that is, a flex subcarrier module is activated for a specific destination. In Fig. 5b, the same S-BVT can fulfill a possible traffic rate increase per destination; thus, two optical flows are reconfigured (through automatic operations carried out by the control plane) to satisfy new demands (for d_1 and d_2). Finally, in Fig. 5c, the whole capacity of the same S-BVT is used for a single destination; thus, all the flex subcarrier modules' flows are directed to one destination. This example highlights the flexible use of an S-BVT, able to vary the rate and direction of the optical flows as the network migrates toward high rates.

Multi-layer restoration is an interesting application for S-BVT. Even without flex-grid, considerable savings can be obtained by reusing transponders during fiber failures. Directionless ROADMs can redirect circuits, allowing restoration while minimizing additional transponders. S-BVTs can improve this further as their output can be sliced into multiple restoration circuits, enabling restoration bandwidth to multiple nodes affected by the failure. Another use can be done to relax requirements on contiguous bandwidth, as in Fig. 5d. Consider a link failure disrupting a 400 Gb/s super-channel. The required overall super-channel bandwidth (75 GHz in the example) is not found along any path connecting the source-destination pair. However, lower bandwidth is available along several paths. In this case, S-BVT is used to partition the super-channel into two media channels directed along different paths, each requiring an amount of available bandwidth (50 GHz in the example), finally successfully restoring all of the traffic. As for the provisioning, during restoration, the required all-optical reach can be achieved through the selection of the proper modulation format (e.g., with proper setting of the DAC in Fig. 3) or code rate (by changing the code in the encoder/decoder in Fig. 3).

Conclusions

This article investigates the next generation optical transponders (S-BVT). An S-BVT architecture is shown including the client side (OTN layer) and the signal generation, providing general architecture modules. The association between OTN streams and media channels is illustrated, highlighting the capability of OTN to fragment high-rate clients into lower-rate streams for effective mapping in the optical layer. Several transmission techniques (NWDM, TFP, and OFDM) are supported by the S-BVT architecture, showing the achievable spectral efficiency and the required hardware. Discussions are provided on components (further studies are required for multi-wavelength sources) and integration challenges to reduce costs, power consumption, and space. Remote control has been discussed, as well as use cases. S-BVTs can be used to cope with traffic dynamicity increase and will find application in the migration toward high rates. Sliceability of a super-channel can also be effective during restoration by reducing the amount of required contiguous bandwidth.

Acknowledgment

This work was supported by the FP-7 IDEALIST project under grant agreement number 317999.

References

[1] M. Jinno *et al.*, "Distance-Adaptive Spectrum Resource Allocation in spectrum-Sliced Elastic Optical Path Network," *IEEE Commun. Mag.*, vol. 48, no. 8, Aug. 2010, pp.138–45.

[2] M. Jinno *et al.*, "Multiflow Optical Transponder for Efficient Multilayer Optical Networking," *IEEE Commun. Mag.*, vol. 50, no. 5, May 2012, pp. 56–65.

[3] M. Jinno *et al.*, "Introducing Elasticity and Adaptation into the Optical Domain Toward More Efficient and Scalable Optical Transport Networks," *ITU-T Kaleidoscope: Beyond the Internet? — Innovations for Future Networks and Services*, 13–15 Dec. 2010, pp. 1, 7.

[4] J. K. Fischer *et al.*, "Bandwidth-Variable Transceivers based on Four-Dimensional Modulation Formats," accepted for publication, *IEEE/OSA JLT*, invited.

[5] R. Schmogrow *et al.*, "Real-Time Digital Nyquist-WDM and OFDM Generation: Spectral Efficiency vs. DSP Complexity," *Proc. ECOC '12*, paper Mo.2.A.4.

[6] M. Svaluto Moreolo *et al.*, "Experimental Validation of an Elastic Low-Complex OFDM-based BVT for Flexi-Grid Metro Networks," *Proc. ECOC '13*, paper We.C1.4.

[7] N. Sambo *et al.*, "Programmable Transponder, Code and Differentiated Filter Configuration in Elastic Optical Networks," *JLT*, vol. 32, no. 11, June 2014, pp. 2079–86.

[8] L. Bahl *et al.*, "Optimal Decoding of Linear Codes for Minimizing Symbol Error Rate," *IEEE Trans. Info. Theory*, 1974.

[9] D. Rafique *et al.*, "Technology Options for 400Gb/s PM-16QAM Flex-Grid Network Upgrades," IEEE PTL 2014.

[10] D. Guckenberger *et al.*, "Advantages of CMOS Photonics for Future Transceiver Applications," *ECOC '10*, Torino, Italy, paper TU.4.C.2.

[11] A. Mishra *et al.*, "Flexible RF-based Comb Generator," *IEEE Photonics Tech. Lett.*, vol. 25, no. 7, 2013, pp. 701–04.

[12] N. Amaya *et al.*, "Introducing Node Architecture Flexibility for Elastic Optical Networks," *J. Opt. Commun. Net.*, vol. 5, 2013, pp. 593–608.

[13] Y. Yan *et al.*, "FPGA-based Optical Network Function Programmable Node," OFC '14, San Francisco, CA, 2014.

[14] G. Colavolpe and T. Foggi, "High Spectral Efficiency for Long-Haul Optical Links: Time-frequency packing vs High-Order Constellations," *ECOC '13)*, 22–26 Sept. 2013, pp.1–3.

[15] L. Potì *et al.*, "Casting 1 Tb/s DP-QPSK Communication into 200 GHz Bandwidth," *ECOC '12*, 16–20 Sept. 2012, pp. 1–3.

Biographies

Nicola Sambo (n.sambo@sssup.it) received his Laurea degree cum laude in telecommunication engineering from the University of Pisa, Italy, in 2004, and his Ph.D. degree from Scuola Superiore Sant'Anna (SSSUP), Pisa, Italy, in 2009. In 2007, he was a visiting student at France Télécom, Lannion. Currently, he is an assistant professor at SSSUP. He is also collaborating with CNIT, Pisa, Italy, and Ericsson. His research interests include optical network architecture, transmission performance modeling, and control plane (GMPLS, PCE, SDN). He is an author of about 80 publications including international journals, conference proceedings, and patents.

PIERO CASTOLDI [SM] has been an associate professor in telecommunications responsible for the Networks and Services area at SSSUP since 2001. His most recent research interests cover flexible optical networks, optical interconnection networks for data centers, software defined network architectures, and networks for industrial applications. He is an author of more than 300 technical papers published in international journals and international conference proceedings in addition to the sole author of a book on CDMA. He has also filed a total of 20 patents.

ANTONIO D'ERRICO received his Ph.D. degree cum laude in telecommunication engineering from SSSUP in 2008. He is currently with Ericsson Research as an experienced researcher. His research interests include advanced technological solutions for optical networks. He is an author of more than 75 papers published in international journals, conference digests, and patents. He is on the Boards of Reviewers for a number of international journals in the fields of optics and photonics.

EMILIO RICCARDI, after graduation in physics, joined Telecom Italia Lab (formerly CSELT) in 1992. His current interest is dedicated to elastic optical networks. He has been involved in several EU Projects and presently leads the Data Plane Working Package of the IDEALIST EU project. He has coauthored more than 40 journal and conference publications.

ANNACHIARA PAGANO received her degree in physics from the University of Turin in 1994. In the same year she joined Telecom Italia (formerly CSELT) to work on optical fiber optics and DWDM metro and long distance communication systems. Her current fields of interest are next-generation optical networks, focusing on both high-capacity network design and testing, together with research activities on optical transmission and scouting of new technologies. She has co-authored more than 50 technical publications covering design and testing aspects of optical components and systems. She has been involved in several European Projects.

MICHELA SVALUTO Moreolo [S'04, A'08, SM'13] received her M.Sc. degree (with honors) in electronics engineering and Ph.D. degree in telecommunications engineering from University Roma Tre, Italy, in 2003 and 2007, respectively. In 2004, she became a teaching assistant in the Engineering Faculty of University Roma Tre and, from January to July 2004, she was a visiting researcher at the Institute of Semiconductor and Solid State Physics, Johannes Kepler University, Linz, Austria. In January 2009, she joined the Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), Castelldefels (Barcelona), Spain. She currently is a senior researcher in the Communication Networks Division and coordinator of the Optical Transmission and Subsystems research line within the Optical Networks and Systems Department. Her research interests include optical/digital signal processing and advanced transmission technologies for future optical networks. She has authored and co-authored more than 70 journal and international conference papers.

JOSEP M. FABREGA [M] received his Ph.D. degree in signal theory and communications from UPC-BarcelonaTech, Spain, in 2010. He is with the Optical Networks and Systems Department at CTTC. During 2004–2010, he was with the Optical Communications Group of UPC-BarcelonaTech. He is the author/coauthor of +60 papers. His research interests include advanced modulation formats. He is member of the IEEE Photonics Society. He received the EuroFOS student research prize in 2010.

DANISH RAFIQUE holds B.Sc. (2007, FAST-NU, Pakistan), M.Sc. (2009, KTH, Sweden), and Ph.D. (2012, Tyndall National Institute, University College Cork, Ireland) degrees, all in electrical engineering. He joined Nokia Siemens Networks in 2012, and is currently pursuing a career with the R&D Labs of Coriant GmbH (former NSN), Germany, where he is working on product development and research programs on next-generation transponder solutions involving flexgrid networks. His interests include nonlinear effects in fiber, optical system design trade-offs, digital signal processing, and dynamic optical networks. He has published more than 55 articles, including peer-reviewed journals, contributed and invited conference publications, and book chapters, and acts as a reviewer for *IEEE Photonics Technology Letters*, *IEEE Journal of Lightwave Technology*, *OSA Optics Express*, and so on.

ANTONIO NAPOLI (antonio.napoli@coriant.com) received M.Sc. and Ph.D. degrees in electrical engineering from Politecnico di Torino in 2002 and 2006, respectively. In 2004 he was a visiting student at UPC, Barcelona and in 2005 at University College London. In 2006, he joined Siemens COM and in 2007 NSN, working on DSP for high-speed optical networks. Since 2013 he is with Coriant. His main interests are DSP and advanced modulation formats for long-haul optical systems. He has (co)-authored 80+ peer-reviewed articles and seven patents

SILVANO FRIGERIO is a member of technical staff within the Optics Product Unit in Vimercate, Italy. He received a degree in electronic engineering at the Politecnico of Milan, Italy. He brings extensive and sustained experience developed as a system architect of the Optical Multi-Service Nodes (OMSN) and as a SONET/SDH ASIC designer. Presently he is focusing on aspects regarding multi-technology solutions and OTN transformation at both the equipment and network levels. He is an Alcatel-Lucent Italia TCT Principal Engineer and holds several patents concerning transmission networks.

EMILIO HUGUES-SALAS (e.huguessalas@bristol.ac.uk) received his M.Sc. and Ph.D. degrees from the University of Essex, United Kingdom. He is a research fellow in the High Performance Networks Group at the University of Bristol. His research interests include the development of high-capacity, reconfigurable, flexible, and multidimensional optical nodes and networks. He is an author and co-author of over 60 international journals and conference papers.

GEORGIOS ZERVAS holds M.Eng. and Ph.D. degrees from the University of Essex. He is a lecturer in optical and high-performance networks at the University of Bristol, United Kingdom, and a member of HPN group. He leads the activities in optical transport and switching, as well as software/hardware programmable systems and networks for backbone and DCN/HPC interconnects. He has published over 150 journal and conference papers including postdeadline papers and has received best-paper awards. He has been a TPC member of numerous leading conferences and a contributor to standards.

MARKUS NOLLE received Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from the Technische Universität Berlin in 2007 and 2013, respectively. Since 2008, he has been with the Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute (FhG-HHI) in Berlin, Germany. Currently he is a researcher and project manager in the Submarine and Core Systems group within the Department of Photonics Networks and Systems. His research interests are spectrally efficient higher-order optical modulation formats and digital distortion mitigation in coherent optical systems.

JOHANNES KARL FISCHER [S'04, M'09] received Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from the Technische Universität Berlin in 2003 and 2009. He is with FhG-HHI, where he heads a research group on digital signal processing. His research interests are advanced modulation formats, digital signal processing for coherent systems. He is a member of the Verband der Elektrotechnik Elektronik Informationstechnik (VDE).

ANDREW LORD joined BT in 1985 after a B.A. in physics from Oxford University. He has worked on a wide range of optical network systems and technologies, including long haul subsea and terrestrial DWDM networks. He currently heads BT's optical core and access research. He has had many years of European Project coordination and currently helps lead the Idealist FP7 project. He regularly speaks at conferences, sits on several organizing committees, including ECOC, and is Technical Program Chair for OFC 2015. He is an Associate Editor of the *Journal of Optical Communications and Networking* and a visiting professor at Essex University.

JUAN PEDRO FERNÁNDEZ-PALACIOS GIMÉNEZ received his M.S. in telecommunications engineering from the Polytechnic University of Valencia in 2000. In September 2000 he joined TIDSA, where he is currently leading the Metro and Core Transport Networks Department. He has been involved in several European projects such as STRONGEST and MAINS as well as other internal projects related to the development of optical networks in the Telefónica Group. Currently he is coordinating the FP7 project IDEALIST.

Next Generation Optical Nodes: The Vision of the European Research Project IDEALIST

Emilio Hugues-Salas, Georgios Zervas, Dimitra Simeonidou, Evangelos Kosmatos, Theofanis Orphanoudakis, Alexandros Stavdas, Marc Bohn, Antonio Napoli, Talha Rahman, Filippo Cugini, Nicola Sambo, Silvano Frigerio, Antonio D'Errico, Annachiara Pagano, Emilio Riccardi, Victor López, and Juan Pedro Fernández-Palacios Giménez

Emilio Hugues-Salas, Georgios Zervas, and Dimitra Simeonidou are with the University of Bristol.

Evangelos Kosmatos, Theofanis Orphanoudakis, and Alexandros Stavdas are with the University of Peloponnese.

Marc Bohn and Antonio Napoli are with Coriant R&D GmbH.

Talha Rahman is with Eindhoven University of Technology.

Filippo Cugini and Nicola Sambo are with CNIT.

Silvano Frigerio is with Alcatel-Lucent Italia.

Antonio D'Errico is with Ericsson Telecomunicazioni SpA.

Annachiara Pagano and Emilio Riccardi are with Telecom Italia.

Victor López and Juan Pedro Fernández-Palacios Giménez are with Telefonica Innovacion y Desarrollo.

ABSTRACT

As traffic demands become more uncertain and newer services continuously arise, novel network elements are needed to provide more flexibility, scalability, resilience and adaptability to today's optical networks. Considering these requirements, within the European project IDEALIST the investigation of elastic optical networks is undertaken with special focus on next generation optical node architectures. As an evolution of existent ROADMs and OXCs, these optical nodes will establish a new paradigm in which the network requirements will be efficiently addressed considering various emerging dimensions. In this article, we describe the drivers, architectures, and technologies that will enable these novel optical nodes. In addition, multivendor traffic interoperability, optical defragmentation, and node cascadability are also described as considerations in the node design.

INTRODUCTION

Current optical transport network infrastructure is supported by nodes with reconfigurable optical add/drop multiplexers (ROADMs). ROADMs provide the network operator with a certain level of flexibility by dynamically reconfiguring optical channels (i.e., wavelengths), without the need for precise knowledge of the traffic growth, enabling remote provisioning and configuration. In addition, optical designs with ROADMs are further simplified by the use of wavelength selective switches (WSSs), which replace multiplexers, demultiplexers, and switching devices by providing filtering and switching functions. These WSS-based ROADMs may have a degree higher than the current two, making them suitable for implementation in mesh networks. Thus, these ROADMs are considered as network elements with multiple-node interconnectivity (i.e., node degree) allowing the transport of optical channels to different destinations within the optical network. However, ROADM architectures offer limited flexibility if selected traffic flows passing through the node demand some additional functionality. On the other hand, if additional functionalities, such as regeneration, are needed only for a few channels, and others, such as compensation, are required for an entire set of lambdas (e.g., for pass-through traffic in a specific direction), agnostic optical cross-connects (OXCs) are preferable because OXCs are able to connect modules with the required functionality efficiently on both single channel units and aggregated channel spectra. Nevertheless, the requirement for a particular signal processing function is often uncertain (e.g., it may be required for some wavelengths at some point in time and for other wavelengths at different moments). As a consequence, the node resources are not optimally used, which is reflected in higher costs of each network element.

In addition to the aforementioned limitations, optical nodes face new challenges as more sophisticated modulation schemes are foreseen to cope with the increased traffic demand. In this context, new types of network elements need to be considered to allow flexible spectra routing in elastic optical networks (EONs) [1]. EONs allow for flexible allocation of an amount of spectrum in order to match individual traffic demands and deliver higher transmission capacities than fixed-grid wavelength-division multiplexing (WDM) networks. For instance, the International Telecommunication Union — Telecommunication Standardization Sector (ITU-T) has foreseen the requirement for a spectrum allocation scheme considering a dense WDM (DWDM) grid with 12.5 GHz frequency slot granularity and 6.25 GHz central frequency granularity that provides more flexibility than a conventional grid. EONs also consider superchannels (> 100 Gb/s) with different subcarrier characteristics, such as spectrum width, modulation format, bit rate, and number of subcarriers. To enhance EONs, more flexibility is required within these optical nodes, and new flexible

 0163-6804/15/$25.00

OXC (Flex-OXC) nodes need to be deployed. In addition, it is necessary to include in the optical nodes features that will reduce their cost and complexity as well as the deployment and operational costs of the entire network.

Initially, elastic node architectures can be obtained by performing a simple upgrade of components. To support elastic spectrum allocation, static ROADM node architectures have been proposed by replacing the fixed-grid WSS with a bandwidth-variable WSS (BV-WSS) [2]. These suggested EON node architectures suffer from scalability limited to a few degrees, complex spectrum fragmentation, and the need for additional components, such as large port-count switches, to provide extra functionalities (time multiplexing, regeneration, etc).

Thus, considering these emerging requirements, and addressing the issues presented above, within the European project IDEALIST (Industry-Driven Elastic and Adaptive Lambda Infrastructure for Service and Transport networks) novel node architectures and subsystems are explored and developed.

This article consists of four parts. The first part presents the new requirements for next generation optical nodes. Subsequently, a detailed description of the architecture-on-demand (AoD)-based Flex-OXC is reported, including static and dynamic optical modules. Hereafter, different interfaces for these Flex-OXC nodes are proposed considering various networking environments. Finally, operation and maintenance issues are addressed in order to contemplate them in the design of optical networks based on novel Flex-OXC nodes.

Requirements for Next Generation Optical Nodes

New requirements need to be addressed in future high-capacity flexi-grid optical nodes. One of these is *node flexibility*. Next generation optical nodes will incorporate high levels of flexibility in different domains, such as switching, bandwidth, and transmission rate. The introduced flexibility would allow dynamic composition of network functions to cope with different generated traffic demands. Optical *node scalability* is another important requisite supported by a component's modularity to achieve a practical approach and enable future services to easily be added to the network functionality.

Node resilience is also accounted for in EONs. Node failures due to faulty subsystems should be efficiently addressed and overcome, guaranteeing continuously rapid provision of services to network users. Generally speaking, Flex-OXC nodes should incorporate redundant components to enhance the node reliability and survivability. In this context, minimum outage times should be achieved by the node for further reduction of operational costs.

Finally, simple *node adaptability* to new network applications is also preferred for these emerging optical nodes. Figure 1 describes the IDEALIST approach to developing the proposed Flex-OXC node architecture, allowing different configurations of a number of functional components (i.e., adaptability). In here, an AoD-based Flex-OXC node is proposed, including a switchless elastic rate node (SERANO) module, which is an alternative to wavelength switching architectures.

Examples of node adaptability are shown in Fig. 1. For instance, if the AoD-based Flex-OXC node is considered within an optical transport network (OTN), network traffic demands are directed through a bandwidth variable transponder (BVT) to the optical node backplane, which in turn can lead this traffic to a SERANO module (use case 1). On the other hand, if an Ethernet infrastructure is considered (use case 2), an elastic BVT interface will interconnect demands, coming from Ethernet switches, to an optical IP elastic layer with a router equipped with BVT optical interfaces through the optical backplane. Moreover, this elastic Ethernet and/or IP layer could be connected directly to a geographical metropolitan area network (MAN) by means of a BV-WSS.

Next generation optical nodes will incorporate high levels of flexibility in different domains, such as switching, bandwidth and transmission rate. The introduced flexibility would allow dynamic composition of network functions to cope with different generated traffic demands.

AoD-Based Flex-OXC Node Architecture

An AoD-based Flex-OXC node consists of an optical backplane, such as a beam-steering switch with as low as 0.5 dB cross-connection, and several signal processing modules including BV-WSS, Erbium-doped fiber amplifiers (EDFAs), SERANO modules, and the node's inputs and outputs attached to the optical backplane. Based on these components, different node arrangements can be constructed by setting up appropriate cross-connections in the optical backplane (Fig. 2a). This combination of various optical modules is achieved gradually, with an initial component synthesis phase considered where a node functionality is required, and subsequent phases are included if the traffic demands change or an additional node functionality is needed. This means the AoD node does not need to be assembled with all the optical components during the first phase, avoiding hardware overprovisioning of the node with extra and unused components. In addition, compared to static node architectures [2], the AoD-based Flex-OXC node provides greater flexibility as the components used for optical processing are not hard-wired, allowing components to be interconnected in an arbitrary manner.

Benefits of the AoD-Based Flex-OXC Node

The aim of the AoD-based Flex-OXC node is to be used in real networks where flexible designs for optical nodes enable additional functionalities. To this extent, a convergence platform of fixed/flex-grid technologies was studied over Telefónica's Madrid network topology [3], and an on(off)-chip field programmable gate array (FPGA) (AoD)-based design was proposed for intra-data centers [4]. Also, analyses of the AoD-based node over the NSF and COST239 network topologies confirm considerable modules savings of up to 40 percent and a 25 percent decrease of network power consumption [5].

The architecture-on-demand (AOD)-based Flex-OXC node brings beneficial features to the network, including scalability, synthetic architectures, self-healing capability, multidimensional adaptability, and bandwidth variation.

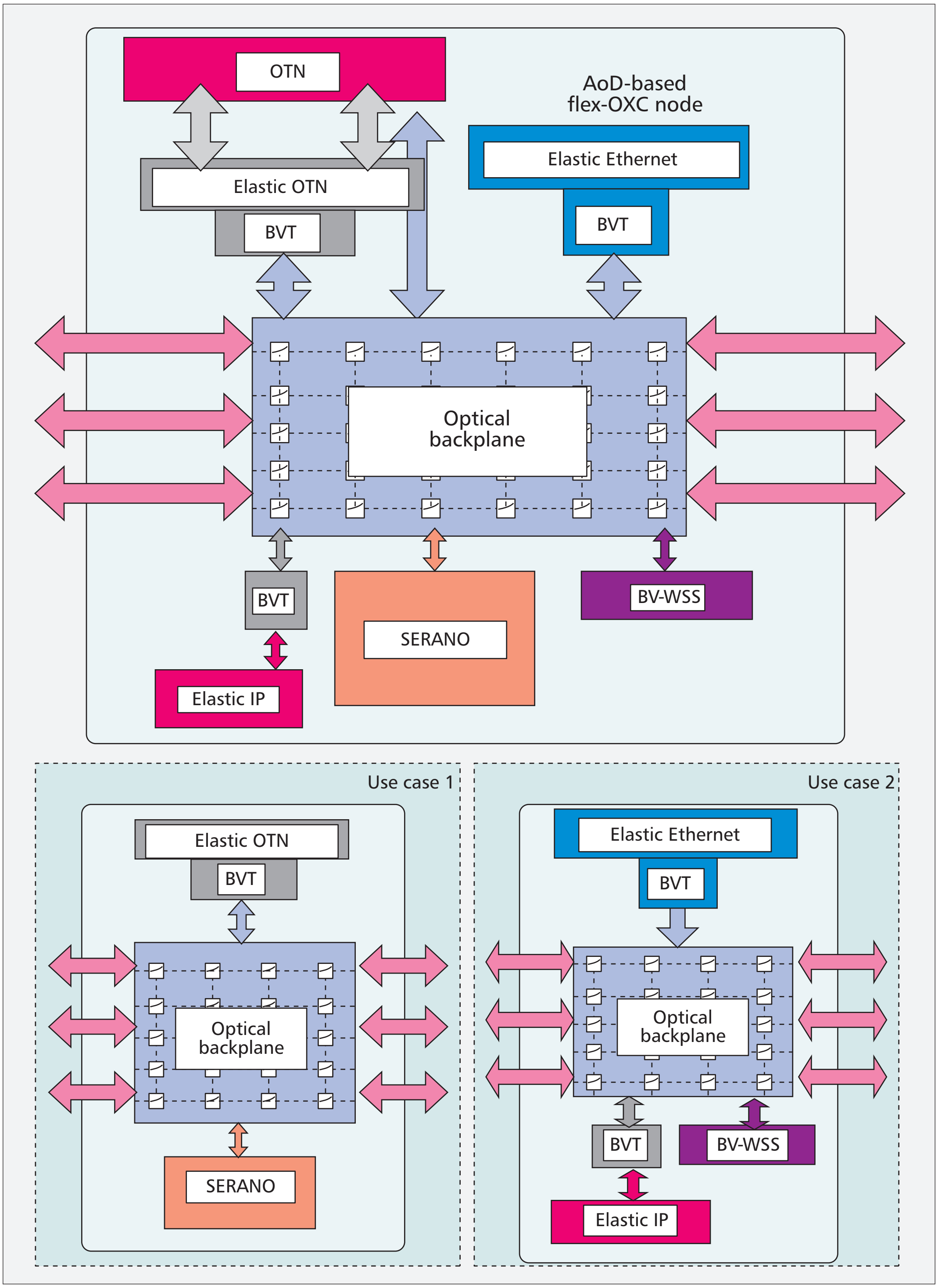


Figure 1. Flex-OXC node adaptability to different use cases.

Considering these network advantages, the AoD-based Flex-OXC node brings beneficial features to the network. These features are:

Enhanced scalability and synthetic architectures. Significant scalability is gained due to the AoD-based node's efficient multi-granular support [6]. Various levels of switching granularity are implemented by selectively introducing, in the synthetic AoD-based node architecture, modules that provide the required switching functionality, such as (de)multiplexing or WSS for DWDM, BV-WSS for flexible DWDM, and fast switch for time multiplexing. Thus, high scalability and module (cost) savings are achieved by switching traffic at the coarsest granularity (i.e., fiber and core switching) so that single backplane cross-connections are able to switch large volumes of traffic. In addition, the system is able to provide switching at finer granularities by inserting modules for extra functionality, which is not possible with static node architectures.

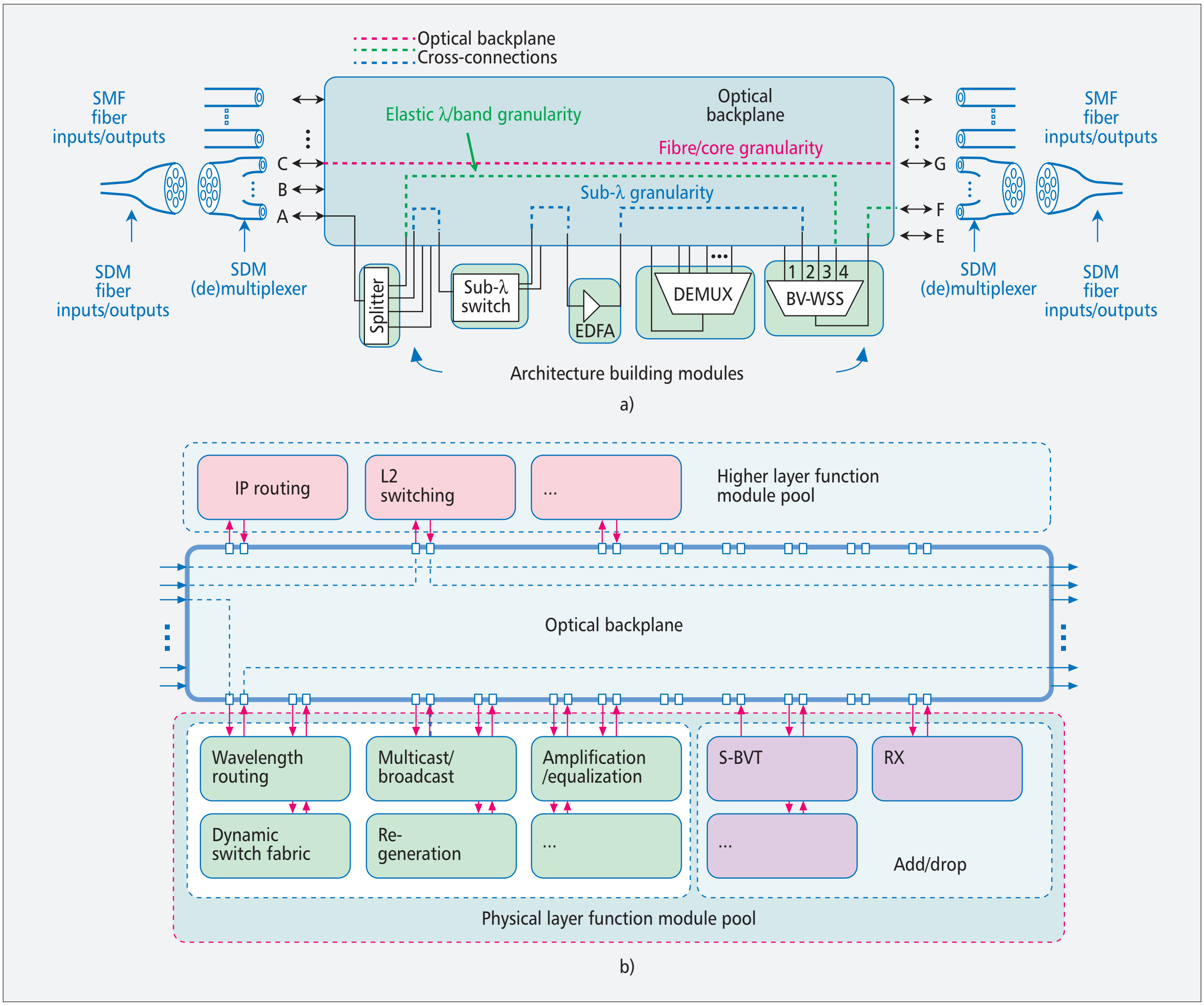


Figure 2. a) AoD-based Flex-OXC node; b) Flex-OXC node supported by NFP.

Self-healing capability. By reconfiguring the system to replace faulty modules, the Flex-OXC node can recover from a component failure, which is one of the advantages of AoD-based node over static node architectures [7]. Based on node subsystem redundancy, an AoD-based Flex-OXC node has demonstrated mean down times 15 times lower than those of a system without redundancy.

Multidimensional adaptability. The AoD-based Flex-OXC node has demonstrated adaptability to different transmission dimensions, able to process optical signals in space, frequency (sub-wavelength, super-wavelength), and time. More specifically, an AoD-based node has shown good end-to-end performance in space-division multiplexing (SDM) infrastructures (e.g., multi-core fiber), incorporating bandwidth-variable self-homodyne spatial super-channels in a network topology [8].

Enabling bandwidth virtualization. Multiple switching granularities supported by AoD-based nodes facilitate the implementation of various types of bandwidth virtualization. Virtual networks can be built by exploiting different domains. For instance, when the space domain is used, entire fibers or fiber cores are assigned to virtual networks. Fibers or cores of the same virtual network are interconnected by backplane cross-connections at the node's points of presence, according to the required virtual network topology. Alternatively, virtual networks can be implemented by interconnecting spectrum slices in a specific fiber/core.

In addition to the aforementioned salient features, AoD-based Flex-OXC can offer network function programmability (NFP) where a set of functions can be programmed utilizing the available resources (Fig. 2b). The functions in the NFP Flex-OXC node can be at higher layers (e.g., routing, switching) and/or the physical layer (e.g., amplification, regeneration) at the request of a network user who is unaware of the optical hardware resources available. Thus, the NFP Flex-OXC node approach will allow operators to access the technology they need for the type of services they provide, wherever they are intended to be deployed.

The functions provided by a SERANO module are changing a central carrier's frequency, spectral slot occupancy and modulation format modification, and tailoring the line rate to the length of the subsequent transparent optical section. In addition, it can guarantee multivendor interoperability.

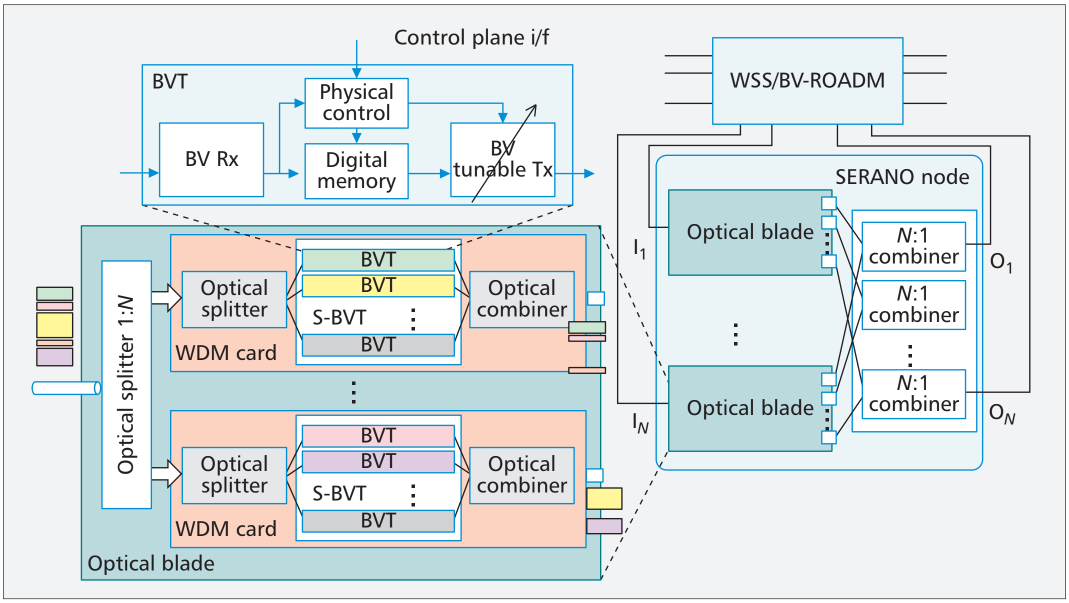


Figure 3. Overview of the SERANO architecture with simplified S-BVT modules embedded.

SERANO Module Architecture for Flex-OXC Nodes

SERANO extends the architectural framework described in [9, 10] to efficiently support AoD-based Flex-OXC nodes and EON operations. Following the multi-granular switching design approach in the context of flex-grid and elastic rate networks, it is assumed that a fraction of the traffic passing through the Flex-OXC node is added/dropped to/by the SERANO block by means of N I/O ports/fibers of a WSS/BV-ROADM, as illustrated in Fig. 3. Each input fiber is terminated in what we call an *optical blade*, which is shown in greater detail in the inset. The first stage in each blade is an optical splitter creating multiple copies of the input signal at all couplers' outputs. Each output is directly interfaced to a WDM card. A WDM card consists of a 1:M optical splitter (possibly in the form of zero-loss couplers) followed by an array of sliceable BVTs (S-BVTs). Such an S-BVT (inset, Fig. 3) consists of a multiflow optical module integrating a fixed-receiver tuneable-multiflow transmitter pair connected back to back, while a buffer is also provisioned for synchronization purposes. Therefore, the functionality of this block is to terminate the entire group of flows dropped from the Flex-OXC, where the selection of the desired channel is made by properly tuning the local oscillator in x phase shift keying (x-PSK)/x-quadrature amplitude modulation (x-QAM) schemes. Thus, the block provides for full 3R regeneration, possibly at a different central carrier frequency, only for selected flows (media channels with specific numbers of spectral slots and/or modulation format), that is, only those desired to be directed to the specific output fiber associated with the specific WDM card.

The WDM card also has a coupler, possibly amplified, that operates as a combiner with M input ports, M being the greatest number of flows a particular input fiber may forward to a particular output fiber. In the final stage of SERANO, the outputs of each WDM card are passively recombined by means of a second, possibly amplified, optical N:1 combiner.

The aforementioned SERANO architecture is flow and link modular, since only the couplers need to be provisioned, and the building blocks (blades and WDM cards) can be added progressively, as traffic increases, following a pay-as-you-grow principle. Thus, all the requirements for flexibility, scalability, resilience, and adaptability set forth in the previous section can be met. It is also worth noting that no optical or electronic switching technology is introduced since the main active building blocks are the multiflow optical arrays integrating transmission and data forwarding. This could lead to massive production of an optimized single element for both functions and lower cost. Moreover, the possibility to integrate couplers, passive or amplified, into these modules presents an interesting possibility for a compact, low-cost, and robust solution.

Conclusively, the SERANO module operates as a flow-forwarding engine in a service transparent/protocol agnostic manner, with 3R regeneration and on-the-fly *adaptation functions* obviating the use of large electronic switches. The functions provided by a SERANO module are changing a central carrier's frequency, spectral slot occupancy and modulation format modification, and tailoring line-rate to the length of the subsequent transparent optical section. In addition, it can guarantee multivendor interoperability, since the embedded multiflow optical receivers and transmitters can be purchased from different vendors.

An important optimization parameter for SERANO is the number of BVTs that are required for efficient operation. Toward this target we need to investigate the operational conditions under which the Flex-OXC operates. Flex-grid networks usually set up paths between two endpoints to serve specific traffic flow demands, such as implementing the path compu-

#	Scenarios												
						First optimization				Second optimization			
	N	M	C_{ch} (G)	Cap. (Tb/s)	Full-scale BVTs (x10³)	α	P_b (x10⁻³)	BVTs (x10³)	Savings (%)	α	P_b (x10⁻³)	BVTs (x10³)	Savings (%)
1	40	100	100	400	160	4.65	4	18.6	88.30	2.6	0.49	10.4	93.50
2	10	80	200	160	8	2.45	4.5	1.96	75.50	1.55	0.56	1.24	84.50
3	20	200	40	160	80	2.5	3.6	10	87.50	1.65	0.53	6.6	91.75
4	20	80	100	160	32	3.5	3.6	5.6	82.50	2.05	0.48	3.28	89.75
5	8	40	100	32	2.56	3	4.7	0.88	65.60	1.6	0.34	0.512	77.20

Table 1. Number of BVTs in different network scenarios.

tation element (PCE). In practical implementations these connections are pre-provisioned together with their protection paths [11]. A key observation in this case is that a given input fiber may forward flows only to a subset of output fibers. Therefore, a data plane architecture of the flow-forwarding engine may take advantage of this characteristic to attain savings in the required building blocks. In general, the total required number of BVTs to construct the SERANO block is

$$S_T = \left\lceil \frac{a \cdot M}{N} \right\rceil \cdot N^2$$

where M is the maximum number of supported flows per input fiber, while a is a resource overprovisioning factor with values in the interval [0, N] and is a design parameter. If $a = N$, the entire flow group of any input port can be forwarded to any output, and the total number of BVTs is equal to $M \times N^2$, which represents a full-scale SERANO implementation. However, not only is this operation better completed via the WSS/BV-ROADM section, but also the majority of BVTs will remain unused or underutilized, wasting considerable amounts of deployed resources. In addition, selecting a lower a value for the parameter might increase the chance of blocking, since fewer BVTs will be deployed in each WDM card. Thus, the design parameter a could be used to trade off the cost of WDM cards (by limiting the size, i.e., number of BVTs) for a limited set of interconnection patterns that could be implemented during each switch reconfiguration.

Therefore, a reduction in the number of BVTs can be achieved as a trade-off between a blocking probability P_b, which can be set to a controllable upper value, and the number of BVTs employed. For a given number of N and M, the value of a can be computed for any acceptable value of P_b based on mathematical analysis. To quantify these results, a number of practical scenarios are examined assuming blocking probability values below 3×10^{-3} under the simplifying assumption that each flow has equal probabilities of selecting an output fiber. Additionally, noticing that the flows entering the SERANO section are most likely directed to only a small number ($c < N$) of preselected output fibers (working fiber, second alternative such as protection, third, etc.) a second simplifying assumption is introduced in our analysis. The scenarios for a number of N, M, and C_{ch} (channel capacity) and the results of the probabilistic analysis are illustrated in Table 1 (selecting $c = 4$ for the second optimization). From Table 1 it is evident that optimization drastically reduced the number of BVTs up to 93.5 percent compared to the full -scale implementation of $M \times N^2$ BVTs. This is beneficial to network operators as it drastically decreases both capital and operational expenditures (CapEx/OpEx) by using only a small fraction of the maximum number of BVTs.

Interfaces for Future Node Architectures

A fundamental matter for operating in EONs is the design of appropriate node interfaces, which are discussed in detail in this section.

Comparison of State-of-the-Art Interfaces

The function of local traffic insertion and drop is currently performed by a WSS-based colorless photonic chain for fully non-blocking per wavelength access and branching. A cost-optimized, scalable, modular design is usually provided to achieve the lowest possible cost, minimizing operational expenditures.

Multiservice provisioning supporting all standard time-division multiplexing (TDM) services, plus L1/L2 Ethernet features is at the moment obtained through a multi-functional digital layer (encompassing adaptation, mapping, and electrical grooming and switching functionalities) partly incorporated in a dedicated electrical node and partly included in the optical interfaces' boards.

State-of-the-art optical interfaces identify different categories of devices — proper transponder, muxponder, and integrated/colored line cards — currently performing the function of

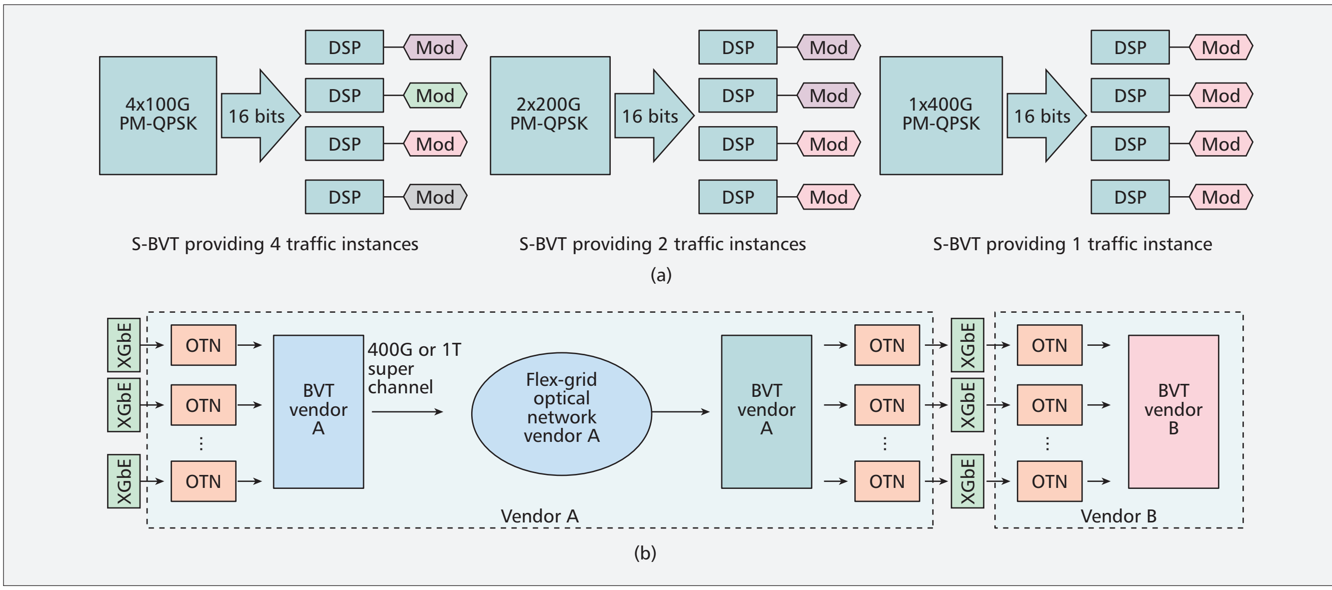


Figure 4. a) Example of S-BVT adaptation in regard to a variable number/size of traffic instances; b) multivendor interoperability using standardized interfaces.

mapping one or more clients into a single DWDM interface: exploitation of pluggable optical modules is desirable, being common for client ports.

SERANO Interface

The S-BVTs used in a SERANO block, facilitating the spectral grooming of flows, are only a "light" version of the corresponding source/destination S-BV transceivers like those described above. Specifically, SERANO is a "service-transparent" module, meaning that no protocol processing or flow granularity adaptation takes place in it, so the S-BVTs consist of a fixed-receiver tunable transmitter pair connected back to back, allowing only for signal carrier, optical bandwdth, and modulation format adaptation of the incoming flows.

Elastic OTN Interface

Elastic and flexible optical interface is currently provided by a flexible optical channel data unit (ODUflex). A flexible OTUflex line is not at present specified; however, a rate-flexible optical transport unit (OTU) line interface, complementing the flexibility provided by ODUflex in the low-order ODU (LO-ODU) service layer, is expected to be one of the distinguishing attributes of the OTN hierarchy evolution "beyond 100G" (B100G). The current predominant thinking in the OTN evolution debate is in favor of defining an "nx100 Gb/s" (with $n \geq 2$) iterative structure, termed OTUCn, meaning that the transport line can be sized in steps of 100 Gb/s (i.e., 200G, 300G, 400G).

An OTN IDEALIST S-BVT will have the capability to generate an elastic OTUCn with the n variable over time to follow the slow load fluctuations (daily, weekly, etc.) occurring in the network. If the active OTU interface does not make use of all the available resources, the spare ones can be reused for other existing traffic demands or future unknown demands. Therefore, the network node interface (NNI) should be modular enough to allow the partition of the common hardware pool into sub-transponders, one per OTU flow set active on that S-BVT.

Network Design Considerations

Based on the listed requirements and the AoD-based Flex-OXC node design discussed in the previous sections, hereafter considerations are described concerning the full functionality of the proposed node.

Interoperability

Based on standardized Ethernet, IP, and OTN elastic interfaces available as optical modules in the AoD-based Flex-OXC, guaranteed interoperability is possible, as depicted in Fig. 4b. Here, multivendor node interoperability is achieved through standard Gigabit Ethernet (GbE) interfaces. An OTN layer is assumed between XGbE ports (e.g., X = 100) and the S-BVT. The OTN may be proprietary to each S-BVT vendor; for this reason, in Fig. 4b, OTNs are reported for both vendors A and B. At the receiver side of vendor A, OTN clients are extracted by the detected super-channel. Then the original XGbE traffic is extracted by the OTN frames and provided to vendor B through standard interfaces.

Moreover, multivendor interoperability is possible if the previous node subsystems are used. For example, by using the SERANO subsystem, multivendor interoperability is guaranteed since the embedded multiflow optical receivers and transmitters can operate as a flow-forwarding engine in a service-transparent/protocol-agnostic vendor-independent manner.

Defragmentation

Push-Pull Approach — In flex-grid optical networks, fragmentation of spectrum resources may significantly affect overall network efficiency. Traditionally, defragmentation techniques are

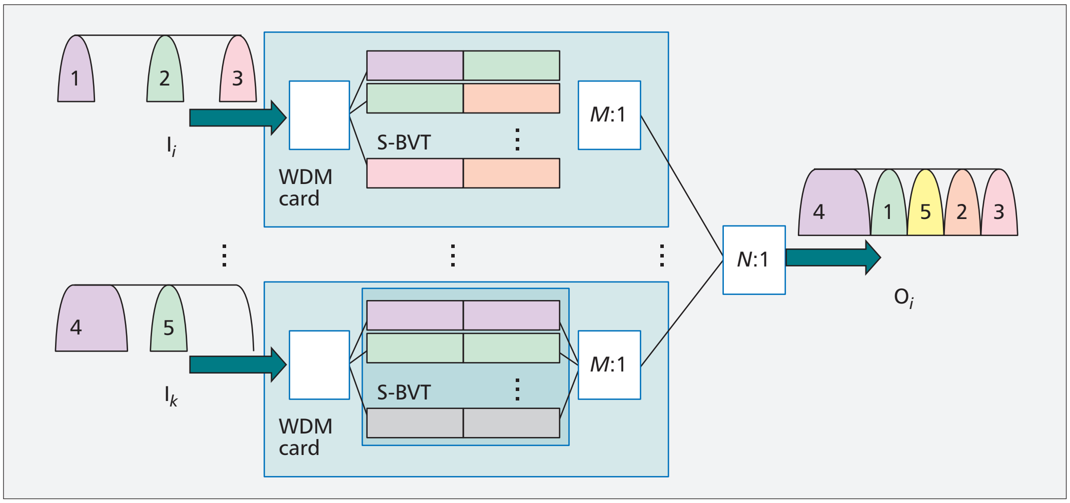


Figure 5. Exploitation of SERANO properties for defragmentation.

By using the SERANO subsystem, multivendor interoperability is guaranteed since the embedded multiflow optical receivers and transmitters can operate as a flow-forwarding engine in a service-transparent/protocol-agnostic vendor-independent manner.

implemented by relying on additional resources, such as spare expensive optical transponders. Alternatively, a novel defragmentation technique, called push-pull, can be applied. The technique is based on dynamic lightpath frequency retuning upon proper reconfiguration of the allocated spectrum resources in the Flex-OXC node. Push-pull defragmentation does not require additional transponders and does not determine traffic disruption. The technique has been successfully performed on a 100 Gb/s PM-QPSK lightpath and on a multi-carrier signal (i.e., super-channel), providing hitless defragmentation [12].

The SERANO Approach — The adaptation functions provided by SERANO node subsystem allow network operators to efficiently implement spectral defragmentation, as shown in Fig. 5, which demonstrates how the BVT transponders can flexibly shift the received slots to any appropriate slot in the output fiber, increasing network utilization.

NODE CASCADABILITY

In the previous sections, we implicitly assumed that optical channels are not degraded when propagating through a cascade of filtering elements that are part of AoD-based Flex-OXC nodes. This assumption is valid as long as the channel does not entirely occupy the selected slot, and the number of hops is limited. In IDEALIST, we target national or pan-European flexible optical networks with the purpose of maximizing their spectral efficiency while also reducing the spectral occupancy of each channel. In this scenario, the penalty induced by narrow filtering plays an important role. In fact, the impact of narrow filtering may also be significant for flows passing through Flex-OXC nodes transparently, that is, not dropped to a SERANO module or a regenerating functional module of the AoD-based Flex-OXC node. This kind of degradation affects media channels, independent of their origin.

Current fixed-grid systems allocate channels in a 50 GHz ITU-grid and transmit at baud rates up to 32 GBaud. By enabling flexible allocation, the frequency slot size can be reduced down to 37.5 GHz, assuming a granularity of 12.5 GHz. In this case, a 128 Gb/s PM-QPSK channel increases its spectral efficiency from ~2.6 to 3.4 b/s/Hz when going from 50 to 37.5 GHz. In [13], the performance of advanced modulation formats was analyzed within a defined worst case scenario for the Flex-OXC cascade, where a channel was filtered each time it passed through an optical node. Here, 32 GBaud over 37.5 GHz was considered with an equivalent –3 dB filter bandwidth, which is already comparable to the symbol rate after only three Flex-OXC nodes (i.e., assuming 2 × BV-WSS/node). Clearly, the induced filter penalty exponentially increases with the number of crossed nodes, leading to severe signal degradation and a dramatic reduction of reach as reported in [13]. Based on this, it was proposed to optically shape the pulse of the channel within each AoD-based Flex-OXC. The aforementioned solution extended the reach, enabling, for example, propagation through 32 BV-WSSs (i.e., 16 Flex-OXCs) with negligible penalty. This result shows that the upgrade to long-haul as well as regional networks employing flex-grid systems is now possible. These findings were experimentally verified in [14], although the overall benefit was limited by the additional losses of the optical shaping.

A further solution to address the detrimental filtering effects consists of configuring different slot widths along the path of an optical channel [15]. Two options can be exploited. The first, differentiated filter configuration, applies to a single optical media channel and enables narrow filtering in only a few nodes along the path, while larger filters are configured in the remaining nodes of the same lightpath. In this way, the optical signal efficiently occupies less spectrum where the network presents limited unreserved resources, while adequate reach is guaranteed given that the other nodes (where the signal occupies more spectral resources, i.e., larger filters are configured) do not induce detrimental filtering effects.

The second option, called a super-filter, applies to multiple optical super-channels, also

Addressing the emerging requirements for flexibility, scalability, resilience and adaptability, we propose the AoD-based Flex-OXC nodes. As one of the building blocks of this Flex-OXC node, the SERANO architecture has been proposed to perform adaptation, regeneration and media channel conversion reducing overall switch complexity and cost.

generated by different source-destination pairs. In this case, each node applies a unique filter configuration (i.e., exploiting the same flat region of the filter) to all lightpaths that are contiguous in frequency and co-routed along its ports. With this strategy, filtering effects are almost negligible on the signals traversing the central region of the super-filter, thus avoiding a reduction in the optical reach or an increase of the reserved spectrum resources.

Conclusions

In this article, an innovative Flex-OXC node solution based on the integration of several functional building blocks was presented. Addressing the emerging requirements for flexibility, scalability, resilience, and adaptability, we propose the AoD-based Flex-OXC nodes. As one of the building blocks of this Flex-OXC node, the SERANO architecture has been proposed to perform adaptation, regeneration, and media channel conversion, reducing overall switch complexity and cost. In addition, the design aspects of two optical interfaces as functional subsystems of the optical node were also presented: the elastic OTN and SERANO interfaces. Finally, specific use cases were discussed exemplifying how interoperability, defragmentation, and node cascadability can be achieved when designing networks adopting the proposed architectural solutions.

Acknowledgment

This work was supported by the FP-7-IDEAL-IST project, grant agreement number 317999.

References

[1] O. Gerstel *et al.*, "Elastic Optical Networking: A New Dawn for the Optical Layer?," *IEEE Commun. Mag.*, vol. 50, Feb. 2012, pp. S12–S20.

[2] N. Amaya *et al.*, "Introducing Node Architecture Flexibility for Elastic Optical Networks," *J. Opt. Commun. Net.*, vol. 5, June 2013, pp. 593–608.

[3] B.R. Rofoee *et al.*, "All Programmable and Synthetic Optical Network: Architecture and Implementation," *J. Opt. Commun. Net.*, vol. 5, Sept. 2013, pp. 1096–1110.

[4] B. R. Rofoee *et al.*, "Programmable On-Chip and Off-Chip Network Architecture on Demand for Flexible Optical Intra-Datacenters," *Opt. Express*, vol. 21, Mar. 2013, pp. 5475–80.

[5] A. Muhammad *et al.*, "Introducing Flexible and Synthetic Optical Networking: Planning and Operation Based on Network Function Programmable ROADMs," *J. Opt. Commun. Net.*, vol. 6, July 2014, pp. 635–48.

[6] M. Garrich *et al.*,"Architecture on Demand: Synthesis and Scalability," ONDM '12, Colchester, UK. April, 2012.

[7] M. Dzanko *et al.*, "Experimental Demonstration and Benefits of Self-Healing Hard-Wired and Synthetic ROADMs," *OFC '14*, W1C.3. San Francisco, CA, Mar. 2014.

[8] N. Amaya *et al.* ,"Software Defined Networking (SDN) over Space Division Multiplexing (SDM) Optical Networks: Features, Benefits and Experimental Demonstration," *Opt. Express.*, vol. 22, no. 3, Feb. 2014, pp. 3638–47.

[9] A. Stavdas *et al.*, "Optical Packet Routers: How They Can Efficiently and Cost-Effectively Scale to Petabit/sec," *OSA J. Opt. Net.*, vol. 7, Oct. 2008, pp. 876–94.

[10] A. Stavdas *et al.*, "A Novel Architecture for Highly Virtualised Software-Defined Optical Clouds," *ECOC '13*, Mo.3.E.4, London, U.K., Sept. 2013.

[11] N. Sambo *et al.*, "Modeling and Distributed Provisioning in 10–40–100-Gb/s Multirate Wavelength Switched Optical Networks," *J. Lightwave Tech.*, vol. 29, May 2011, pp. 1248–57.

[12] F. Cugini *et al.*, "Push-Pull Defragmentation Without Traffic Disruption in Flexible Grid Optical Networks," *J. Lightwave Tech.*, vol.31, Jan. 2013, pp. 125–33.

[13] T. Rahman *et al.*, "On the Mitigation of Optical Filtering Penalties Originating From ROADM Cascade," *IEEE Photon. Tech. Lett.*, vol. 26, Jan. 2014, pp. 154–57.

[14] T. Rahman *et al.*, "Mitigation of Filtering Cascade Penalties Using Spectral Shaping in Optical Nodes," *ECOC '14*, P.4.19. Cannes, France, Sept. 2014.

[15] F. Paolucci *et al.*, "Filter Optimization in SDN-Based Flexgrid Networks," *OFC '14*, Th3I.4, San Francisco, CA, Mar. 2014.

Biographies

Emilio Hugues-Salas (e.huguessalas@bristol.ac.uk) received his M.Sc. and Ph.D. degrees from the University of Essex, United Kingdom. He is a research fellow in the High Performance Networks Group at the University of Bristol, United Kingdom. His research interests include the development of high-capacity, reconfigurable, flexible, and multidimensional optical nodes and networks. He is an author and co-author of over 60 international journal and conference papers.

Georgios Zervas holds M.Eng. and Ph.D. degrees from the University of Essex. He is a senior lecturer in optical and high performance networks at the University of Bristol and a member of the High Performance Networks (HPN) group. He leads the activities in optical transport and switching, as well as software/hardware programmable systems and networks for backbone and DCN/HPC interconnects. He has published over 150 journal and conference papers, including post-deadline papers, and has received best paper awards. He is a TPC member of numerous leading conferences and a contributor to standards.

Dimitra Simeonidou is a professor with the HPN group at the University of Bristol. She is a leading academic in the field of optical networks, a pioneer in future Internet research and experimentation (FIRE), an influential researcher in the fields of grid/cloud computing, and a founder of Transport SDN. She has chaired several international conferences (ECOC '13, OFC '15) and committees across these technical fields. She is an author/co-author of over 400 publications, many of which have received best paper awards, 12 patents, and several major contributions to standards.

Evaggelos Kosmatos received Dipl.Ing. and Ph.D. degrees, both in electrical and computer engineering, from the National Technical University of Athens (NTUA), Greece, in 2002 and 2007, respectively. His research interests lie in the area of network control plane design, resource allocation in broadband networks, intelligent multimedia content delivery, and optical/wireless network integration. He has participated in several EU and national research projects in the above areas.

Theofanis Orphanoudakis [SM] received his Dipl.Ing. and Ph.D. degrees, both in electrical and computer engineering, from NTUA in 1995 and 1998, respectively. Currently he is an assistant professor at the Hellenic Open University and a research associate of the ONG group. He has participated in several EU research programs, and has led projects in the areas of broadband networking components and high-speed core and access networks. He is a Senior Member of the Technical Chamber of Greece.

Alexandros Stavdas has a B.Sc. in physics from theUniversity of Athens, an M.Sc. in optoelectronics and laser devices from Heriot-Watt /St Andrews University), and a Ph.D. from University College London (UCL) in WDM networks. Currently, he is a professor of optical systems and networking in the Department of Informatics and Telecommunications of the University of Parma. He is an author of over 140 journal and conference papers. He has also served on TPCs of many international conferences. His current research interests include future Internet architectures, heterogeneous access networks, physical layer modeling of optical networks, OXC architectures, and optical packet/burst switching.

Marc Bohn received his Dipl.-Ing. degree in electrical engineering in 1999 and his Ph.D. degree in 2005 from the University of Kiel, Germany. He was with Siemens, Nokia Siemens Networks, and Coriant working on transmission aspects in optical communications in different research and has headed the Research and Technology Department inside Coriant.

ANTONIO NAPOLI (antonio.napoli@coriant.com) received M.Sc. and Ph.D. degrees in electrical engineering from Politecnico di Torino in 2002 and 2006, respectively. In 2004 he was a visiting student at Universitat Pólitecnica de Catalunya (UPC), Barcelona, Spain, and in 2005 at UCL. In 2006, he joined Siemens COM and in 2007 NSN, working on DSP for high-speed optical networks. Since 2013 he has been with Coriant. His main interests are DSP and advanced modulation formats for long-haul optical systems. He has (co)-authored more than 80 peer-reviewed articles and seven patents.

TALHA RAHMAN received his M.Sc. (Hons.) degree in advanced optical technologies with specialization in optical communication systems from FAU Erlangen-Nürnberg, Germany, in 2012. He conducted his internship and Master thesis research at Alcatel-Lucent Bell Labs Stuttgart. He is currently employed by Eindhoven University of Technology as a Ph.D. candidate under EU-FP7 project IDEALIST and conducts his research at Coriant R&D Labs in Munich, Germany. His area of research includes flexible optical transceivers focusing on 400 Gb/s and 1 Tb/s superchannels.

FILIPPO CUGINI received his M.S. degree in telecommunication engineering from the University of Parma, Italy. Since 2001, he has been with the National Laboratory of Photonic Networks, CNIT, Pisa, Italy. His main research interests include theoretical and experimental studies in the field of optical communications. He is a co-author of 12 international patents and more than 100 IEEE publications.

NICOLA SAMBO (n.sambo@sssup.it) received his M.Sc. cum laude in telecommunication engineering from the University of Pisa, Italy, in 2004, and his Ph.D. from Scuola Superiore Sant'Anna (SSSUP), Pisa, Italy, in 2009. In 2007, he was a visiting student at France Télécom, Lannion, France. Currently, he is an assistant professor at SSSUP. He is also collaborating with CNIT and Ericsson Italy. His research interests include optical network architecture, transmission performance modeling, and control plane. He is an author of about 80 publications including international journals, conference proceedings, and patents.

SILVANO FRIGERIO is a member of technical staff within the Optics Product Unit, Vimercate, Italy. He received a degree in electronic engineering from Politecnico di Milano, Italy. He brings extensive and sustained experience developed as a system architect of optical multi-service nodes and a SONET/SDH ASIC designer. Presently he is focusing on aspects regarding multi-technology solutions and OTN transformation at both the equipment and network levels. He is an Alcatel-Lucent Italia TCT Principal Engineer and holds several patents concerning transmission networks.

ANTONIO D'ERRICO received his Ph.D. degree cum laude in telecommunication engineering from SSSUP in 2008. He is currently with Ericsson Research as an experienced researcher. His research interests include advanced technological solutions for optical networks. He is an author of more than 75 papers published in international journals, conference digests, and patents.

ANNACHIARA PAGANO received her degree in physics from the University of Turin in 1994. In the same year she joined Telecom Italia (formerly CSELT) to work on fiber optics and DWDM metro and long distance communication systems. Her current fields of interest are next-generation optical networks, focusing both on high-capacity network design and testing, together with research activities on optical transmission and scouting of new technologies. She has co-authored more than 50 technical publications covering design and testing aspects of optical components and systems. She has been involved in several European Projects.

EMILIO RICCARDI, after graduating in physics, joined Telecom Italia Lab (formerly CSELT) in 1992. His current interest is dedicated to elastic optical networks. He has been involved in several EU Projects and presently leads the Data Plane Working Package of the IDEALIST EU Project. He has coauthored more than 40 journal and conference publications.

VICTOR LÓPEZ is a telecommunication engineer (M.Sc., Universidad de Alcalá de Henares, 2005, and Ph.D., Universidad Autónoma de Madrid, UAM, 2009). In 2004, he joined Telefónica I+D as a researcher, where he was involved in transport networks projects. In 2006, he worked as a researcher and assistant professor at UAM. In 2011, he joined Telefonica I+D as a technology specialist. He has co-authored more than 100 publications and contributed to IETF drafts. His research interests include the integration of Internet services over IP/MPLS and optical networks, and control plane technologies (PCE, SDN, GMPLS).

JUAN PEDRO FERNÁNDEZ-PALACIOS GIMÉNEZ received his M.S. in telecommunications engineering from the Polytechnic University of Valencia in 2000. In September 2000 he joined TIDSA, where he is currently leading the Metro and Core Transport Networks Department. He has been involved in several European Projects, such as Eurescom P1014 TWIN, and IST projects DAVID, NOBEL, NOBEL-2, STRONGEST, BONE, and MAINS, as well as other internal projects related to the development of optical networks in the Telefónica Group. Currently he is coordinating the FP7 project IDEALIST as well as the standardization activities within the CAON cluster, where he is acting as Co-Chair.

400 Gigabit Ethernet Using Advanced Modulation Formats: Performance, Complexity, and Power Dissipation

Jinlong Wei, Qixiang Cheng, Richard V. Penty, Ian H. White, and David G. Cunningham

ABSTRACT

We review possible architectures for 400 Gigabit Ethernet links based on advanced modulation formats for the first time. Their optical link power budget, digital complexity, and power dissipation are compared via simulations. The challenges of implementing the physical layer are discussed.

INTRODUCTION

Internet traffic continues to grow exponentially, fueled by applications such as high definition TV, video on demand, and clod computing. This drives the need for massive data centers with flexible scalability. Optical interconnects play a critical role in intra- and inter-data-center communications. The CPU computing speed and aggregate network traffic double every 18 months [1]. Today's fastest standardized Ethernet rate is 100 Gb/s, completed in 2010 (IEEE 802.3ba) with the previous two rates being 10 Gb/s and 40 Gb/s. In accordance with the 4× and 10× bit rate scaling model, IEEE 802.3 started the 400 Gigabit Ethernet (GbE) Study Group in March 2013, which successfully justified the creation of the IEEE P802.3bs 400GbE Task Force in May 2014 [1]. The 400 GbE Task Force agreed on objectives for both single-mode fiber (SMF) and multimode fiber (MMF) links to cover the historical reaches of 100 m over MMF, and 2 km and 10 km over SMF. In addition, a new 500 m SMF objective was agreed on to support the longer intra-data-center connection required for the emergent massive data centers, which MMF cannot support at 400 Gb/s.

Currently, the challenge for Ethernet evolution mainly exists at the physical layer. Figure 1a shows an idealized functional diagram of SMF-based high-speed Ethernet including both the electrical interface and the optical module. The application-specific integrated circuit (ASIC) sends high-bandwidth input/output signals through the electrical interface and a printed circuit board (PCB) channel to the optical module. The optical module contains optical transceivers responsible for electrical-to-optical (E/O) and O/E conversion. Intensity modulation and direct detection (IMDD) are preferred due to its cost effectiveness. First embodiments of the higher rate standard might include optical transceiver components similar to those shown in Fig. 1a. Not all blocks in the transceiver are needed, depending on the chosen electrical and optical lane rate and component performance. As implementations mature, some functions could be subsumed into the system ASIC to simplify the optical module, such as error correction, gearbox, and advanced modulation.

LANE RATE

The highest lane rate supported by IEEE 802.3 electrical interfaces is 25 Gb/s, the wide deployment of which will soon occur. Electrical interfaces beyond 25 Gb/s are very challenging. Thus, to allow practical electrical PCB trace length/loss, multilevel modulation formats and error correction are likely to be adopted, as shown in Fig. 1b. 100GBASE-KP4 has adopted pulse amplitude modulation-4 (PAM-4) for electrical backplane interconnects. Also, the Optical Internetworking Forum is developing electrical interfaces at 50–56 Gb/s per lane, and baseline draft implementer's agreements for NRZ using low loss electrical channels and PAM-4 for higher loss electrical channels are being developed.

For optical Ethernet, the fastest lane rate developed to date is 25 Gb/s, beyond which the optoelectronic component performance becomes demanding on both bandwidth capability and fiber dispersion. For SMF applications, wavelength-division multiplexing (WDM) is a natural choice for data rate upgrade: an example is 100GBASE-LR4. However, WDM inevitably has high system cost and power dissipation as a result of the number of optoelectronic components and wavelengths used. Alternatives to WDM are transmission over a set of parallel fibers or transmission at a higher lane rate over a single wavelength. Both options were studied as part of the 400G Study Group. Although high-speed optical components supporting 112 Gb/s non-return to zero (NRZ) single-wavelength transmission have been reported [2], further cost reduction is needed to make them commercially viable. In contrast, advanced modulation formats might enable the 100G per wavelength with fewer advanced optical components.

Jinlong Wei, Qixiang Cheng, Richard V. Penty, and Ian H. White are with the University of Cambridge.

David G. Cunningham is with Avago Technologies.

 0163-6804/15/$25.00

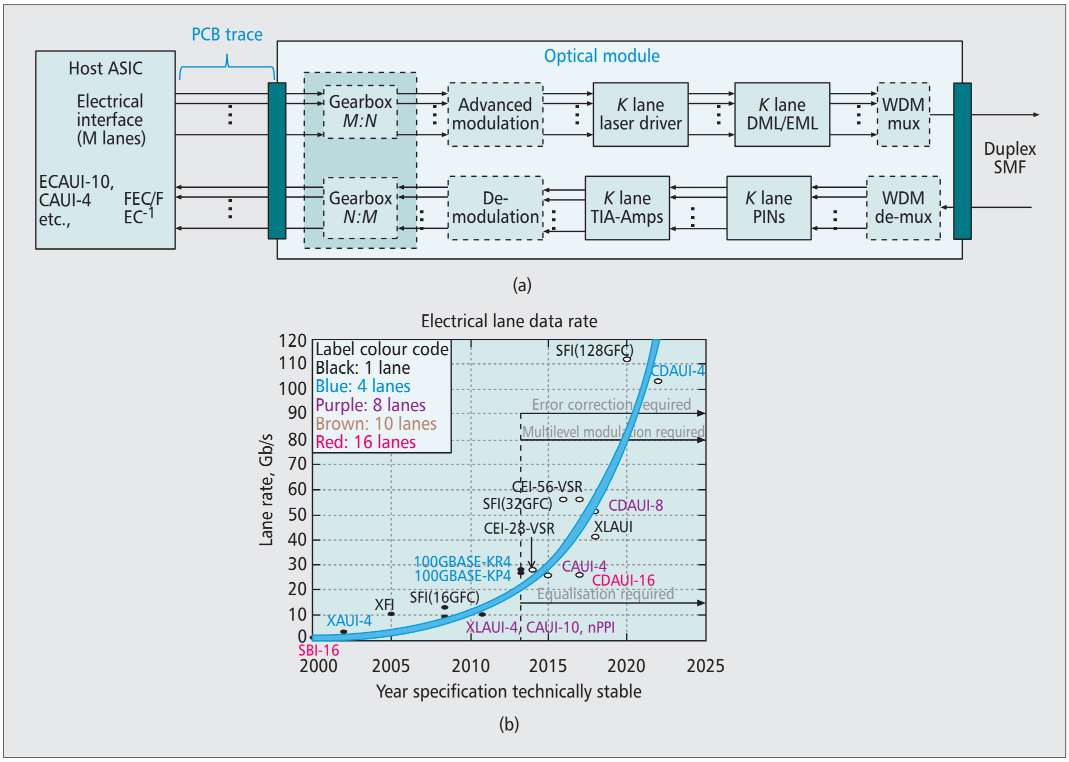


Figure 1. a) Diagram of a high-speed Ethernet including electrical interface and optical module; b) trends for standards-based electrical interfaces. Filled circles indicate developed, while empty circles represent under development or discussion.

The highest lane rate supported by IEEE 802.3 electrical interfaces is 25 Gb/s, the wide deployment of which will soon occur. Electrical interfaces beyond 25 Gb/s are very challenging. Thus, to allow practical electrical PCB trace length/loss, multilevel modulation formats and error correction are likely to be adopted.

400 GbE

The 400 GbE Task Force is interested in adopting 50 Gb/s per electrical lane, and 50 or 100 Gb/s per optical wavelength. The Task Force has taken votes implying that the preferred solution would be 100 Gb/s per wavelength. However, given the challenges of implementing 100 Gb/s per wavelength, 50 Gb/s per wavelength is still a strong contender. While for higher lane rates, the Task Force is still open to the choice of modulation format and number of wavelengths per fiber [1], NRZ and PAM4 at 50 Gb/s and PAM4 and orthogonal frequency division multiplexing (OFDM) at 100 Gb/s per wavelength dominate the debate. For SMF, the simplest solution is to leverage the 100GBASE-LR4 modules by implementing 16 lanes each operating at the existing 25 Gb/s rate (Fig. 2a). In this scheme, four duplex fiber pairs (DPs) are required, each connected to a 100GBASE-LR4 module with four WDM lanes per fiber. This solution, however, inherits the cost and power dissipation of 100GBASE-LR4. It is envisaged that first generation 400 GbE might be based on eight lanes with 50 Gb/s per lane using two DPs, as shown in Fig. 2b. The second generation might adopt one DP with four lanes each operating at 100 Gb/s [1]. The 400 GbE MMF links might adopt a configuration similar to that shown in Fig. 2a but using parallel links rather than WDM to fully leverage the 100GBASE-SR4 standard. However, if 50 Gb/s per vertical cavity surface emitting laser (VCSEL) lane using NRZ or PAM-4 can be developed for MMF links, the configuration in Fig. 2b with WDM or parallel fibers for each lane might also be possible.

Advanced Modulation Formats

Unless electrical and optical components with much higher bandwidth become commercially available, the use of advanced modulation formats seems unavoidable for 400 GbE. For SMF, there have been a number of demonstrations of various modulation formats, for example, 50 Gb/s per polarization using PAM-4 over 100 m SMF [3], 50 Gb/s carrierless amplitude and phase (CAP) modulation over 5 km SMF link [1], 103 Gb/s duobinary transmission over 1 km SMF [4], 400 Gb/s (100 Gb/s per lane) multiband carrierless amplitude/phase (CAP) signal transmission over 20 km (or 40 km optically amplified) SMF [5], 112 Gb/s quadrature amplitude modulation-16 (QAM-16) over 4 km SMF link [6], and 448 Gb/s (112 Gb/s per wavelength) optical OFDM over 10 km SMF [7]. These demonstrations are either based on real-time [1] or offline processing [3–7]. This article offers the first known comprehensive review of a wide range of these schemes that takes into account the key criteria, including optical link power budget, system complexity, and power dissipation for 400 GbE scenarios. The advanced modulation formats considered include duobinary, PAM, hybrid CAP/QAM, and optical OFDM.

Duobinary

Doubinary operates at the full NRZ data rate and introduces controllable intersymbol interference (ISI) to reduce the signal bandwidth by two [4]. To shape doubinary pulses one can leverage the natural frequency rolloff of the optical transceiver [4], although commercial products

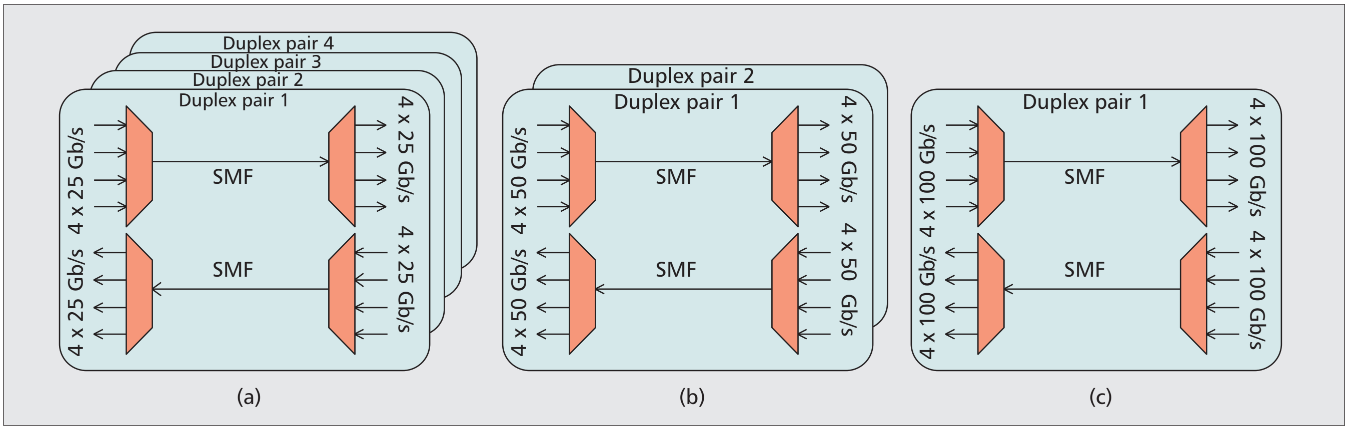


Figure 2. 400 GbE solutions with different numbers of lanes with different bit rates per lane.

would probably include an equalizer (Fig. 3a) to optimize link performance and provide tolerance to frequency response variations of the optical components. This is a simple and low-cost implementation. At the receiver side, a simple three-level slicer architecture can be used to demodulate a doubinary signal [4].

PAM

PAM is a modulation technique where information bits are encoded in the amplitude of a series of signal pulses. Figure 3b shows a typical digital signal processing (DSP) implementation of an optical PAM link. PAM symbols often first pass through a pre-emphasis filter before conversion to analog pulses via a digital-to-analog converter (DAC). After E/O conversion, the optical PAM signal propagates through a fiber link and is converted into an electrical signal by a photo-diode (PD) followed by signal processing in the receiver which is the inverse of that in the transmitter. The transceiver equalizers are optional depending on the channel conditions.

Hybrid CAP/QAM

Figure 3c shows a hybrid CAP/QAM system that combines a CAP transmitter and a modified QAM receiver. The CAP transmitter generates a QAM-like signal by combining two multilevel signals using two electrical shaping filters the impulse responses of which form an orthogonal Hilbert pair [8]. The combined signal is up-converted to an optical carrier by an optical modulator. The PD detected radio frequency (RF) signal is down-converted into baseband signals using electrical mixers and a local oscillator (LO). Two matched filters are used in the receiver to recover the transmitted in-phase (I) and quadrature (Q) signals. A CAP transmitter has the advantage of easy implementation as no mixers or RF carrier sources are involved, while a modified QAM receiver helps eliminate the crosstalk between I and Q channels that a CAP receiver otherwise has [8].

Optical OFDM

Figure 3d depicts an optical OFDM system. The major operations required to generate an electrical OFDM signal include data modulation format mapping using QAM constellations, power loading, inverse fast Fourier transform (IFFT), cyclic prefix insertion, OFDM symbol serialization, and DAC. The generated electrical OFDM signal is then converted into an optical signal. On the receiver side, after O/E conversion, the received electrical OFDM signal is then digitized and processed with a procedure inverse to that adopted in the OFDM transmitter.

DML vs. EML

IMDD links can be realized by using a directly modulated laser (DML) or an externally modulated laser (EML). The DML has advantages over EML configurations such as low cost, compactness, low power consumption, and high optical output. However, when using a DML, the optical performance is limited by laser dynamic nonlinearity [9]. Despite this, from a commercial viewpoint, DMLs are preferable to EMLs in short-haul scenarios.

Figure 4 shows the eye diagrams of optical data links using various modulation formats. To obtain Fig. 4, the DML, EML, and optical receiver are modeled based on rate equations, a Gaussian response, and a fourth order Bessel filter, respectively, and their corresponding 3 dB bandwidths are 17 GHz, 18.6 GHz (both 12 ps 20–80 percent rise time), and 21 GHz, respectively [9]. The choice of each modulation format for a certain single-lane bit rate is based on two factors: 1) the best trade-off between spectral efficiency and signal-to-noise ratio (SNR); 2) the baud rate being lower than two times the DML bandwidth [9]. For single-lane 25 Gb/s NRZ and 50 Gb/s PAM-4 links, Fig. 4 indicates that the effect of DML nonlinearity on the signal quality is negligible compared to the EML cases. Unless a higher transceiver bandwidth of > 25 GHz is used, when duobinary, PAM-4, or PAM-8 is used for 100 Gb/s per lane, the link fails due to the DML nonlinearity [9].

The challenge for a single-lane 100 Gb/s DML link can be addressed by using hybrid CAP-16/QAM-16 and QAM-16-OFDM. Although use of a DML leads to a closed eye diagram in the QAM-16 receiver (Fig. 4d), a feed forward equalizer (FFE) prior to the slicer shown in Fig. 3c can mitigate ISI and achieve the same equalized signal quality as the EML counterpart. For 100 Gb/s QAM-16-OFDM, the effect of DML nonlinearity on both the unequalized and equalized constellation diagrams

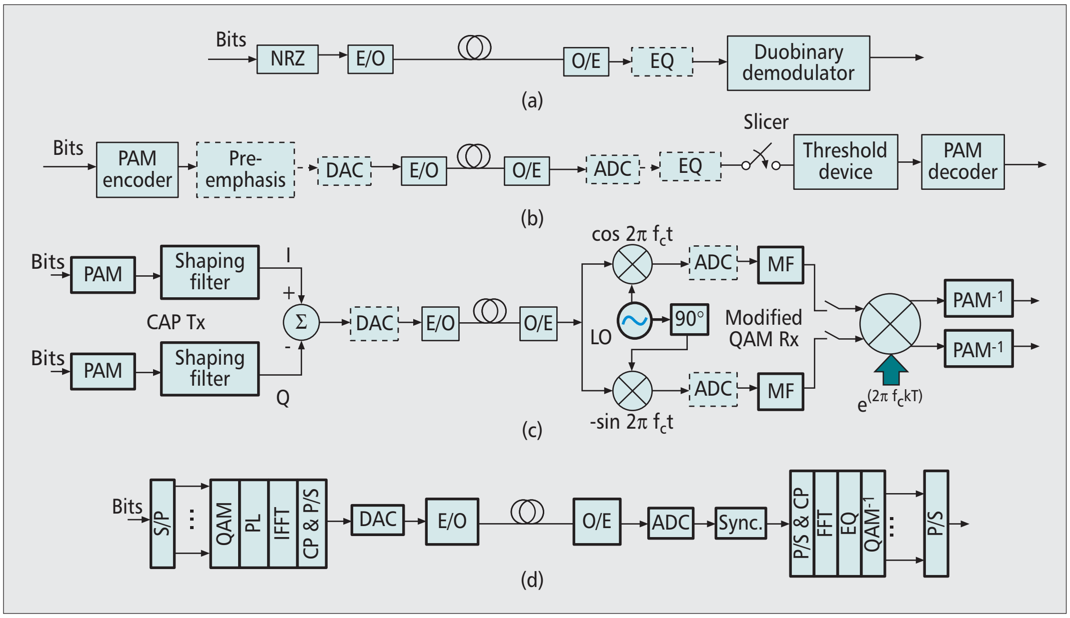


Figure 3. System diagram of an optical data link using: a) duobinary; b) PAM; c) hybrid CAP/QAM; d) QAM-OFDM. E/O: electrical-to-optical conversion; EQ: equalizer; MF: matched filter; PL: power loading; P/S: parallel-to-serial conversion; CP: cyclic prefix; Sync: synchronization; (I)FFT: (inverse) fast Fourier transform; DAC/ADC: digital-to-analog/analog-to digital conversion.

For high speed optical data links using multilevel modulation schemes, FEC has to be adopted to offset the multilevel penalty. The IEEE 400 GbE Task Force has started to develop an architecture from MAC to PMD including how different FEC's per PMD can be supported.

(Figs. 4f and 4g) is obvious compared to the EML case (Figs. 4m and 4n)]. However, the penalty for DML nonlinearity at a FEC threshold BER (typically on the order of 1×10^{-3}) is almost negligible [9].

Digital versus Analogue Implementation

DSP is now widely used in long haul transmissions using coherent optical transmitters and receivers [10]. Recently, use of DSP in short-haul IMDD applications has attracted attention due to advances in complementary metal oxide semiconductor (CMOS) technology [5–7]. The complexity of the required DSP is a critical factor that determines the system cost and power dissipation. The DSP blocks (including the DACs and ADCs) are typically the most power consuming elements of a typical high-speed DSP-based optical data link [9].

Analog implementation based on specialized analog circuits eliminates the need for DAC and ADC modules, and has the potential to be lower cost and power. PAM-4 transmitters [3], CAP systems [1], and hybrid CAP/QAM systems [8] have been demonstrated based on analog implementations. However, analog implementation of an optical OFDM system requires hundreds of carriers and filters in either the RF or optical domain, which makes it impractical.

FEC

For high-speed optical data links using multi-level modulation schemes, FEC has to be adopted to offset the multi-level penalty. The IEEE 400 GbE Task Force has started to develop an architecture from medium access control (MAC) to phase physical medium dependent (PMD) sublayer including how different FECs per PMD can be supported. For example, the IEEE 802.3bj FECs for PAM-4 and Bose-Chaudhuri-Hocquenghem (BCH) FEC for OFDM have been proposed within the 400 GbE Task Force [1]. In this article , RS(544,514) defined in IEEE 802.3bm with an FEC overhead of about 6 percent and code gain of about 7 dB is considered in simulations. The IEEE 400 GbE Task Force has also suggested architectures allowing flexible implementation of FEC in either a module, a switch ASIC, or a separate physical layer (PHY) card.

Optical Link Power Budget

Figure 5 shows the link power penalties for the various 400 GbE schemes considered here. The penalties include the relative receiver sensitivity, dispersion, relative intensity noise (RIN), link loss, timing jitter, mixer (for hybrid CAP/QAM only), WDM demultiplexer, and unallocated penalty (i.e., power margin). A per-lane transmit power of 3 dBm is available since it can be supported by most 100 GBASE-LR4 CFP transceivers. Thus, considering a receiver sensitivity of –10 dBm at a bit error rate (BER) of 10^{-12} for a reference 25 Gb/s link [9], the overall power budget is 13 dBo for the 400 Gigabit NRZ link without FEC, and 16.7 dBo for the other links with FEC. The component parameters used in Fig. 4 are adopted here.

DML-based systems can successfully support transmission over 10 km SMF (Fig. 5a). NRZ exhibits the best performance with approximately 1.7 dBo, 1.6 dBo, 2.8 dBo, and 3.7 dBo power margin relative to duobinary, PAM-4, hybrid CAP/QAM, and optical OFDM, respectively. Doubinary has a comparable power margin to PAM-4. However, PAM-4 has half the symbol rate of duobinary for the same bit rate, and is preferable if the electrical circuit speed is limited. Optical OFDM has the largest relative

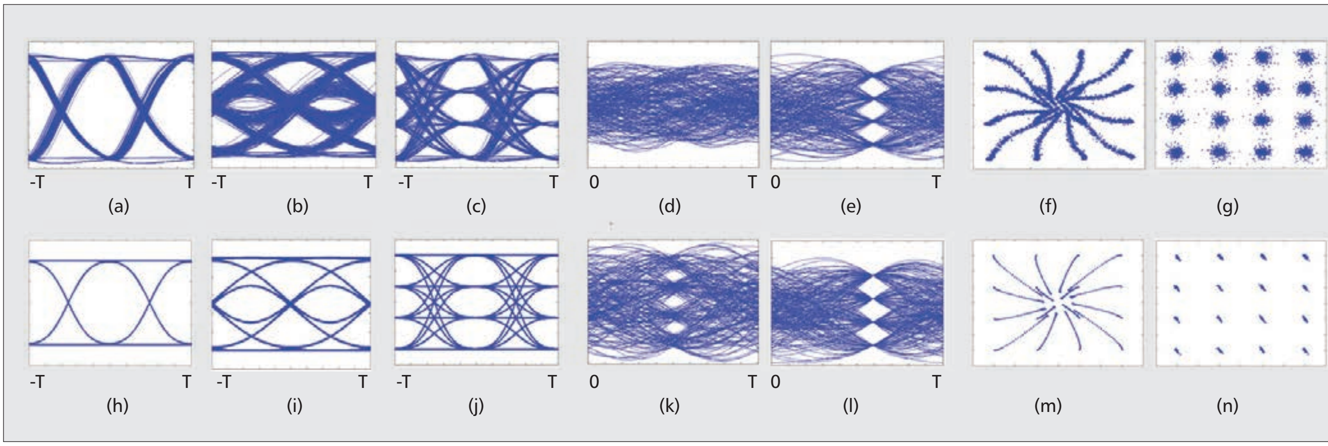


Figure 4. 4 Representative eye diagrams and constellation diagrams for 25 Gb/s NRZ, 50 Gb/s doubinary and PAM-4, and 100 Gb/s hybrid CAP-16/QAM-16 and QAM-16-OFDM. (a–g) are for the DML case; (h–n) are for the EML case.

receiver sensitivity penalty due to multilevel penalty, and the extra penalty introduced by clipping and quantization [9]. The other penalties are similar for all schemes. The mixer penalty only exists in the hybrid CAP/QAM system, and is mainly determined by the noise figure of the RF amplifier required prior to the mixer.

For comparison, the power budget for various modulation formats using EMLs is presented in Fig. 5b. It shows that EML-based duobinary (PAM-4) can support a single-lane 100 Gb/s data transmission over 5 km (10 km) SMF by using a 10-tap T/2 space FFE and 3-tap DFE (5-tap T space FFE and 3-tap DFE), although it can only support a 50 Gb/s single-lane bit rate for DML links. Note that the total power penalty of the 400 Gb/s PAM-4 over 10 km SMF case per our independent calculations agrees well with the IEEE proposal [1]. PAM-4 outperforms duobinary because doubinary has double the symbol rate leading to higher noise enhancement from equalization. The EML-based single-wavelength 100 Gb/s PAM-4 scheme and the DML-based dual-wavelength 100 Gb/s PAM-4 scheme both have strong support within the 400 GbE Task Force. EML-based hybrid CAP-16/QAM-16 or QAM-16-OFDM have comparable power margin to their DML counterparts, verifying that DML nonlinearity is negligible due to their lower symbol rate/signal bandwidth. For this reason OFDM is also a strong contender within the 400GbE Task Force.

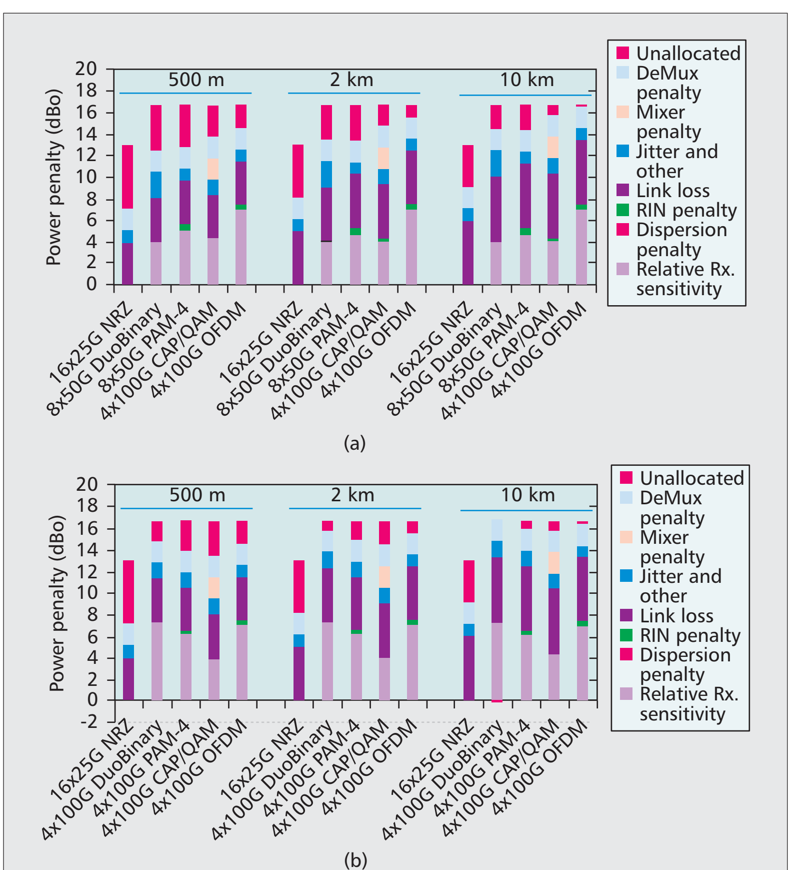


Figure 5. Link optical power penalties for various 400 GbE links using different modulation formats under: a) DML case; b) EML case.

DIGITAL COMPLEXITY

Since NRZ or duobinary is unlikely to adopt digital implementation, we only analyze the digital complexity of the other three cases shown in Fig. 3. The digital complexity is evaluated as the number of real-valued arithmetic operations required per second.

The DSP blocks considered contain the transmitter components prior to the DAC and the receiver components after the ADC, shown in Fig. 3. For PAM-4, the transceiver blocks include the PAM-4 encoder/decoder, equalizers, and threshold device (TD). The arithmetic operations of the PAM-4 transceiver include four comparisons per symbol in the worst case for the PAM-4 encoder/decoder based on a lookup table (LUT), and three comparisons in the worst case for the TD. A simple 2-tap T- spaced FFE-based pre-emphasis or EQ might be implemented to mitigate the ISI caused by the RF circuit/path, which requires two multiplications and one addition. In total, 17 operations per symbol are required for PAM-4. Thus, the complexity of 8 × 50 Gb/s PAM-4 Ethernet link is 8 × $17/T_S$ where T_S is the symbol rate required to achieve 50 Gb/s net bit rate per lane considering the FEC overhead.

For hybrid CAP-16/QAM-16, the transceiver includes two PAM-4 encoders/decoders and two 12-tap T/4 spaced shaping filters/matched filters, two 20-tap $T/4$ spaced FFE filters, one symbol rotator operating at the symbol rate, and two TDs. The shaping filter can be implemented as a LUT. For each output symbol, it requires 4 comparisons and 12 additions due to partial superposition between adjacent symbols. For each symbol, the matched filter needs 12 multiplications and 11 additions. The FFE filter has 20 multiplications and 19 additions for each symbol. The symbol rotator involves one complex multiplication equivalent to four basic real-valued multiplications and two additions. The TD and PAM-4 decoder needs three and four comparisons, respectively. Thus, in total the hybrid CAP-16/QAM-16 transceiver has 184 operations per symbol, and 4 × 100 Gb/s complexity is 4 × $184/T_S$ with T_S being the single-lane CAP symbol rate considering FEC overhead.

The QAM-16-OFDM transceiver involves a number of DSP blocks as shown in Fig. 3d. The input of the IFFT with size N satisfies Hermitian symmetry and allows only $N/2 - 1$ subcarriers to carry user information [9]. Suppose the OFDM frame time duration is T_f and a cyclic prefix of η_{CP}, $N(1 + \eta_{CP})$ samples are transmitted in T_f. The QAM-16 encoder/decoder is implemented as a LUT, which requires four comparisons for real or imaginary channel of each subcarrier. The power loading involves $N/2 - 1$ multiplications. The IFFT/FFT requires $N/2\log_2 N$ complex multiplications and $N\log_2 N$ complex additions [10], equivalent to $5N\log_2 N$ basic real-valued operations. The receiver symbol synchronization requires $6N(1 + \eta_{CP})$ operations in T_f [11]. Assume a pilot-tone overhead of η_{PT}, the channel estimation requires $(N/2 - 1)\eta_{PT}$ complex divisions, and the one-tap equalization needs $(N/2 - 1)$ complex multiplications, equivalent to $6(N/2 - 1)(1 + \eta_{PT})$ real operations. The total computational effort for the 4 × 100 Gb/s QAM-16-OFDM transceiver is

$$M = 23\left(\frac{N}{2}-1\right) + 2*5N\log_2 N + 6N(1+\eta_{CP}) + 6\left(\frac{N}{2}-1\right)(1+\eta_{PT}) \qquad (1)$$

and the complexity is denoted as

$$C_{OFDM} = 4\times\frac{M}{T_f} = 4\times M \frac{R}{\left(\frac{N}{2}-1\right)\log_2 16}(1+\eta_{FEC}+\eta_{CP}+\eta_{PT}) \qquad (2)$$

The following parameters are considered: η_{FEC} = 6 percent for all schemes, single lane net bit rate R = 100 Gb/s (50 Gb/s) for hybrid CAP/QAM and optical OFDM links (PAM-4 link), FFT size η_{CP} = 12.5 percent, and η_{PT} = 5 percent. Based on the above quantitative analysis, the relative complexities of the 8 × 50 Gb/s PAM-4, 4 × 100 Gb/s CAP-16/QAM-16, and 4 × 100 Gb/s QAM-16-OFDM are 1 : 5.4 : 5.7.

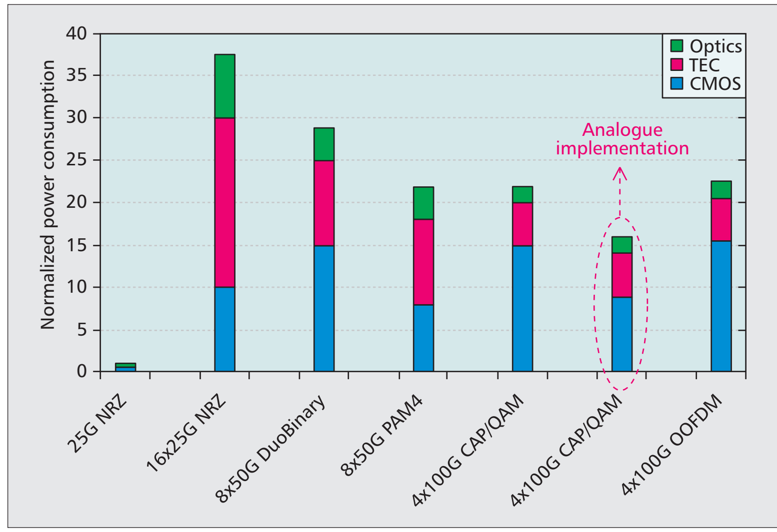


Figure 6. Estimated power dissipation of various SMF transceivers to support 400 Gb/s bit rate.

The complexity of the CAP-16/QAM-16 system is comparable to QAM-16-OFDM. However, the complexity of the OFDM system increases considerably with increasing the IFFT/FFT size according to Eqs. 1 and 2. For example, at N = 256, the relative complexities of the three schemes become approximately 1 : 5 : 7.

Power Dissipation

The power dissipation of the various 400 GbE links are estimated based on state-of-the-art CMOS technology and a benchmark 100GBASE-LR4 link that adopts four parallel WDM 25 Gb/s lanes. The power dissipation is normalized to that of a 25 Gb/s reference NRZ system without considering a thermoelectric cooler (TEC) [9]. For this comparison, 16 × 25 Gb/s NRZ and 8 × 50 Gb/s doubinary are based on analog implementation. Digital implementation is considered for 8 × 50 Gb/s PAM-4 and optical OFDM. Both are considered for 4 × 100 Gb/s hybrid CAP/QAM.

The normalized power dissipation for major transceiver constituent discrete components of each 400 GbE link is given in Fig. 6. The CMOS power consumption refers to that of electrical interface and relevant module transceiver electrical signal processing components including DAC/ADC if used as shown in Fig. 3 for each scheme, while the optics considers that of E/O and O/E converters as well as wavelength multiplexing/demultiplexing. The results take into account IEEE P802.3bs 400 GbE Task Force proposals [1] as well as recent publications [9]. For example, the CMOS power dissipation of 4 × 100 Gb/s optical OFDM is about twice that of 8 × 50 Gb/s PAM-4, which is agreed within the 400 GbE Task Force [1]. It shows that the benchmark 16 × 25 Gb/s NRZ system consumes the most power attributed to the TEC used in each transmitter to stabilize laser wavelength, which typically each consume 1 W [12]. The power dissipations of 8 × 50 Gb/s doubinary and PAM-4, digital 4 × 100 Gb/s hybrid CAP/QAM, and optical OFDM are about 0.76, 0.58, 0.57, and 0.6

Looking ahead, the demand for higher bit rates will continue. Conventional NRZ will become even more demanding on the performance of electrical and optical components. However, advanced modulation formats, together with photonic integration, are likely to be able to address the challenges of further increases in lane rate.

times that of the benchmark NRZ link, respectively, as a direct result of the reduced number of TECs and other components used. The 4 × 100 Gb/s hybrid CAP/QAM based on analogue transversal filters [8] shows as the most power efficient scheme, only 0.42 times that of the benchmark NRZ link as DAC/ADC is not needed. Duobinary consumes more power than the other three schemes because the duobinary demodulator has comparable power dissipation to that of a TEC [4].

The energy efficiency of 400 GbE links has the potential to be improved. Advances in CMOS technology will bring about energy saving in the RF components. This is especially true for optical OFDM. For example, the power consumption of the ADC is almost halved using 40 nm technology compared to 65 nm CMOS [1]. Similar performance can also be expected for DSP blocks. Second, next generation 400 GbE may adopt CWDM configurations that avoid the use of TECs. Additionally, if parallel SMF is used instead of WDM, TECs' and WDM optics can be eliminated, which would reduce the power dissipation and increase the optical budget/margin. Hence, although not shown on the graph, SMF parallel transmission at 50 Gb/s per fiber is a strong contender for the 500 m SMF and future 100 m MMF cases.

Discussion

For high-speed optical Ethernet, the challenges of implementing advanced modulation formats exists in the cost, size, and power dissipation. In order to keep low cost and power, for OFDM, DSP-efficient FFT and simplified algorithms to reduce the peak-to-average power ratio (PAPR) are attractive. Next generation DMLs [13] with > 20 GHz bandwidth and typically 5 dBm per channel optical power need to be used to relax the system power budget and DSP complexity.

At 50 Gb/s per lane, duobinary is potentially a low-power scheme if the optical components can only operate at half the lane data rate but the electronics can operate at the full lane data rate. On the other hand, PAM4 is a good choice as both the electronics and optical components operate at half the lane data rate. At 100 Gb/s per lane, the hybrid CAP/QAM scheme shows excellent potential for power efficiency if analog implementation is considered. Analog transversal filter performance is determined by tap count, tap bandwidth, and tap spacing. CMOS-based RF transversal filters used as equalizers for NRZ signals at 40 Gb/s or beyond have been reported [14]. The major challenge lies in how to optimize the filter parameters while achieving low power and noise.

Photonic integration can effectively improve system power efficiency [1, 15]. Integration not only increases the port density but also facilitates the sharing of components. For example, four integrated WDM lasers can share one TEC instead of four.

Conclusions

We have reviewed various architectures for 400 GbE based on advanced modulation formats including 16 channels × 25 Gb/s NRZ, 8 channels × 50 Gb/s duobinary and PAM-4, 4 channels × 100 Gb/s PAM4, hybrid CAP-16/QAM-16, and QAM-16-OFDM. We have shown that all 25 Gb/s and 50 Gb/s per optical lane schemes considered can be supported via DMLs or EMLs. However, only hybrid CAP-16/QAM-16 and QAM-16-OFDM can support 100 Gb/s per optical lane with DMLs. In contrast, EMLs can enable all advanced schemes considered to transmit 100 Gb/s on a single optical lane. The 16 × 25 Gb/s NRZ scheme offers the best optical power budget margin, but it represents the most power consuming solution. For the other schemes, even though FEC is needed, significant power efficiency improvement is observed compared to NRZ. For digital implementations it was found that PAM-4 has the least digital complexity, while hybrid CAP-16/QAM-16 and QAM-16-OFDM exhibit about five and six times the digital complexity of PAM-4, respectively.

Looking ahead, the demand for higher bit rates will continue. Conventional NRZ will become even more demanding on the performance of electrical and optical components. However, advanced modulation formats, together with photonic integration, are likely to be able to address the challenges of further increases in lane rate.

Acknowledgment

This work was supported by the UK EPSRC via the INTERNET project.

References

[1] IEEE 802.3, http://www.ieee802.org/3.

[2] R. H. Derksen *et al.*, "Cost-Efficient High-Speed Components for 100 Gigabit Ethernet Transmission on One Wavelength Only: Results of the HECTO Project," *IEEE Commun. Mag.*, vol. 51, no. 5, May 2013, pp. 136–44.

[3] R. Rodes *et al.*, "100 Gb/s Single VCSEL Data Transmission Link," *OFC*, 2012, Paper PDP5D.

[4] J. Lee *et al.*, "Serial 103.125 Gb/s Transmission over 1 km SSMF for Low-Cost, Short-Reach Optical Interconnects," *OFC*, 2014, Paper PDP Th5A.5.

[5] T. Zuo *et al.*, "O-band 400 Gbit/s Client Side Optical Transmission Link," *OFC*, 2014, Paper M2E.4.

[6] J. C. Cartledge *et al.*, "100 Gbit/s Using Intensity Modulation and Direct Detection," *ECOC*, 2013, Paper We. 4. C. 3.

[7] T. Tanaka *et al.*, "Experimental Demonstration of 448-Gbps+ DMT Transmission over 30-km SMF," *OFC*, 2014, Paper M2I.5.

[8] J. L. Wei *et al.*, "Experimental Demonstration of Optical Data Links Using a Hybrid CAP/QAM Modulation Scheme," *Opt. Letters*, vol. 39, no. 6, Mar. 2014, pp. 1402–05.

[9] J. L. Wei *et al.*, "Study of 100G Ethernet Using Carrierless Amplitude/Phase Modulation and Optical OFDM," *J. Lightwave Tech.*, vol. 31, no. 9, May 2013, pp. 1367–73.

[10] B. Spinnler, "Equalizer Design and Complexity for Digital Coherent Receivers," *IEEE J. Sel. Topics Quantum Elect.*, vol. 16, no. 5, Oct. 2010, pp. 1180–92.

[11] X. Q. Jin *et al.*, "Real-Time Experimental Demonstration of Optical OFDM Symbol Synchronization in Directly Modulated DFB Laser-Based 25 km SMF IMDD Systems," *Opt. Express*, vol. 18, no. 20, Sept. 2010, pp. 21,100–10.

[12] J. J. Lee *et al.*, "Predication of TEC Power Consumption for Cooled Laser Diode Module," *LEOS*, 2004, Paper WW3.

[13] K. Nakahara *et al.*, "28-Gb/s Directly Modulated InGaAlAs ACPM DFB Lasers with High Mask Margin of 22% at 55°C," *OFC*, 2013, Paper OTh4H.3.

[14] J. Sewter *et al.*, "A 40 Gb/s Transversal filter in 0.18 μm CMOS using Distributed Amplifiers," *CICC*, 2005, pp. 417–20.

[15] Y. A. Vlasov, "Silicon CMOS-Integrated Nano-Photonics for Computer and Data Communications Beyond 100G," *IEEE Commun. Mag.*, vol. 50, no. 2, Feb. 2012, pp. s67–s72.

BIOGRAPHIES

JINLONG WEI [S'09, M'11] received his Ph.D. degree in electronic engineering from Bangor University, United Kingdom, in 2011. He joined the University of Cambridge, United Kingdom, as a research associate in August 2011 and worked on high-speed energy-efficient optical communication systems. He was awarded a Marie Curie grant and joined ADVA Optical Networking SE as a senior engineer in September 2014. He has authored/co-authored over 70 journal and conference publications, and holds 7 patents.

QIXIANG CHENG received his B.Sc. degree in opto-electronics from Huazhong University of Science and Technology, Wuhan, China in 2010. He is currently a Ph.D. student at the Centre for Photonic Systems, Department of Engineering, University of Cambridge. He has over 20 journal and conference publications, and his current research interests include energy efficient Datacommunicaitons and optical switching.

RICHARD V. PENTY [SM'10] is currently a professor of photonics and Master of Sidney Sussex College, University of Cambridge. He received his Ph.D. degree in engineering from Cambridge in 1989 and then was a SERC Information Technology Research Fellow. He has held academic posts at the University of Bath, United Kingdom, and the University of Bristol, United Kingdom. He has authored/co-authored more than 700 publications. He is Editor-in-Chief of the *IET Optoelectronics Journal* and co-founder of Zinwave Ltd and PervasID Ltd.

IAN H. WHITE [S'82, M'83, SM'00, F'05] is currently a Master of Jesus College, and deputy vice-chancellor and head of photonics research at the Engineering Department, University of Cambridge. He received his B.A. and Ph.D. degrees from the University of Cambridge in 1980 and 1984, respectively. He has published over 900 journal/conference papers and holds 28 patents. He is a Fellow of the Royal Academy of Engineering and the IEE. He is an Editor-in-Chief of *Electronics Letters*.

DAVID G. CUNNINGHAM is an Avago Technologies R&D Fellow. He received his Ph.D. in laser physics from Queens University, Belfast, Northern Ireland, in 1985. He has worked at British Telecommunications Laboratories, Hewlett-Packard Laboratories, Agilent Technologies, and Avago Technologies. The IEEE 802.3 and Fibre Channel (INCITS T11) committees have recognized his contributions to their 100 Mb/s to 32 Gb/s standards. He is currently interested in the application of advanced modulation schemes to multimode fiber links.

A West-East Bridge Based SDN Inter-Domain Testbed

Pingping Lin, Jun Bi, Stephen Wolff, Yangyang Wang, Anmin Xu, Ze Chen, Hongyu Hu, Yikai Lin

ABSTRACT

SDN [1] is considered to be a promising way to re-architect the Internet. However, the Internet is managed by owners of different administrative domains, so the centralized control model of SDN must be extended to account for inter-domain traffic. Thus, this article proposes a WE-Bridge mechanism to enable different SDN administrative domains to peer and cooperate. Based on WE-Bridge, we further designed two innovative inter-domain routing applications as use cases. To verify our design, we implemented the WE-Bridge and the two use cases by building an international testbed on which WE-Bridge, together with the two use cases, are deployed. The testbed is composed of four SDN networks: CERNET, Internet2 in USA, CSTNET, and SURFnet.

INTRODUCTION

Currently, the architecture of the network devices are closed systems. Closed systems do not favor network innovation, especially considering protocol development or network service evolution. SDN separates the network control plane from the network data plane, and moves it to a centralized software controller. In this article we treat a controller as network operating system (NOS) plus control applications above. The controller manages and controls the entire intra-domain information, including routes, bandwidth, and so on. Switches in the network are SDN switches that retain only the basic data forwarding function. Thus, SDN decouples the vertical and tightly coupled network architecture. At the same time, it standardizes forwarding technology in the data plane, and opens up the control plane and the associated protocols. Furthermore, the controller can run on a normal host or server to control data packet forwarding in SDN switches through a standardized protocol and an optional SSL (secure sockets layer) channel. In this way, all networking people, rather than only vendors, can contribute to SDN and promote the rapid innovation and the evolution of the network.

The idea of SDN is well received by academic researchers and industry researchers, network operators, and the networking industry. SDN is considered to be a promising way to re-architect networks. The Open Networking Foundation (ONF) is leading SDN standardization and has gained support from more than 100 companies who jointly accelerate the creation of standards, products, and applications, such as NEC, Google, IBM, and VMware.

PROBLEM STATEMENT

SDN works as a centralized control model. However, the Internet is managed by owners of different administrative domains, so the centralized control model of SDN must be extended to account for inter-domain traffic. An inter-domain protocol for SDN is necessary. Thus, this article proposes a West-East Bridge (WE-Bridge) mechanism to enable different SDN administrative domains to peer and cooperate. Beyond just achieving basic inter-domain routing, we leverage SDN to improve inter-domain routing by announcing domain-views containing rich/fine-granularity information/policies, to enable various inter-domain innovations based on network information.

SDN DOMAIN

The SDN domain in this article refers to the administrative SDN domain. One SDN domain may include multiple ASs (autonomous systems).

SDN already has some protocol implementations for the communication between the control plane and the data plane such as OpenFlow [2], NetOpen [3], and Grainflow [4]. OpenFlow is the most popular implementation; it has been deployed by many universities and research institutions around the world. To give readers a more concrete picture, we choose OpenFlow as an example to explain the whole article.

In the following sections this article rethinks inter-domain routing in SDN, describes the design of WE-Bridge, and the international SDN testbed infrastructure. Then we introduce two inter-domain applications — fine-granularity inter-domain routing and end-to-end QoS (Quality of Service) routing — to prove the feasibility of WE-Bridge. Finally, we conclude the article.

RETHINKING INTER-DOMAIN ROUTING IN SDN

SDN currently only changed how the intra-domain network works; it did not change the packet format but it changed intra-domain pack-

Pingping Lin, Jun Bi (corresponding author), Yangyang Wang, Anmin Xu, Ze Chen, Hongyu Hu, and Yikai Lin are with Tsinghua University.

Stephen Wolff is with Internet 2.

 0163-6804/15/$25.00

et forwarding. In theory, BGP (Border Gateway Protocol Version 4, RFC4271) still can be used between inter-domains. We have already successfully applied BGP between SDN and IP domains or between SDN and SDN domains [5], which are feasible and verified by implementation and deployments. However, by further analyzing the characteristic of SDN, we found that we can do more revolutionary innovations for the SDN inter-domain beyond just applying BGP:

• Besides forwarding based on a destination IP address as in legacy IP networks, the OpenFlow protocol extended more fields (source IP address, source MAC address, destination MAC address, port, etc.) to be matched during packet forwarding, which is called flow-based forwarding. Inspired by this, the inter-domain can also apply such multi-fields matching forwarding to achieve inter-domain fine-grain routing.

• SDN is centralized control; the controller can control intra-domain packet routing, and adjust the traffic status in the data plane in real time. In this way, the controller also has the ability to control how packets from other domains transit its domain. Many newly proposed routing policies such as local transit policies [6], which are impossible to achieve in the current distributed routing network, become possible in SDN.

• On the other side, BGP still has much room for improvement. Some of its shortcomings are: lacking QoS routing; lacking support for multipath routing [7]; limited policy expressiveness; and lacking path diversity. Therefore, simply applying BGP to the SDN inter-domain is not an ideal solution.

Inspired by the analysis above, we designed a new inter-domain mechanism named West-East Bridge (WE-Bridge) for inter-domain SDN, which is used for different SDN administrative domains to peer. WE-Bridge itself is not an inter-domain routing protocol, but a platform to exchange basic network information between different domains, and enable the third party to carry out SDN inter-domain innovations.

WE-Bridge

By providing the intra-domain network view to applications above the NOS, SDN promoted the intra-domain network protocol innovations. Similarly, to enable inter-domain innovations, WE-Bridge needs to exchange the basic inter-domain network information, and provide it to the applications above the NOS through a north bound API (application programming interface) as shown in Fig. 1. We design what information should be exchanged among domains, how to exchange such information in high performance, and the north bound interface for providing such information to the applications above.

We call the information that is needed to be shared among different domains a "virtual network view." WE-Bridge is designed to be compatible with different third-party NOSs and network view storage systems. In the following section we will give definitions to the network view, network view abstraction, and network view learning. Considering the network privacy and policy, each SDN domain may only want to expose part of its domain information to its peers, which urges us to further design the network view virtualization. Then we describe the data expression format for virtual network view during transfer among different SDN domains. After that we introduce the network view distribution, and the north bound API.

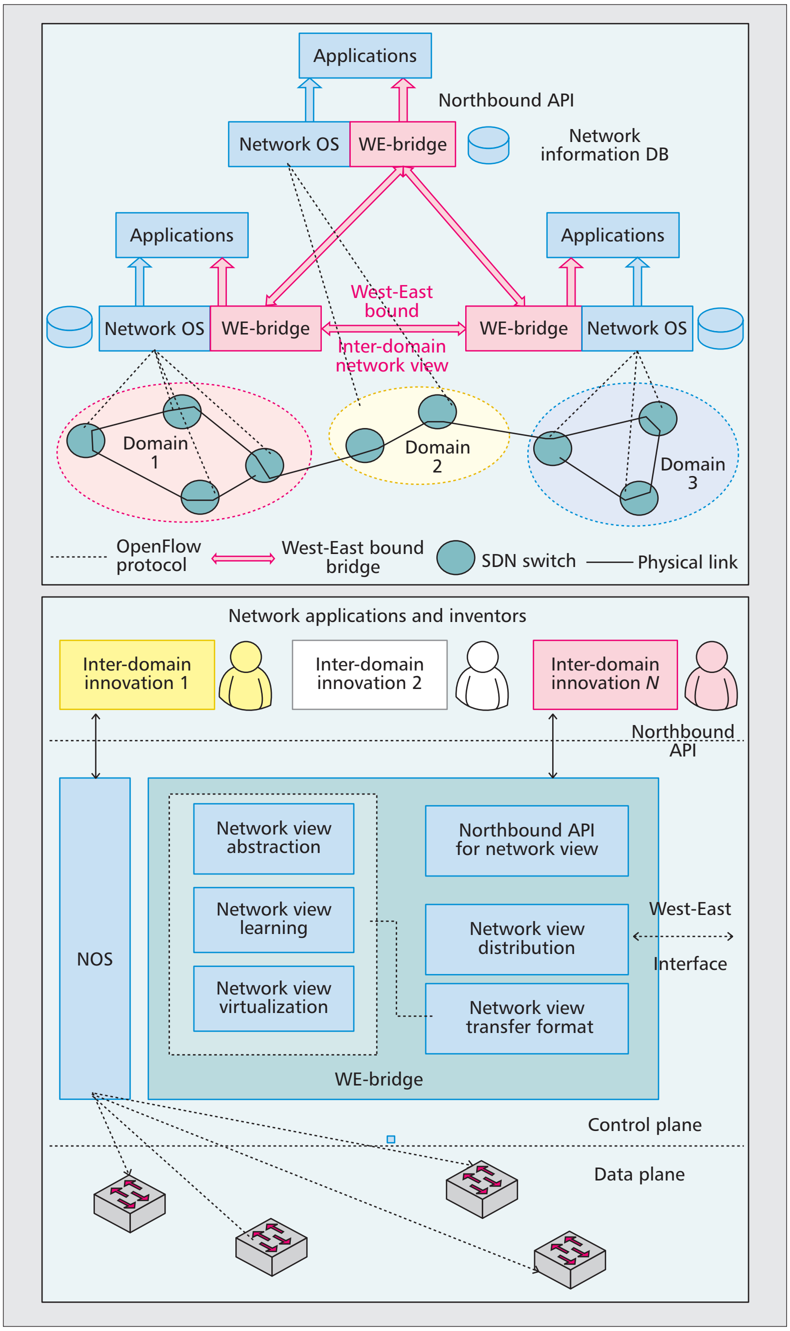


Figure 1. West-East Bridge for SDN inter-domain communication (top); overview of WE-Bridge (bottom).

Benefits from WE-Bridge:

• For network innovators: by providing the global basic network information to the applications above, WE-Bridge provides a platform for network researchers to innovate SDN inter-domain protocols;

• For the SDN network: WE-Bridge provides a way to enable multiple SDN inter-domain protocols to coexist, which is verified by the two use cases shown later. Multiple inter-

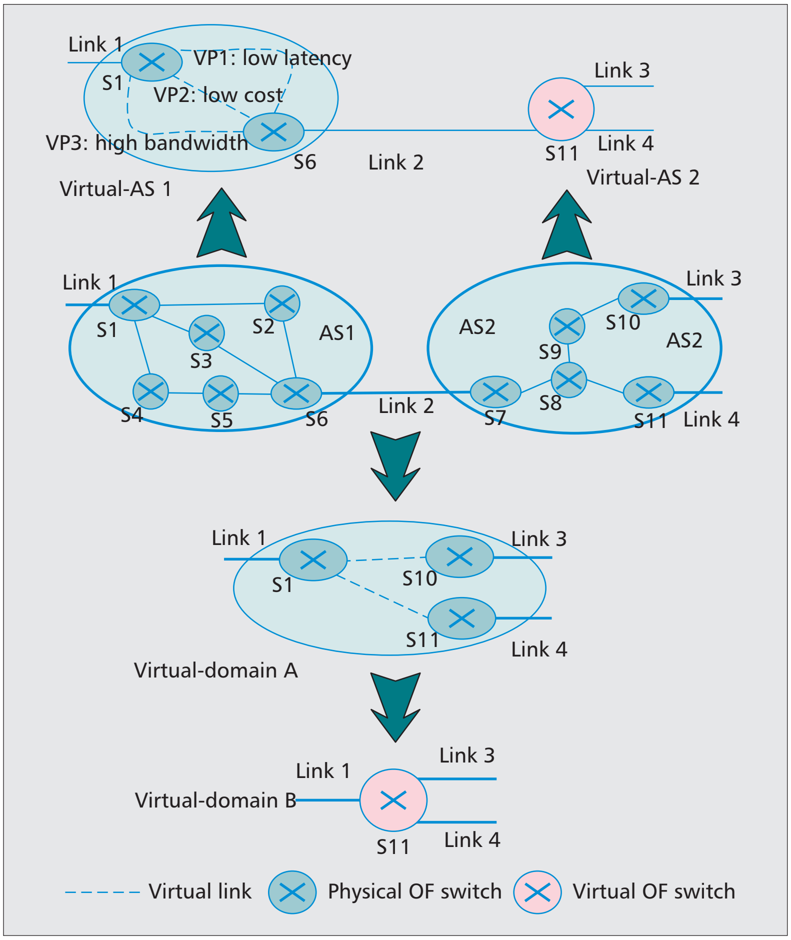


Figure 2. Network view virtualization (VP: virtual path; OF: OpenFlow; S: switch).

domain protocols can form a complementary or competitive relationship and promote the SDN inter-domain evolution.

Network View Abstraction

For the inter-domain circumstance, controlling the flow of data packets in a global network requires each controller to have a relative global network view to determine the next controller hop. Hence, NOSs are required to exchange reachability and topology information between inter-domain networks. We describe all the network entities with a network view and further divide it into two types: physical network view and virtual network view.

Network View — Both physical network view and virtual network view include two aspects: the network static information aspect and the dynamic information aspect. The network static aspect includes the following information:

- Reachability: IP addresses.
- Topology: nodes (e.g. switches, servers, hosts, controllers, firewalls, balancers, others), links, link attributes, port throughput, link connections.
- Network service capabilities, such as SLA (service level agreement), GRE (generic routing encapsulation), SSL (secure sockets layer), OpenFlow version, numbers of FlowTables in each switch, and how many flow entries each FlowTable supports.
- Forwarding capability parameters, such as latency, reliability, packet loss rate, availability, maximum throughput, time delay variation, and cost.

The network dynamic information mainly includes the network status, such as FlowTable entries information in each switch, real-time bandwidth utilization in the topology, and all the flow paths in the network.

Physical Network View Learning — Currently, LLDP (Link Layer Discovery Protocol) is used by controllers to discover the network topology. Usually, the controller in each SDN domain instructs each of the connected OpenFlow switches to send LLDP packets out from all the ports (the LLDP packet carries the source switch identity, out-port, and other capabilities). Once the neighbor switches receive the LLDP packets, they will send it directly to the controller. Then the controller abstracts and analyzes the information from the LLDP packet to determine if the source switch identity belongs to its domain and the LLDP packet received by a neighbor is the same as the one sent out from the source OpenFlow switch. If this is true, the controller will then create a direct intra-domain link between the source switch and this neighbor. For the inter-domain link, we extended LLDP in the NOS by adding a network view driver: if the source switch identity does not belong to its domain, then the controller can infer that this packet is from another domain, and will create an inter-domain link like (S6, S7) in Fig. 2 according to the source switch identity, source switch out-port, and the destination switch (who received the LLDP packet in its SDN domain) identity with the in-port. The inter-domain links should be stored in both the neighbor domains' local network views.

To learn more network view information such as OpenFlow version and number of the FlowTables on each node, link utilities, and flow entries, we extended the LLDP by adding an LLDP extension module. By counting the total number of packets related to a certain port in all the FlowTables in an OpenFlow switch, the LLDP extension can learn the link utilities. By the flow stats APIs defined in OpenFlow, the LLDP extension can learn the OpenFlow version, number of FlowTables, and flow entries in each switch.

Network View Virtualization

Each SDN domain may be willing to expose only a part of its domain information to its peers, rather than the entire network view, due to policy concerns. We can abstract a physical network view to a virtual network view in such situations. As shown in Fig. 2, we design three kinds of virtual network views:

- Abstract a physical network view into a virtual network view with only the edge switches, like AS 1 to virtual AS 1. Route path segments (like VP 1, VP 2, VP 3) from the ingress switch to the egress switch in the virtual network can have SLA (service level agreement)-level path attributes such as time latency, reliability, bandwidth, and packet loss rate.

- Abstract a physical network view into a virtual node, like AS 2 to virtual AS 2. The virtual node only retains three physical inter-domain (cross-domain) links: link2, link3, link4. After network virtualization, each NOS should store a mapping table between the physical network view and the virtual network view.
- In addition, one administrative domain may include several ASs. Assume AS1 and AS2 in Fig. 2 belong to the same administrative domain. We design to virtualize multiple ASs in one administrative domain to the same virtual domain like virtual domain A, or virtual domain B.

To compute an end-to-end or global routing path with QoS, the path computing application on the NOS needs the network views in other domains, or should at least know the abstracted virtual network views of other domains as shown in Fig. 2. After WE-Bridge exchanges the local virtual network views, each NOS can construct the global network view based on all the local virtual network views plus the inter-domain links and their attributes, and provide it to network applications above. Then the path computing application can compute an end-to-end path, cooperate with other peer networks to set up an end-to-end path, and send the cross-domain packet directly to the edge switch output port along the routing path.

Virtual Network View Transfer Format and Distribution

Different domains sharing network view (including the reachability information) requires the network view to be expressed in a manner WE-Bridges from all domains can understand. Thus, an application-independent language to enable peering between heterogeneous NOSs is needed such as JSON (JavaScript Object Notation), XML (eXtensible Markup Language), and YAML (YAML Ain't Markup Language). We chose JSON in our design since JSON is more light-weight.

After defining and expressing the network view, WE-Bridge needs to set up peer connections and to deliver the network view data. All SDN network controllers/peers are equivalent to each other, and they construct a peer-to-peer control plane. We design the network view distribution mechanism with the following principles. On one side, since the more connections each peer sets up, the more stable the peer control plane will be, each peer should set up as many connections as its resources can support (Principle 1). On the other side, the shorter the average hops in the peer control plane, the shorter the convergence time will be for network view update message delivery (Principle 2). Based on these two principles, we propose a maximum connection degree based connection algorithm [8].

By now the entire work-flow of WE-Bridge is as follows: WE-Bridge in each SDN domain collects and virtualizes the local physical network view. The reachability information of the domain can be configured by network operators, which is similar to the situation when network operators configure BGP routers with routes in the current IP network. Then all WE-Bridges from different SDN domains exchange the basic reachability (route) and virtual network information in JSON format, construct the relative global network view, and provide such information to the applications above. Then inter-domain routing applications in different domains translate all the routes into OpenFlow flow paths and cooperate with each other to set up cross-domain routing paths.

North Bound API

To promote SDN inter-domain innovation such as new inter-domain routing, we need to provide as many types of inter-domain information as we can to applications above the NOS/WE-Bridge in an easy-to-use manner. The north bound API is designed for such a purpose. After exchanging virtual network views by WE-Bridge, WE-Bridge in each network can construct a relative global network view and provide it to network applications above. However, to better serve the application above, WE-Bridge provides various types of network views, which mainly fell into two categories:

Original global network view: This includes local network view and virtual network views learned from all other domains.

Specific network view: after data processing to the original global network view, which is valuable for particular applications, such as the virtual network view data of a specific SDN domain or a set of SDN domains, the routes information of a specific SDN domain or a set of SDN domains, the topology information of a specific SDN domain or a set of SDN domains.

We design the following two approaches to pass the network views including both original global network view and specific network view from WE-Bridge to applications above the NOS:

Information subscribe/publication: applications above the NOS can register themselves with WE-Bridge for certain information requirements. Each time a WE-Bridge module receives a corresponding information event, it will notify all the applications who subscribed before.

North bound API: WE-Bridge provides all kinds of network information and provides it to the applications above by north bound REST (representational state transfer) API.

To promote SDN inter-domain innovation such as new inter-domain routing, we need to provide as many types of inter-domain information as we can to applications above the NOS/WE-Bridge in an easy-to-use manner. The north bound API is designed for s uch a purpose.

Testbed for Inter-Domain Innovation

We implemented the WE-Bridge, and for deployment, we built an international federal SDN testbed in July 2013, and deployed WE-Bridge to this testbed as shown in Fig. 3. The testbed includes four SDN networks: Internet2 (United States open national research and education network), CERNET (China Education and Research Network), CSTNET (China Science and Technology Network), and SURFnet (the national research and education network of the Netherlands). WE-Bridge successfully connected those four SDN networks. The testbed

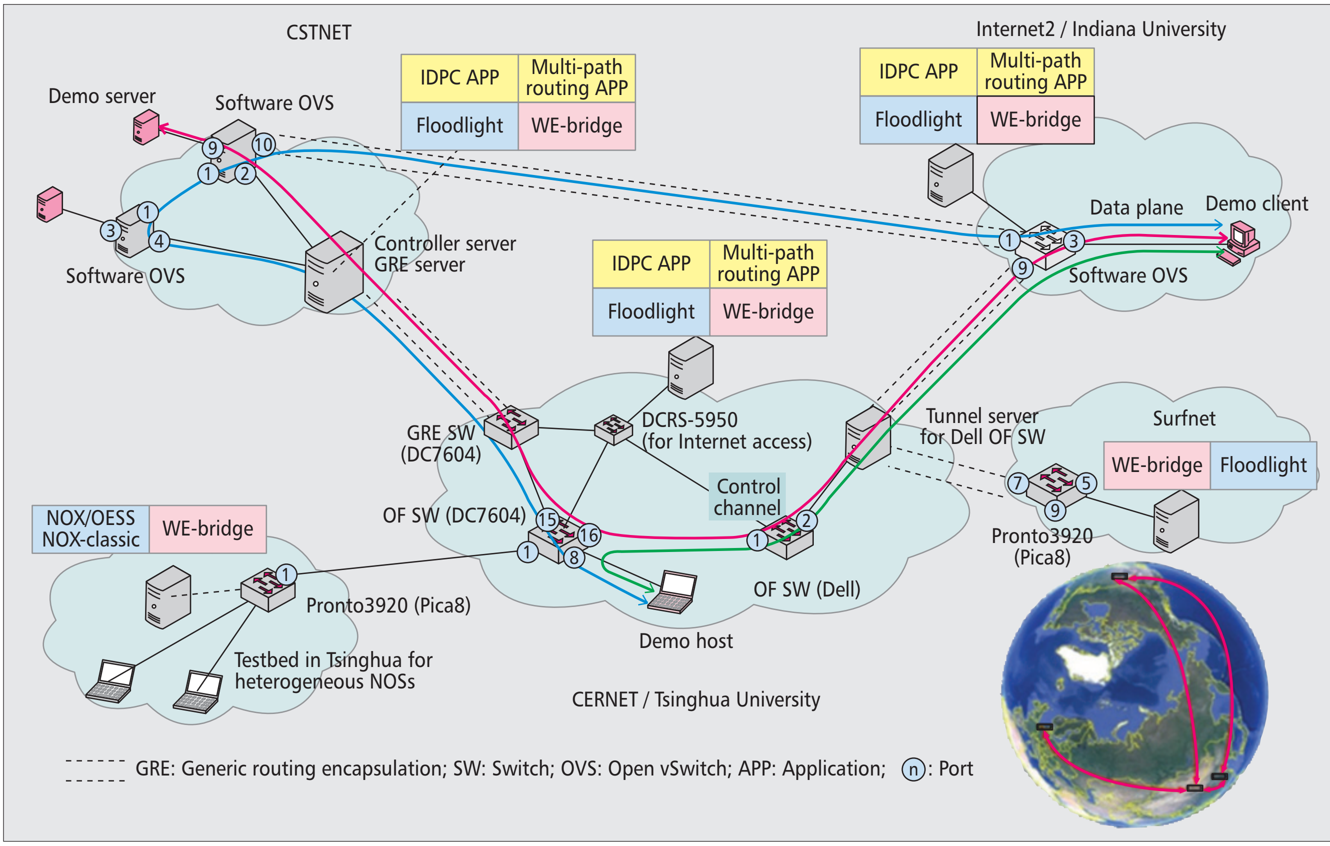


Figure 3. Global SDN Federated Testbed: Internet2 (United States open national research and education network), CERNET (China education and research network), CSTNET (China science and technology network), and SURFnet (the national research and education network of the Netherlands).

has been running since last summer as an international collaboration project and users in CSTNET and CERNET have kept using it as a production network for trans-Pacific genomic data transfer.

Hardware Configuration

We deployed three hardware OpenFlow switches (DC7604 from Digital China, Pronoto3920 from Pica8, one Dell OpenFlow switch) in CERNET, one hardware GRE (generic routing encapsulation) switch (DC7604 from Digital China), and one aggregation switch (DCRS-5950 from Digital China). There are two Open vSwitches (software switch) [9], one GRE server, two demo servers in CSTNET. Internet2 also has an Open vSwitch, one server running the controller, and one host for the traffic test. SURFnet used one hardware switch Pronoto3920 form Pica8.

Software Configuration

The controller used in this implementation is Floodlight [10], and the software switch used is Open vSwitch [9]. Each network deployed a WE-Bridge. The NOX/NOX-Classic/OESS [11, 12] in Tsinghua are used for trying out heterogeneous NOS peering with WE-Bridge. Currently, the four SDN testbeds in CERNET, CSTNET, Internet2, and SURFnet are not neighbors in layer 2. Therefore, we connected those four OpenFlow testbeds by GRE tunnels shown in Fig. 3. Then from the viewpoint of each SDN network, all four testbeds are neighbors.

Two Use Cases

Based on such a testbed and to verify the feasibility of the WE-Bridge platform, we further designed and carried out two innovative inter-domain routing applications as two use cases to the SDN inter-domain routing:

- Source-address based multi-path routing. Based on a path vector algorithm like BGP, we apply the flow based routing into SDN inter-domain by adding more fields (such as Ether type, source IP address) to the route announcement message to achieve fine-granularity inter-domain routing.
- Inter-domain path computation (IDPC). Considering the path attributes such as bandwidth and time latency, we let different domains negotiate the path attributes to achieve end-to-end QoS routing. Both of them are implemented as network applications running on NOS.

We also implemented those two applications, and deployed them to the testbed. The first use case (fine-granularity multi-path routing) achieves the global infrastructure connection. Then WE-Bridge in each network exchanges the reachability and topology information with its peers, and the global network view is constructed. Then, inter-domain path computation applications in each domain cooperate together and set up end-to-end paths. Such deployment to the real SDN network proves the feasibility of WE-Bridge and its ability to enable inter-domain innovations.

Use Case 1: Source-Address Based Multi-Path Routing — This use case presents a new fine granularity inter-domain routing, which supports multipath inter-domain routing to the same destination based on different source IP addresses.

As shown in Fig. 3, each network works as an AS. CERNET at Tsinghua university sets host IP addresses 101.6.30.100 and 101.6.30.101, CSTNET has a block of IP addresses denoted as 159.226.61.*, and Internet2 has a host with IP address 149.165.130.212.

AS4 (Internet2) announces three routes with the same destination IP address with different source addresses, while Tsinghua announces two routes as shown in Fig. 4 (a). All the routes are propagated along AS paths indicated respectively by the green, red, and yellow arrows. Thus, the flow matching (src: 101.6.30.100, dst: 149.165.130.212) will traverse AS path (AS1, AS3, AS4), and the flow matching (src: 101.6.30.101, dst: 149.165.130.212) will be forwarded along AS path (AS1, AS4). All other flows matching (dst: 149.165.130.212) take the AS path (AS1, AS4). We used the ping command to ping the demo client address 149.165.130.212 in Internet2 from different source addresses 101.6.30.100 and 101.6.30.101 on the demo host in Tsinghua. Then the two traffic flows are routed along two different AS-level paths shown with the blue and green color lines in Fig. 3. Figure 4b presents a similar case with one more SDN network, SURFnet. The routes that traverse AS path (AS4, AS1) in Fig. 4a change to traverse AS path (AS4, AS2, AS1) in Fig. 4b.

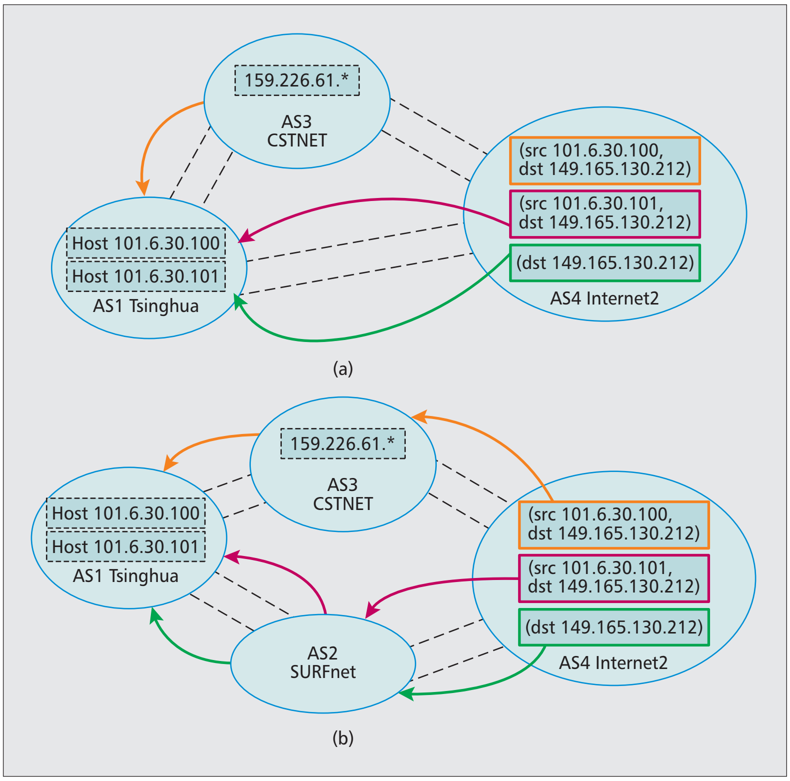


Figure 4. A concept illustration of source address based multipath routing.

Use Case 2: Inter-Domain Path Computation — Inter-domain path computation (IDPC) application is also developed and installed as an application. The purpose of IDPC is to achieve end-to-end QoS routing. After WE-Bridge modules in all the domains exchange local virtual network views including the bandwidth information, each domain can construct a relative global network view.

According to the specification of OpenFlow, each time there is a new packet coming and if there is no rule for this packet in the switch, this packet will be transferred to the controller. Then the IDPC application reads the global virtual view information provided by WE-Bridge and judges the location of the destination IP: whether the destination IP of an incoming data flow is located in the intra-domain or in other domains. If the destination IP is in the intra-domain, it will carry out re-active flow calculation and installation. If the destination IP is in other domains, it will compute an end-to-end path according to the global network view, and negotiate with other IDPC applications along the path with path segment request (ingress switch and port, egress switch and port, matching fields, path attributes) to set up an end-to-end path with QoS attribute. Then each IDPC application translates the corresponding path segment request into flow entries and installs them to the OpenFlow switches. Then an end-to-end path (satisfying the QoS requirement of user traffic) is set up. In this use case the QoS requirement is simplified as the minimal bandwidth requirement of user traffic.

The bandwidths of the GRE tunnel link between Tsinghua and Internet2 and between CSTNET and Tsinghua are larger than that between CSTNET and Internet2. Therefore, the flow tables in each switch for the traffic will be set up automatically by IDPC applications as follows: CSTNET->Tsinghua->Internet2.

At last, we delivered one Terabyte of trans-Pacific genomic data (FTP: File Transfer Protocol) from the server in CSTNET to the client in Internet2. We monitored the traffic on each interface of the OpenFlow switches along the routing path by SNMP (Simple Network Management Protocol). As shown in Fig. 5, we can see the average speed of the traffic is 20 MB/s.

The sizes of the virtual network views of CSTNET, CERNET, and Internet2 are 452 bytes, 486 bytes, and 319 bytes. Compared with the traffic data, the bandwidth occupied by the control plane data is very small. So we are able to conclude that WE-Bridge is a lightweight solution to enable different SDN domains to cooperate.

Related Work

To enable incremental deployment of SDN, a research study of SDN-IP network peering [5] was conducted by us in 2013. This work focuses on the interaction between BGP-based SDN domain and legacy IP domain. This solution applies BGP between the SDN and IP domains or between SDN and SDN domains without improving the inter-domain routing.

RouteFlow [13] is one of the first implemen-

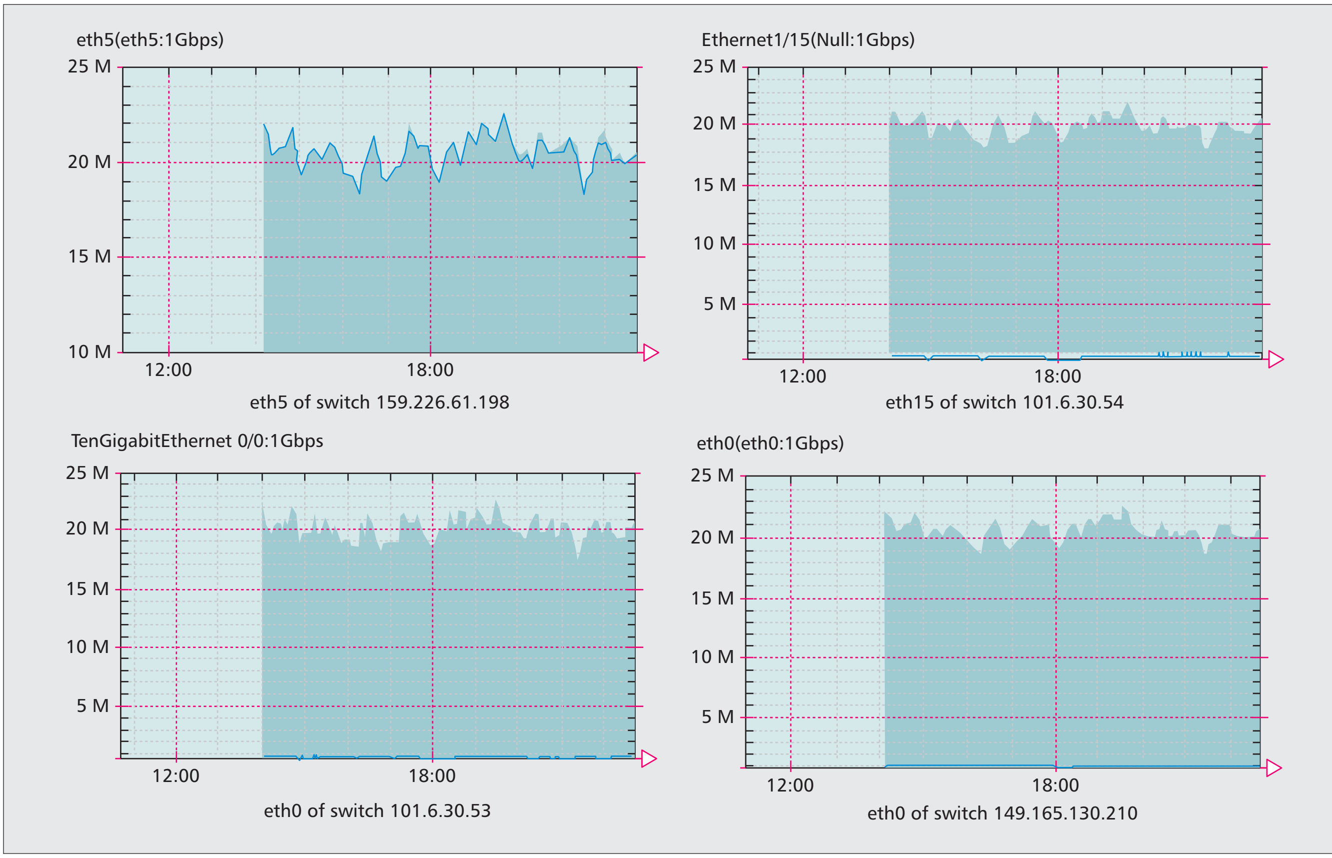


Figure 5. Traffic at interfaces of OpenFlow switches along the routing path.

tations of IP routing on OpenFlow switches. RouteFlow instantiates a VM for each OpenFlow switch with as many virtual network interfaces as there are active ports in the corresponding device, and runs a stack of open-source routing protocols on the virtual topology. All control messages are exchanged between VMs as if they are running as a distributed control plane. The routing engine is still based on traditional open source software like Quagga [14]. Inter-domain routing is still standard BGP. Such a solution incurs the overhead of distribution without the benefits of scale.

Feamster *et al*. [15] presented a software defined Internet exchange (SDX) to bring SDN enabled features to today's BGP inter-domain routing. This article aims to take advantage of Internet exchange points (IXPs) and to enable more expressive policies than conventional destination IP address based forwarding with BGP. However, domains that are not connected to IXPs cannot benefit from this.

SDN inter-domain peering is currently still a challenge. No single solution has been successfully deployed in large scale yet.

Conclusion

This article designs a West-East Bridge for SDN inter-domain peering. We first defined what network information can be exchanged and how such information is efficiently exchanged among inter-domain SDN peers. With this technology, we built an international SDN testbed with four SDN networks: CERNET, CSTNET, Internet2, and SURFnet. To verify the WE-Bridge platform, we further designed, implemented, and deployed two SDN inter-domain routing innovations. From the two innovations, we proved it is easy to do innovations and implementations on WE-Bridge. Some new features of inter-domain that cannot be achieved on current networks can be achieved in the SDN environments with the WE-Bridge platform, such as fine-granularity inter-domain routing and end-to-end QoS routing, as shown in the two use cases. WE-Bridge has great potential to enable SDN inter-domain innovations.

In the future, we plan to expand our SDN testbed by connecting more SDN networks with WE-Bridge and attracting more real inter-domain traffic.

Acknowledgments

We are grateful for the valuable feedback and discussions with Steve Wolff's team members in Internet2: Mike Sullivan, Scott Brim, Dale Finkelson, Edward Moynihan, Wendy Huntoon, and also Frank Gao and Tony Zhang from the National Center for Biotechnology Information of USA. We thank the testing team: Yang Yu from Tsinghua and Dr. Yulei Wu from CSTNET. We thank the infrastructure team: Yayuan Sun and Prof. Jilong Wang's operation team in Tsinghua, Dr. Yuepeng E from CSTNET, John Hicks and Brent Sweeny from Indiana University, and Ronald van der Pol from SURFnet. We also thank the application participators: Jianning

Chen from CSTNET and Shaoping Ling from Beijing Institute of Genomics. Finally, we thank Di Lu of Internet2 for organizing all the international architecture meetings. This work is supported by the National High-Tech R&D Program ("863" Program) of China (Grant No. 2013AA013505), the National Science Foundation of China (Grant No. 61472213), and the CANS Future Internet Working Group (https://china.internet2.edu/workinggroups/sdnfuture-internet-working-group/).

REFERENCES

[1] N. McKeown, "Keynote talk: software-defined networking," *Proc. IEEE INFOCOM*, 2009.
[2] N. McKeown *et al.*, "OpenFlow: Enabling Innovation in Campus Networks," *SIGCOMM Computer Communication Review*, vol. 38, no. 2, March 2008, pp.69–74.
[3] N. Kim and J. Kim, "Building Netopen Networking Services over Open Flow-Based Programmable Networks," *Proc. Information Networking (ICOIN)*, 2011.
[4] Z. Liu *et al.*, "Grainflow: A Per-Bit Customizable Scheme for Data Plane Innovation on Programmable Hardware," *Proc. ACM Conf. emerging Networking EXperiments and Technologies (CoNEXT) Student Workshop*, 2011.
[5] P. Lin *et al.*, "Seamless Interworking of SDN and IP," Demo, *Proc. SIGCOMM 2013*.
[6] P. B. Godfrey *et al.*, "Pathlet Routing," *Proc. SIGCOMM 2009*.
[7] H. T. Kaur *et al.*, "BANANAS: An Evolutionary Framework for Explicit and Multipath Routing in the Internet," *Proc. Future Directions in Network Architecture*, 2003.
[8] P. Lin, J. Bi, and Y. Wang, white paper: "SDN Interdomain Network Peering," http://netarchlab.tsinghua.edu.cn/~junbi/THU-NetArchLab-SDN-TR-WEBridge-20131001.pdf.
[9] Open vSwitch: http://openvswitch.org/.
[10] Floodlight: http://www.projectfloodlight.org/floodlight/.
[11] N. Gude *et al.*, "NOX: Towards an Operating System for Networks," *SIGCOMM Computer Communication Review*, vol.38, no.3, July 2008, pp.105–10.
[12] OESS: http://globalnoc.iu.edu/sdn/oess.html.
[13] M.R. Nascimento *et al.*, "Virtual Routers as a Service: The Routeflow Approach Leveraging Software-Defined Networks," *Proc. 6th Int'l. Conf. Future Internet Technologies*, 2011.
[14] Quagga: http://www.nongnu.org/quagga/.
[15] N. Feamster *et al.*, "SDX: A Software Defined Internet Exchange," Open Network Summit (ONS) 2013.

BIOGRAPHIES

PINGPING LIN (pingpinglin2012@gmail.com) received her B.S. degree in computer science from the University of Electronic Science and Technology of China (UESTC) in 2008. She received her Ph.D. degree in computer science from Tsinghua University in 2014. The work in this article was carried out during her Ph.D. research. Currently she is working at Open Networking Lab in the US. Her research interests include evolvable Internet architecture, source address validation, and software defined networking.

JUN BI (junbi@tsinghua.edu.cn) received B.S., M.S., and Ph.D. degrees in computer science from Tsinghua University. Currently he is a full professor and the director of the Network Architecture and IPv6 Research Division, Institute for Network Sciences and Cyberspace, Tsinghua University. He is a senior member of IEEE and ACM, co-chair of the AsiaFI (Asia Future Internet Forum) steering group, and co-chair of the CANS Future Internet/SDN working group. His research interests include network architecture and Future Internet.

STEPHEN WOLFF received a Ph.D. in electrical engineering from Princeton University. He is a life member of IEEE, a member of the American Association for the Advancement of Science (AAAS), of the Association of Computing Machinery (ACM), and of the Internet Society from whom he received the Postel Award in 2002. He was elected to the Internet Hall of Fame in 2013. He is currently the interim vice president & chief technology officer of Internet2.

YANGYANG WANG received his B.S. degree in computer science and technology from Shandong University, Jinan, China, in 2002, his M.S. degree from Capital Normal University, Beijing, China, in 2005, and his Ph.D. in computer science from Tsinghua University in 2013. He is currently a postdoctoral scholar in computer science at Tsinghua University. His research interests include Internet routing architecture, future Internet design, and SDN.

ANMIN XU is a Ph.D. student in computer science at Tsinghua University.

ZE CHEN is a master's student in computer science at Tsinghua University.

HONGYU HU received her Ph.D. degree in computer application major from Beijing Institute of Technology, Beijing, China, in 2009. She worked as a post-doctor at Tsinghua University from 2009 to 2012. She is now a teacher at Tsinghua University. Her current research interests include SDN, OpenFlow, and Internet architecture.

YIKAI LIN is a research intern at Tsinghua University. He is an undergraduate student in communication engineering at Beijing University of Posts and Telecommunications (BUPT).

Some new features of inter-domain that cannot be achieved on current networks can be achieved in the SDN environments with the WE-Bridge platform, such as fine-granularity inter-domain routing and end-to-end QoS routing, as shown in the two use cases. WE-Bridge has great potential to enable SDN inter-domain innovations.

Accepted from Open Call

5th Generation Mobile Networks: A New Opportunity for the Convergence of Mobile Broadband and Broadcast Services

Jordi Calabuig, Jose F. Monserrat, and David Gómez-Barquero

Abstract

This article analyzes the challenges and opportunities that the upcoming definition of future 5G mobile networks brings to the mobile broadband and broadcast industries to form a single converged network. It reviews the state-of-the-art in mobile and broadcast technologies and the current trends for convergence between both industries. This article describes the requirements and functionalities that the future 5G must address in order to make an efficient and flexible cellular-broadcasting convergence. Both industries would benefit from this convergence by exploiting synergies and enabling an optimum use of spectrum based on coordinated spectrum sharing.

Introduction

The mobile communications sector is characterized by a rapid worldwide increase in traffic demands due to the continuously evolving requirements and expectations of both users and operators. In 2013 global mobile data traffic grew 81 percent, and it is expected that this will increase nearly 11-fold between 2013 and 2018 [1], primarily driven by the increasing usage of mobile video services on devices with large screen sizes such as smartphones and tablets.

The most representative video service is mobile TV, which is often identified with linear TV and broadcast (point-to-multipoint) distribution. However, the behavior of mobile users is different from traditional fixed TV because content is mainly consumed on-demand with unicast (point-to-point) connections. The convergence of linear TV and on-demand content represents a challenge that requires a combined broadcast/unicast delivery model.

The mobile and broadcast industries have both developed several mobile broadcast technologies to support large-scale consumption of mass multimedia services on mobile devices. So far, the adoption of mobile TV services has not fulfilled the initial expectations due mainly to the lack of a successful business model and the high costs associated with the deployment of new mobile broadcasting networks. However, both industries are currently taking steps toward possible broadcast solutions to provide mobile TV services.

On the one hand, the mobile industry has adopted Long Term Evolution (LTE) and its evolution LTE-Advanced (LTE-A) [2] as worldwide Fourth-Generation (4G) cellular technologies. Both LTE and LTE-A support broadcast transmissions by means of the evolved Multimedia Broadcast Multicast Services (eMBMS) [3], which is commercially known as LTE Broadcast. eMBMS allows mobile network operators (MNOs) to cope, to some extent, with the increasing demand of mobile video data using point-to-multipoint transmissions. As this broadcast solution is supported over the same frequency carrier as unicast services, the eMBMS deployment costs are significantly lower than other broadcast alternatives. However, it also entails the use of dense networks compared with terrestrial broadcast networks and the reduction of system capacity for unicast services, which hampers the eMBMS business model.

Nowadays, while MNOs are investing in 4G network deployments, the mobile industry is already working on the Fifth-Generation (5G), the future mobile communications beyond 2020. The three main requirements for 5G wireless networks are [4]: to support massive capacity and connectivity; to carry a diverse set of services, applications and users with extremely diverging requirements; and to make flexible and efficient use of the available spectrum, whether contiguous or not, supporting wildly different network deployment scenarios.

Concerning the broadcast industry, traditional digital terrestrial TV (DTT) networks are more efficient for the delivery of very popular data, such as live video services, at high quality to a large number of subscribers. While tradi-

The authors are with Universitat Politècnica de València.

0163-6804/15/$25.00 © 2015 IEEE

tional broadcasting systems use high power high tower (HPHT) networks to cover large areas, cellular networks need thousands of base stations to cover the same area.

Moreover, the growing interest of the broadcast industry to enable mobile TV services on smartphones and especially tablets has renewed the interest in mobile broadcasting after the failure of first-generation mobile broadcast technologies such as DVB-H (Digital Video Broadcasting Handheld) in Europe, Media FLO and ATSC-M/H (Advanced Television Systems Committee Mobile/Handheld) in North America, or CMMB (China Mobile Multimedia Broadcasting). Furthermore, broadcasters can benefit from the capacity and coverage improvements brought by such new-generation broadcast standards as the Second-Generation Terrestrial DVB (DVB-T2), DVB Next Generation Handheld (DVB-NGH), or the upcoming ATSC 3.0.

So far, the highly fragmented DTT market has demotivated manufacturers from including broadcast TV capability on mobile devices. However, the current interest in developing a worldwide new-generation broadcast terrestrial standard targeting both fixed and mobile reception could favor economies of scale, thus allowing for its implementation on mobile devices [5].

The scenario in case the broadcast industry agrees on a worldwide broadcast terrestrial standard and the mobile industry includes a multicast/broadcast transmission mode in the future 5G system, would end up with two completely different industries, with different network infrastructures and business models competing for market and spectrum.

This article discusses a different approach by which the definition of the future 5G mobile broadband communication system could bring together the cellular and broadcast industries to form a single fixed and mobile converged network. The next section presents the state-of-the-art in the evolution of mobile and broadcast technologies. Afterward, the first approaches for the cooperation between mobile broadband and broadcast are described. They are the starting point for the definition of a joint solution within 5G. Finally, the challenges of the definition of such a converged solution are discussed.

Evolution of Mobile Communications and Terrestrial Broadcasting Standards

Figure 1 shows the mobile communications and DTT standards landscape. On the one hand, the mobile industry evolves toward the future 5G mobile broadband communications, whereas the broadcast industry evolve toward a global DTT standard.

Mobile Industry: Evolution Toward the Future 5G System

From the second-generation (2G) of mobile cellular networks, the mobile communications sector has seen many competing radio standards mainly developed by the 3GPP, the 3GPP2, and the IEEE. However, the dynamics of 4G are changing this landscape since all existing mobile cellular systems are converging toward a single technology, i.e. LTE. LTE was designed from the beginning with the goal of evolving radio access technologies under the assumption that all services would be packet-switched. Unlike previous technologies, LTE adopted Orthogonal Frequency Division Multiplexing (OFDM) as its radio access technology, which is the standard dominating the latest evolution of all mobile radio standards. This change was accompanied by an evolution of the non-radio aspects of the complete system toward a flat and all-IP system architecture. It is worth noting that the latest cellular generation developed by the IEEE, known as 802.16m, has similar targets as the evolution of LTE, i.e. LTE-A. However, the IEEE 802.16 family has not been designed with the same emphasis on mobility and compatibility with operators' core networks as the 3GPP technology family.

Concerning the state of the art on 4G network deployments, the number of commercial LTE networks and handsets is increasing very rapidly worldwide. By the end of 2014, 360 LTE networks have commercially been launched worldwide and this number is expected to reach around 450 networks by the end of 2015 [6]. This worldwide adoption gives to eMBMS, the broadcast mode of LTE, the edge over all its mobile broadcast competitors to provide mass multimedia services to mobile devices. The main advantages of eMBMS are the end-to-end IP architecture that enables the coexistence of unicast and broadcast services with high capacity, high bandwidth, and high scalability. Moreover, the deployment costs are significantly lower than other broadcast alternatives due to its easy integration with LTE infrastructure and mobile device chipsets.

Concerning LTE Broadcast, mobile users have seen some milestones related to eMBMS deployments in 2014 [6]. Several MNOs have already deployed their eMBMS networks, successfully completing the first trials by the end of January 2014. As examples, Verizon Wireless used the Super Bowl in New York as a test case for the LTE Broadcast technology, and the Australian operator Telstra completed a live demonstration of LTE Broadcast in a stadium environment during a cricket match in Melbourne. In addition, Korea Telecom completed the world's first commercial launch of LTE Broadcast services using eMBMS technology and, later on, Vodafone Germany and KPN conducted Europe's first trials of LTE Broadcast in a football stadium.

Although MNOs are currently investing in 4G network deployments, the mobile industry is already working on the definition of the future 5G mobile communication system. Following the current requirements and expectations from both users and operators, it is expected that future 5G wireless networks will also include the efficient provision of mass mobile multimedia services through one or several broadcast transmission modes.

Following the current requirements and expectations from both users and operators, it is expected that the future 5G wireless networks will also include the efficient provision of mass mobile multimedia services through one or several broadcast transmission modes.

The worldwide adoption of LTE and LTE-A as 4G technologies gives eMBMS a head start over all its mobile broadcast competitors to provide mobile TV services to smartphones and tablets. eMBMS is currently only supported in a mixed carrier mode, where broadcast and unicast data share the carrier capacity.

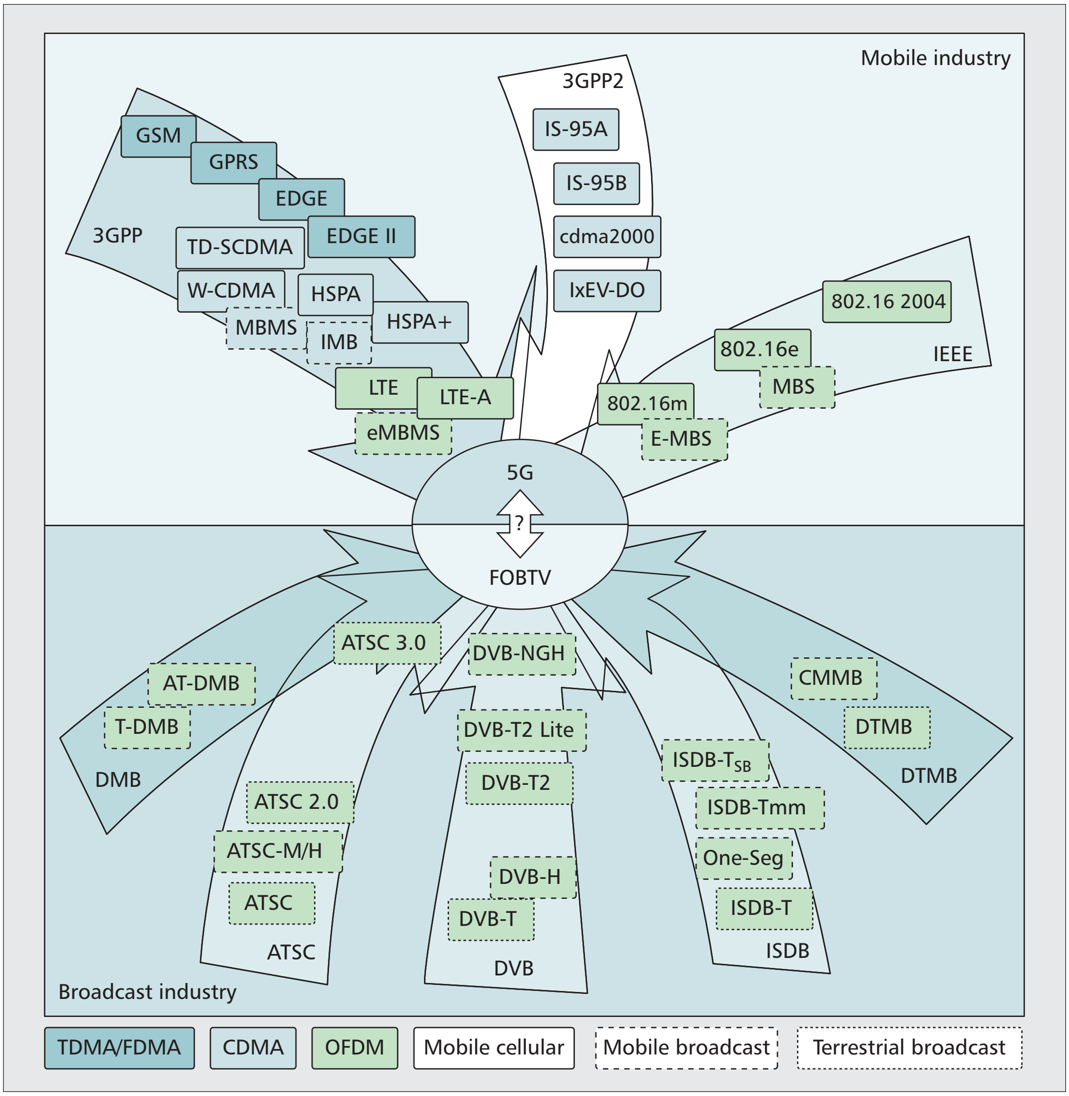


Figure 1. Mobile communications and DTT standards landscape.

Broadcasters: Evolution Toward a Global Terrestrial Standard

The need for replacing analogue TV to make a more efficient use of the broadcast radio spectrum and to improve the quality of TV services motivated the development of DTT systems [7]. There are five digital TV standardization organizations in the world: ATSC in North America; DVB in Europe; Digital Terrestrial Multimedia Broadcast (DTMB) in China; Integrated Services Digital Broadcasting (ISDB) in Japan; and Digital Multimedia Broadcasting (DMB) in Korea.

DVB standards are the most widely adopted worldwide; currently two of these standards, DVB-T2 and DVB-NGH, are the most advanced DTT systems for fixed and mobile devices, respectively. DVB-T2 provides at least 50 percent capacity increase over the other existing standards and its evolution to handhelds, DVB-NGH, includes the use of multiple antennas at the transmitter and receiver side and other enhancements with respect to DVB-T2.

Despite the fact that DVB-NGH outperforms previous DTT systems to support large-scale consumption of mass multimedia services in mobile devices, there are no plans for its commercial deployment. So far, in many countries, mobile TV services either did not reach the market or were launched and promptly stopped. The only successful stories among the mobile broadcasting systems are One-Seg in Japan and Terrestrial-DMB (T-DMB) in South Korea. However, in these two cases the transmission of mobile broadcasting services is requested by the national regulators.

The high fragmentation of the DTT market and the need to benefit from the economies of scale have recently motivated the approaches and wishes of cooperation among the different broadcasting standard developing organizations. The Future of Broadcast TV (FOBTV) initiative was launched in 2012 with the aim of creating a common working scenario for the future generation of broadcast terrestrial systems to avoid competing standards, overlap, and inefficient deployment of new services [5]. For example, since mobile devices are likely to move across borders, it is highly desirable that the specification contains core technologies that will have broad international acceptance and enable global interoperability.

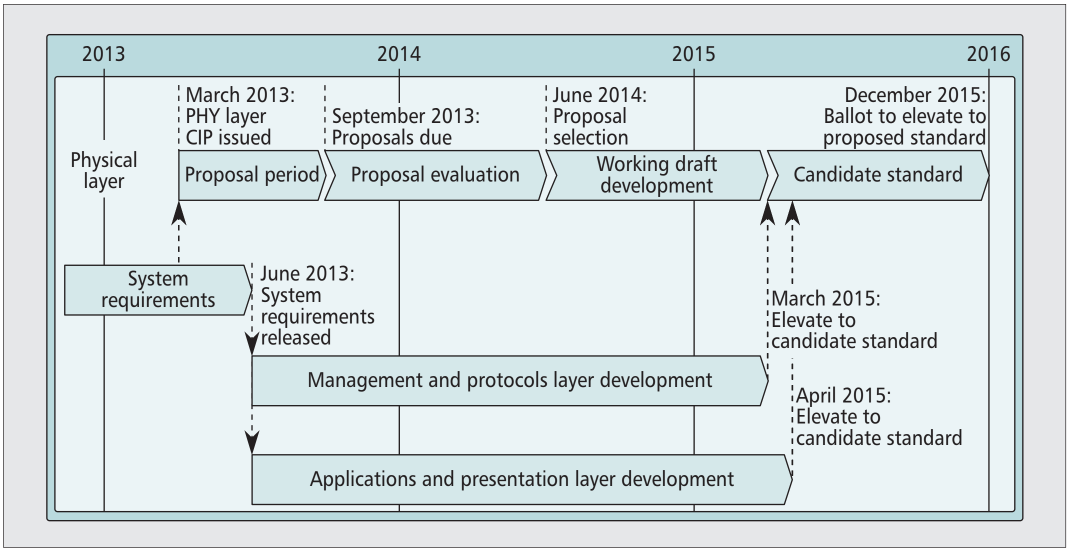


Figure 2. ATSC 3.0 standard development timeline.

Traditional broadcast networks are more efficient for broadcasting TV services to a large number of subscribers due their HPHT networks, which cover large areas. Typical SFN distances used in traditional broadcast networks range between 50 and 90 km.

The latest milestone in the development of next-generation digital broadcast technologies is the ATSC 3.0 standard, which will replace the current systems used in the United States and aims at being a global standard. Figure 2 shows the development timeline of ATSC 3.0 [8], which involves the physical layer, as well as the management and protocols, and applications and presentation layers. In September 2013, 11 detailed physical layer proposals were received [9], taking many of these technical proposals in the DVB-T2 and/or DVB-NGH standards as a baseline. Therefore, it is likely that the upcoming physical layer of the ATSC 3.0 will have significant similarities with the last DVB standards.

First Attempts for the Mobile Broadband-Broadcast Cooperation

The worldwide adoption of LTE and LTE-A as 4G technologies gives eMBMS a head start over all its mobile broadcast competitors to provide mobile TV services to smartphones and tablets. eMBMS is currently only supported in a mixed carrier mode, where broadcast and unicast data share the carrier capacity. In particular, up to 60 percent of the total LTE resources can be reserved for eMBMS. The subframes are named MBMS over single frequency network (MBSFN) subframes, and use ODFM symbols with a longer cyclic prefix (CP) of 16.67 μs. The use of this longer CP allows the construction of SFNs between multiple cells with a maximum of 5 km inter-site distance. For MBSFN dedicated carrier deployments, it is possible to double the CP length to 33.33 μs (10 km SFN distance). Although the physical-layer features of MBSFN dedicated carrier are already defined, it is not supported in current releases of LTE and LTE-A.

From the broadcasters' point of view, current eMBMS features do not address all mobile broadcast use cases, since its low delay spread tolerance entails that eMBMS only can be deployed on dense networks. Traditional broadcast networks are more efficient for broadcasting TV services to a large number of subscribers due their HPHT networks, which cover large areas. Typical SFN distances used in traditional broadcast networks range between 50 and 90 km.

The next section gives an overview of possible ways of complementing wireless broadband systems using terrestrial broadcast networks.

Technical Approaches for 3GPP and DVB Cooperation

In November 2010 the DVB forum contacted 3GPP to consider a potential collaboration in the area of mobile broadcasting. A joint workshop took place in March 2011 with presentations from 3GPP and DVB standardization activities and, two months later, the creation of a study item was proposed to the 3GPP [10]. This proposal introduced the concept of common broadcasting specification (CBS) to be used in 3GPP mobile communications networks and DVB-based-broadcasting networks.

Although this proposal was not accepted by 3GPP due to lack of support from MNOs, the broadcast industry and academia continued working on 3GPP and DVB cooperation and several technical approaches have recently been proposed. The main assumption is the evolution of the broadcast capability of current 3GPP eMBMS in terms of capacity, coverage, and quality of service to match traditional broadcast networks.

First, the French M3 project evaluated and analyzed convergence possibilities of the CBS physical layer providing broadcast capabilities to the 3GPP LTE and DVB-NGH systems [11]. This technical proposal addressed key topics such as broadcast system architecture, modulation, system parameters, channel coding, time interleaving, and reference pilots. In particular, it proposes the use of time interleaving schemes at the physical layer based on enlarged transmission time interval and the time slicing concept to exploit time diversity. Moreover, it proposed some eMBMS enhancements in order to inte-

The reception on the dedicated broadcast carrier should be enabled by proper signaling located in LTE-A unicast carrier. Moreover, the tower overlay resources could be shared by multiple MNOs if all of them inform in their unicast carriers about the availability of dedicated carriers for broadcast services.

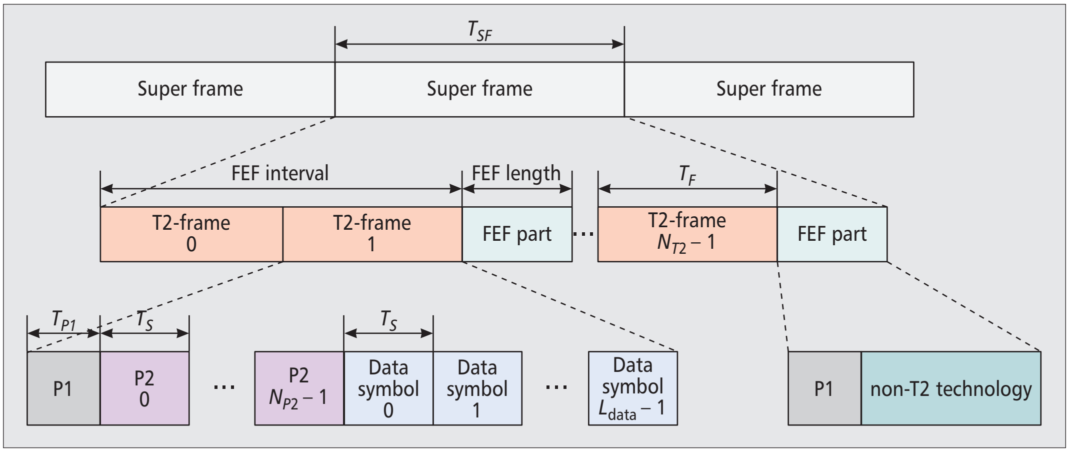


Figure 3. The DVB-T2 frame structure, showing the division into super-frames, T2-frames and FEF.

grate typical requirements from broadcasters such as the use of a longer CP and an optimized downlink-only system, which would also allow for an efficient use of HPHT networks to provide mobile video services.

Initially, CBS was assumed to be transmitted on a dedicated carrier either by MNOs using mobile cellular networks, or by broadcasters using HPHT networks. However, CBS could also be transmitted in-band with LTE/LTE-A unicast, as is the case in the current eMBMS specifications, or in-band with DVB-T2 broadcast, by using the future extension frames (FEF) concept shown in Fig. 3.

Precisely the FEF concept together with the LTE-A carrier aggregation concept are the basis of the other proposal for 3GPP and DVB cooperation, known as "tower overlay over LTE-A" [12]. On the one hand, the DVB-T2 framing structure defines two types of frames: T2 frames and FEFs. T2 frames contain preambles P1 and P2, which provide control information to DVB-T2 receivers, and T2 data symbols. FEF starts with P1 and the rest can be used for either extensions of the standard or other technologies, which allows time multiplexing various signal formats on the same frequency carrier. The use of FEF does not have an impact on existing DVB-T2 receivers since they work as if in a discontinuous transmission. On the other hand, carrier aggregation, one of the key features of LTE-A Release 10 [2], allows for the use of wider bandwidth and consequently increases transmission capacity by means of the aggregation of up to five carriers — one primary cell and up to four secondary cells — which can belong to either continuous or discontinuous spectrum.

Figure 4 shows the tower overlay over LTE-A system concept. Based on FEFs and carrier aggregation, it proposes the use of a hybrid carrier integrating an eMBMS dedicated carrier into a DVB-T2 data stream, which can be transmitted using existing HPHT networks for broadcast service delivery to both fixed and mobile devices.

The feasibility of this approach requires some extension or changes in both systems. Regarding DVB-T2 networks, they need to be extended to support the modulation of LTE Broadcast signals and their integration into FEFs. It implies the use of hybrid modulators. Concerning LTE/LTE-A networks, it is required to complete the definition of a MBSFN dedicated carrier within eMBMS standard and to support larger CPs. The integration of a dedicated MBSFN carrier into the existing LTE system can be achieved by carrier aggregation, which allows the simultaneous reception of both broadcast and unicast services by using a normal LTE-A unicast carrier as the primary cell and the MBSFN dedicated carrier as the secondary cell. The reception on the dedicated broadcast carrier should be enabled by proper signaling located in the LTE-A unicast carrier. Moreover, the tower overlay resources could be shared by multiple MNOs if all of them inform in their unicast carriers about the availability of dedicated carriers for broadcast services.

In order to comply with the timing regulations of both standards, special consideration is required on the length of the FEFs as well as T2 frames. Both LTE and DVB-T2 standards operate with different elemental periods defined as the inverse of their sampling frequency. To ensure a constant length of both frame types over a complete T2 superframe, the length of the two different parts in which the hybrid carrier is divided must be an integer multiple of 10 ms, that is, the length of a single LTE radio frame. This fulfills the requirements of the DVB-T2 standard regarding FEF and enables the synchronization between the LTE unicast carrier and the dedicated broadcast carrier. Thus, T2 frames, the initial P1 symbol of the following FEF, and a synchronization buffer whose length is an integer multiple of the DVB-T2 elemental period are contained in the DVB-T2 part, whereas the LTE dedicated broadcast part consists of an integer number of LTE radio frames.

LTE Broadcast as ATSC 3.0 Physical Layer

Currently, ATSC is working on the development of the ATSC 3.0 physical layer. Since most of the proposals take DVB-T2 and/or DVB-NGH as a baseline, it is likely that ATSC 3.0 will have significant similarities with these two DVB standards. However, one of the proposals from Qualcomm and Ericsson is based on LTE Broadcast

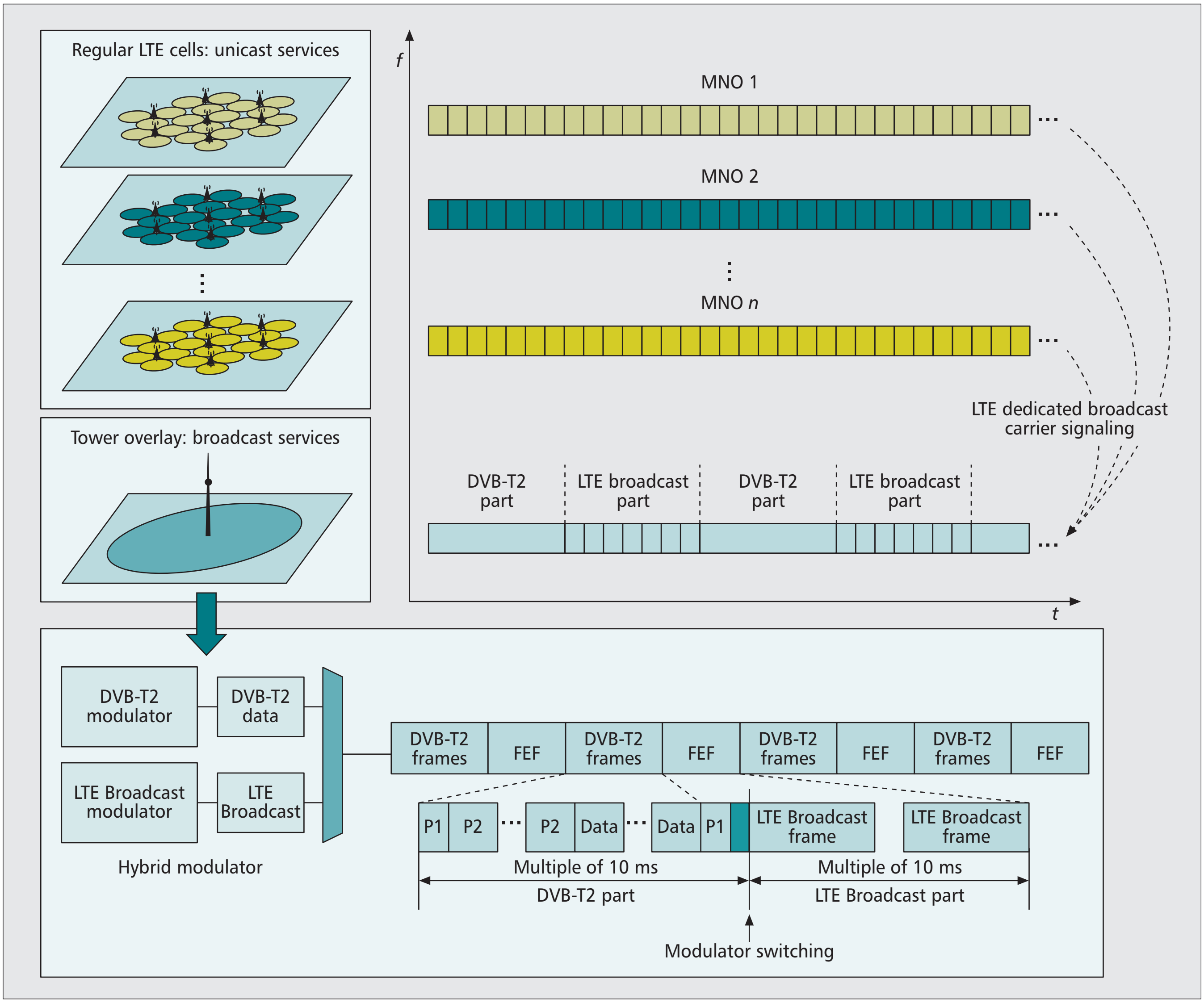


Figure 4. A Tower Overlay over LTE-A with the integration of an eMBMS dedicated carrier into DVB-T2 FEFs.

and proposes some enhancements of this standard in order to use DTT infrastructure to address both fixed and mobile use case applications [9].

A key motivation of this proposal is the announced intentions to deploy LTE Broadcast by multiple MNOs worldwide, where Qualcomm and Ericsson support these efforts with end-to-end integrated solutions. Moreover, LTE Broadcast provides an all-IP solution and, with some enhancements on the physical layer, it could be adapted for traditionally fixed TV broadcast services.

Qualcomm proposed a work item to 3GPP in order to improve the eMBMS efficiency from the radio access perspective [13]. In particular, it proposed the use of longer CPs for a large SFN delay spread environment, and the adoption of the MBSFN dedicated carrier option completing the upper layer support. The definition of a dedicated carrier would enable new MBSFN use cases, where the control information would be delivered on another unicast or mixed carrier. However, this work item proposal was not considered by the 3GPP.

Although the ATSC 3.0 physical layer is currently under development, it is likely that the proposal based on eMBMS goes in the same direction as the 3GPP work item proposal. That is, to complete the definition of eMBMS dedicated carriers and to define longer CPs to address deployment scenarios using HPHT broadcasting networks.

Challenges for a Future Mobile Converged Network

Historically, the business modeling for mobile broadcasting has not been attractive, considering the absence of a clear and viable economic model that resolves the monetary conflicts between cellular and broadcast operators, one of the main issues behind the failure of mobile TV services [14]. However, both industries are separately taking steps toward the definition of broadcast solutions to cope with future mass mobile multimedia services. The mobile industry has nowadays a good opportunity with eMBMS

By developing a single 5G standard to cope with both mobile a broadcast industry demands, an optimum usage of the spectrum based on spectrum sharing would be enabled. A coordinated spectrum access, as opposed to a competitive and mutually-interfering access, would also enable infrastructure sharing and simplified international frequency coordination.

Spectrum bands	Current user	WRC-15 target
Parts of 500–600 MHz [470–around 694 MHz]	TV broadcasting PMSE	Regional identification for IMT usage Need cooperation with broadcast industry
700 MHz [694–790 MHz]	TV broadcasting PMSE	IMT identification for Region 1 Proposed in WRC-12 and included in the WRC-15 AI 1.2
Parts of 1.4 GHz [1350–1525 MHz]	Digital radio Fixed service Scientific	Global identification for IMT usage Scientific use, only in a part of frequencies and some parts of regions
2700–2900 MHz	Radar	Global identification for IMT usage
3.4–3.6 GHz	IMT Satellite	Global identification for IMT usage
3.6–3.8 GHz	IMT Satellite	Global identification for IMT usage
Parts of 3.8–4.2 GHz	Satellite	Global identification for IMT usage
Parts of 4.4–4.99 GHz	Satellite	Global identification for IMT usage

Table 1. Possible candidate bands for IMT under WRC-15 Agenda Item 1.1.

and its future evolutions, and broadcasters are facing the definition of a worldwide DTT standard capable of supporting both rooftop and mobile reception.

In this framework, it is quite likely that these two industries, with their different network infrastructures, will compete for market and spectrum. However, the above mentioned trends for cooperation highlight the mutual benefit of a complementary use of both networks. For example, broadcasters need more infrastructure to provide mobile services to indoor users. The definition of the future 5G system brings a new opportunity for convergence, which would benefit both industries alike.

Spectrum Usage: From a Problem to a Solution

Future 5G mobile communications systems are required to make a flexible and efficient use of the available spectrum, whether contiguous or not, supporting wildly different network deployment scenarios [4]. In 2020 the envisaged spectrum requirements range between 1280 MHz and 1720 MHz [15]. However, the current spectrum allocated by the International Telecommunication Union (ITU) is much lower than these values in the three ITU Regions. The last decisions of the ITU regarding spectrum reallocations focus on reducing this deficit. In particular, ITU decided in the World Radiocommunications Conference 2007 (WRC-07) to allocate part of the UHF band traditionally used for analogue TV broadcasting to International Mobile Telecommunications (IMT) technologies. This band, also known as digital dividend, ranges from 790 MHz to 862 MHz in Region 1, and from 698 MHz to 790 MHz in Region 2 and Region 3, and requires additional guard bands to avoid interference between cellular and broadcast technologies. The worldwide allocation of the 700 MHz band to IMT technologies is on the agenda for the next WRC-15. Table 1 shows candidate bands for IMT under the WRC-15 agenda [15].

By developing a single 5G standard to cope with both mobile a broadcast industry demands, an optimum usage of the spectrum based on spectrum sharing would be enabled. A coordinated spectrum access, as opposed to a competitive and mutually-interfering access, would also enable infrastructure sharing and simplified international frequency coordination.

New Opportunities and Attractive Business Models

5G could smooth out current dissension between industry players and pave the way for a new converged industry, with new classes of services and new business models in which all parties, mobile network operators, content providers, broadcasters, spectrum holders, and advertisers, could benefit from this convergence solution and offer a full alternative to terrestrial TV broadcasting as a universal service.

Regarding the mobile industry, the higher transmission efficiency of broadcasting as compared with unicast delivery would reduce network overload while diminishing the CAPEX associated with the ultra-dense deployment of transmission points. Moreover, the possibility of sharing spectrum would reduce the investment in spectrum auctions, improve the quality of service, and enable new services and revenues. Concerning broadcasters, the definition of a worldwide DTT standard would boost economies of scale, enabling the integration of mobile TV into handheld devices. This will make current broadcasting services grow in terms of scope and availability. In addition, the future 5G networks could also be used to provide TV services to fixed receivers, which would entail the implementation of mobile cellular chips into TV sets.

Requirements for Successful Convergence

Although past generations of terrestrial broadcast and mobile systems are quite similar in their fundamentals — for example, they are OFDM-based — the optimal radio configuration comprising signaling, procedures, and transmission modes is different for broadcast services due to their particular features. Future 5G networks must address several requirements for a successful convergence of both industries.

First, the converged solution in 5G shall include a flexible and scalable broadcast mode able to allow the transmission of mass multimedia services to mobile and stationary receivers through different network infrastructures. This broadcast mode shall support the efficient transmission of mobile HDTV services. In addition, it shall support the use of traditional broadcasting HPHT infrastructure to increase the coverage range and should be able to share the same frequency used for transmitting fixed Ultra-HDTV services using time multiplexing. The future air interface shall be flexible enough to accommodate different sharing scenarios by using dynamic profiles that could switch from unicast to broadcast mode. In this sense, some radio functions could be activated/ deactivated/modified on demand depending on the specific needs of the service and the status of the network.

Conclusions

This article has reviewed the state-of-the-art and the current trends in mobile multimedia broadcasting, and promoted the convergence of cellular and broadcast networks in the future 5G mobile networks. Although the integration of both networks is on the roadmap of both industries, there are still some challenges that have been identified throughout this article. Potential technical enablers for such convergence are the future extension frames of second-generation DVB-T2 networks, which allows combining in the same frequency broadcast and cellular transmissions via time multiplexing, or the carrier aggregation feature of LTE-Advanced, which allows temporarily using a broadcast frequency for cellular transmissions. However, the main issue is that LTE broadcast eMBMS technology does not incorporate a sheer broadcast mode, and especially that it cannot be efficiently deployed in HPHT traditional broadcasting infrastructure due to their short guard intervals. Those considerations should be the starting point for the definition of a joint solution within the 5G framework.

Acknowledgment

The authors would like to thank the funding received from the Spanish Ministry of Science and Innovation within the Project number TEC2011-27723-C02-02.

References

[1] J. Zander and P. Mähönen, "Riding the Data Tsunami in the Cloud: Myths and Challenges in Future Wireless Access," *IEEE Commun. Mag.*, vol. 51, no. 3, March 2013, pp. 145–51.

[2] S. Parkvall, A. Furuskar, and E. Dahlman, "Evolution of LTE toward IMT-Advanced," *IEEE Commun. Mag.*, vol. 49, no. 2, Feb. 2011, pp. 84–91.

[3] J. F. Monserrat *et al.*, "Joint Delivery of Unicast and E-MBMS Services in LTE Networks," *IEEE Trans. Broadcast.*, vol. 58, no. 2, June 2012, pp. 157–67.

[4] A. Osseiran *et al.*, "The Foundation of the Mobile and Wireless Communications System for 2020 and Beyond: Challenges, Enablers and Technology Solutions," *Proc. IEEE 77th VTC Spring*, 2–5 June 2013, pp. 1–5.

[5] W. Zhang *et al.*, "FOBTV: Worldwide Efforts in Developing Next-Generation Broadcasting System," *IEEE Trans. Broadcast.*, vol. 60, no. 2, June 2014, pp. 154–59.

[6] GSA, "Evolution to LTE Report," Information Papers, Jan. 2015.

[7] D. Gómez-Barquero, *Next Generation Mobile Broadcasting*, CRC Press, March 2013.

[8] Jerry Whitaker, "ATSC Update: New Technologies for Today and Tomorrow," SBE CHAPTER 24, Inc., 2012.

[9] ATSC, "Summaries of Responses to ATSC 3.0 Physical Layer Call for Proposals," Oct. 2013.

[10] DVB Project, "A Cooperation between 3GPP and the DVB Project on a Common Downlink-Only Standard?" *Proc. 6th 3GPP PCG meeting*, Kansas City, March 2011.

[11] M. Crusire *et al.*, "A Unified Broadcast Layer for Horizon 2020 Delivery of Multimedia Services," *IEEE Trans. Broadcast.*, vol. 60, no. 2, June 2014, pp. 193–207.

[12] D. Rother, S. Ilsen, and F. Juretzek, "A Software Defined Radio Based Implementation of the 'Tower Overlay over LTE-A+' System," *Proc. IEEE Int'l Symposium Broadband Multimedia Systems and Broadcasting (BMSB) 2014*, Beijing, June 2014, pp. 1–6.

[13] G. K. Walker *et al.*, "Relationship Between LTE Broadcast/eMBMS and Next Generation Broadcast Television," *IEEE Trans. Broadcast.*, vol. 60, no. 2, June 2014, pp. 185–92.

[14] A. A. Razzac *et al.*, "Dimensioning and Profit Sharing in Hybrid LTE/DVB Systems to Offer Mobile TV Services," *IEEE Trans. Wireless Commun.*, vol. 21, no. 12, Dec. 2013, pp. 6314–27.

[15] Huawei, "Whitepaper on Spectrum," Feb. 2013.

Biographies

Jordi Calabuig (jorcagas@upv.es) received a telecommunications engineer degree and an M.Sc. degree in communications technologies, systems and networks from the Universitat Politècnica de València (UPV), Spain, in 2009 and 2010, respectively. Currently he is a Ph.D. student at the Institute of Telecommunications and Multimedia Applications, UPV. In 2012 he was a guest researcher with the Technische Universität Braunschweig, Germany. His research interests are focused on broadcast services in 4G mobile communications networks and its evolution toward 5G.

Jose F. Monserrat (jomondel@iteam.upv.es) received his M.Sc. degree with high honors and Ph.D. degree in telecommunications engineering from the Universitat Politècnica de València (UPV) in 2003 and 2007, respectively. He is currently an associate professor in the Communications Department of the UPV. He has been involved in several European Projects, being especially significant is his current participation in METIS, where he is leading the simulation work. He is co-author of the Wiley book *Mobile and Wireless Communications for IMT-Advanced and Beyond*.

David Gómez-Barquero (dagobar@iteam.upv.es) received a Ph.D. degree in telecommunications engineering from the Universitat Politècnica de València (UPV), Spain, in 2009. He is currently a senior researcher (Ramon & Cajal Fellow) at UPV's Institute of Telecommunications and Multimedia Applications (iTEAM), where he leads a research group working on the development of next-generation digital terrestrial broadcast technologies. He is the editor of the book *Next Generation Mobile Broadcasting* (CRC Press, 2013), and the vice-chairman of the Modulation and Coding Ad-Hoc Group of the standardization process of the next-generation U.S. TV broadcasting system ATSC 3.0.

LTE broadcast eMBMS technology does not incorporate a sheer broadcast mode and especially that it cannot be efficiently deployed in HPHT traditional broadcasting infrastructure due to their short guard intervals. Those considerations should be the starting point for the definition of a joint solution within the 5G framework.

Detection of False Data Injection Attacks in Smart-Grid Systems

Po-Yu Chen, Shusen Yang, Julie A. McCann, Jie Lin, Xinyu Yang

ABSTRACT

Smart grids are essentially electric grids that use information and communication technology to provide reliable, efficient electricity transmission and distribution. Security and trust are of paramount importance. Among various emerging security issues, FDI attacks are one of the most substantial ones, which can significantly increase the cost of the energy distribution process. However, most current research focuses on countermeasures to FDIs for traditional power grids rather smart grid infrastructures. We propose an efficient and real-time scheme to detect FDI attacks in smart grids by exploiting spatial-temporal correlations between grid components. Through realistic simulations based on the US smart grid, we demonstrate that the proposed scheme provides an accurate and reliable solution.

INTRODUCTION

Cyber-physical systems (CPSs) embedded in the environment are used to monitor, understand behaviors of, and control, the physical world [1]. As a representative emerging CPS application, the proliferation of smart grids has been observed in our daily life. A smart grid is a relatively new type of energy distribution system that consists of the power lines as with traditional power grids, as well as information and communication technology (ICT) infrastructures connected to smart meters that may take the form of specialized devices, or laptops, cell phones, and so on. Some of these smart grid components allow information systems to perform prediction analysis, which balances the power production and consumption in the grid system. For example, real-time pricing gives both consumers and suppliers valuable indications to help manage their energy demands and supplies, respectively. Therefore, energy distribution (which controls the energy generation, consumption, and transmission processes) can be performed in a more dynamic and efficient manner [2, 3].

However, the heterogeneity, diversity, and complexity of smart-grids pose critical challenges in ensuring overall system integrity [2]. This is because in smart grids, grid-status inference and decision making may be performed on local smart devices rather than in well-protected control centers. Therefore, unlike traditional power grids where the majority of attacks and failures come from physical accesses to critical facilities [4], the ubiquity of smart-grid components further invite these abnormalities from cyber-infrastructures.

A typical attack found in smart grids is false data injection (FDI), which can be utilized to distort real energy demand and supply figures. Energy distribution may therefore be erroneous, which results in additional costs or more devastating hazards. One example was the Northeast blackout of 2003 in the USA caused by a lack of accurate real-time condition information (http://en.wikipedia.org/wiki/Northeast_blackout_of_2003).

It is imperative that we trust such systems and get the security right as national security, not just cyber-security, is at risk. However, recent countermeasures against FDIs focus more on traditional power-grid scenarios [2], in which FDI attacks are launched at physical meters rather than smart components [4]. Without considering the cyber-attacks and the distributed design of smart grid infrastructures, such approaches may not be able to provide comprehensive protection that is immediate to each local smart meter or device making decisions based on status information.

In response to this problem, we propose a lightweight approach that can be easily installed on any smart component for real-time FDI detection. Compared to traditional solutions, it identifies anomalies in *state estimations* (which are inferred from physical meter readings), as these *state estimations* are the ultimate targets for both traditional FDI attacks and cyber-attacks. Furthermore, our approach scales to large smart grid systems compared to other similar anomaly-detection mechanisms in cyber systems [5]. Results of simulations based on real-world data demonstrate the effectiveness of our proposed approach in terms of reducing extra energy transmission costs and user outage rates caused by FDIs.

SMART GRID AND ENERGY DISTRIBUTION

In a smart grid, entities are connected to the grid via ICT infrastructures and power lines. Energy is transmitted from energy-rich to ener-

Po-Yu. Chen, Shusen Yang (corresponding author), and Julie A. McCann are with Imperial College London.

Ji Lin and Xinyu Yang are with Xi'an Jiaotong University.

 0163-6804/15/$25.00 © 2015 IEEE

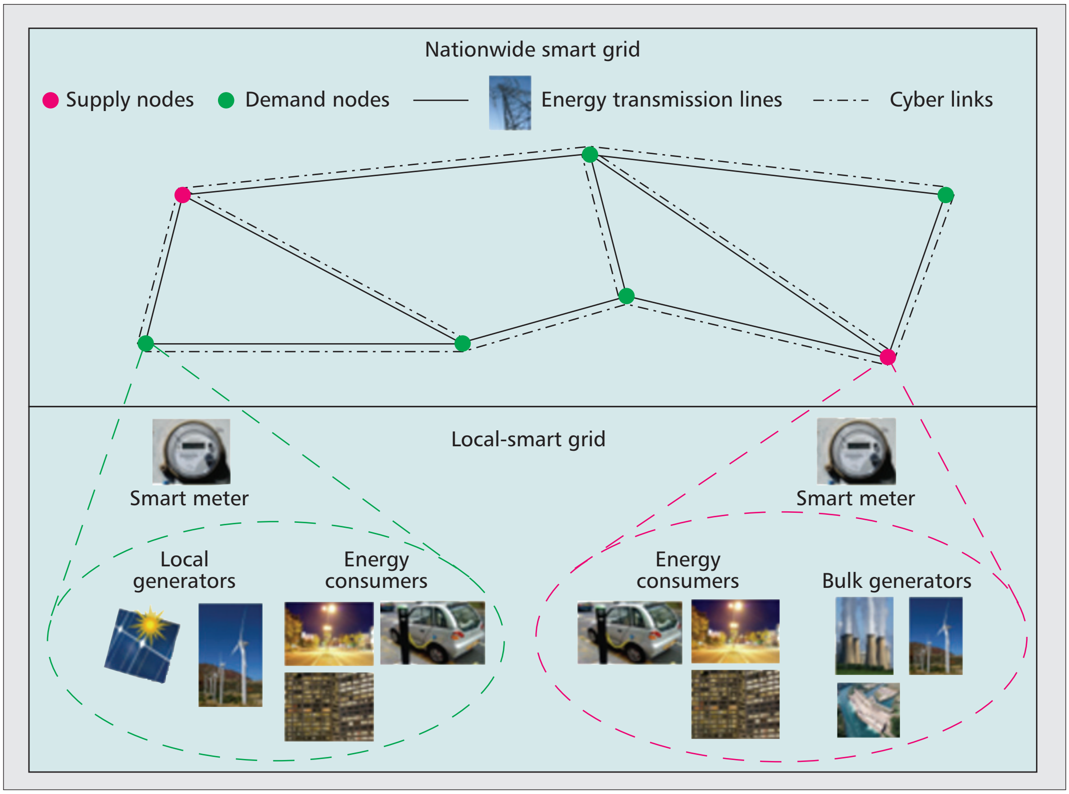


Figure 1. A conceptual illustration of cyber-physical topologies in a smart grid system.

In smart grids, these components are typically connections to the local networks or the Internet. Therefore, as the number of access points increase, compared with traditional power grids, smart grids are more prone to cyber-attacks.

gy-poor entities as per energy-distribution mechanisms. Figure 1 illustrates an example of a two-layer topology of a smart grid.

The upper layer represents the nationwide backbone power transmission grids and ICT infrastructures, where each *node* is an abstraction that represents the average energy demand/ supply of a local area (e.g. a town). Due to the difference in local energy needs and heterogeneous efficiencies of energy generation, each node in the grid will generate and consume different amounts of energy. When a node consumes more energy than it generates, the node is denoted as an *energy-demand node*, which needs to pull energy from the grid; similarly, when a node consumes less energy than it generates, the node is denoted as an *energy-supply node*, which pushes residual energy into the grid.

At the lower layer, smart meters are installed in every entity (e.g. house, factory, and school) that contains both power consuming facilities and energy producing equipment. Different from the traditional power meters that gather *raw measurements* (e.g. voltage and electrical current), smart meters are capable of further inferring *state estimations* (e.g. energy demands/ supplies) and make preliminary decisions (e.g. data fusion) before these estimations reach control centers. With the information obtained from smart meters, energy distribution can be optimized with regard to various power grid performance metrics. For example, they may maximize the network utility and energy usage efficiency, while minimizing energy transmission costs and user outage numbers [6].

FDI Attacks in Smart Grids

Although smart components such as smart meters play an important role in smart grids, they also increase the *vulnerability* of the grid system. In smart grids these components are typically connections to the local networks or the Internet. Therefore, as the number of access points increase, compared with traditional power grids, smart grids are more prone to cyber-attacks [2, 6]. Among these attacks, the most critical attack is false data injection (FDI) [2]. Adversaries can launch these attacks to compromise *raw measurements* or *state estimations* [4, 6].

These FDI attacks have been found to be launched by *consumers* and *adversaries* on purpose [2]. Consumers typically cheat the system by manipulating their energy consumption to reduce their energy bills. Adversaries, who can be opponent companies or countries, would aim to compromise *state estimation* (as described previously) to increase the cost of energy distribution and smart-grid operations. Considering the degree of these effects, the latter attacks are more critical for smart grids. False energy demand and supply can disrupt the energy distribution process, which may result in significant financial loss and even devastating outcomes.

FDI Attack Targets

The four most critical *state estimations* and the potential FDI attack targets related to these are listed as follows.

Energy Demand: The untruthful values of this state estimation are critical as they can dramatically increase financial costs to both the energy users and providers due to the extra cost

Detection is the most essential step in minimizing the damages resulting from the aforementioned FDI attacks. Therefore, the efficiency and effectiveness of FDI detection techniques have significant impacts on the overall performance of smart grid systems.

of power transmission and the waste of energy that can occur. Moreover, they can result in an even more devastating crisis, i.e. a power outage, where energy requests to the smart grid are less than the energy demand that nodes truly require. Since these nodes consist of energy consumers, such as personal users or companies, the adversary may launch such attacks either through infected personal devices such as laptops, or through poorly configured firewalls [2].

Energy Supply: The value of this state estimation is mainly provided by energy-supply nodes. FDIs can falsely decrease the advertised quantity of their supplied energy, which typically results in the starvation of energy-demand nodes who are not receiving what they request. Reversely, a falsely advertised increase in energy supply can incur an increase of wasted energy. These attacks can be achieved using malware to infect the servers of power suppliers, falsifying the quantity of energy it can truly provide.

Grid-Network States: The grid-network states are the configurations and conditions of the power grids, such as grid topologies and power-line capacities. For instance, adversaries can isolate nodes from the power grid by invalidating their power-line connection. Maliciously forging the network states can seriously mislead the energy distribution, again resulting in devastating power shortages or extra energy transmission costs.

Electricity Pricing: Dynamic electricity pricing can greatly reduce the electricity bill of consumers, such as energy-utility companies and houses, as well as help balance the power loads between peak and off-peak periods. Therefore, fake electricity pricing would incur significant damage to both the financial and physical subsystems that make up a smart grid, negating their advantages of optimum supply efficiencies. For instance, adversaries can falsely reduce prices during peak hours, which will ultimately result in grid system overload. Also, illegal users may use malware to modify their energy bills, leading to a loss of utility company revenue.

Detecting FDI Attacks

Detection is the most essential step in minimizing the damages resulting from the aforementioned FDI attacks. Therefore, the efficiency and effectiveness of FDI detection techniques have a significant impact on the overall performance of smart grid systems.

Traditional Solutions for FDI Detection

Since the smart grid is a relatively new topic compared to traditional power-grid systems, the main research on countermeasures against FDIs are mainly system-theoretic approaches [4, 7, 8]. In these approaches, the relationship between raw meter measurements and system states are simply described as $\vec{z} = \mathbf{H}\vec{x} + \vec{e}$, where the matrix $\mathbf{H}$ represents system configuration, and vectors $\vec{z}$, $\vec{x}$, and $\vec{e}$ represent the meter measurements, state estimations, and meter measurement errors, respectively. Therefore, for a predicted state $\hat{\mathbf{x}}$, abnormal values can be detected by computing the 2-Norm of its measurement residual $||\mathbf{z} - \mathbf{H}\hat{\mathbf{x}}||$. $\mathbf{z}$ will be considered to be abnormal, if $||\mathbf{z} - \mathbf{H}\hat{\mathbf{x}}||$ is larger than a given threshold τ.

Recent research [4] demonstrates that this traditional countermeasure approach can be easily bypassed if adversaries are capable of accessing enough meters and inject malicious data to the grid with specific values. Also, this traditional solution can only detect abnormalities at the group level (each of which contains several meters). Abnormalities in each individual meter are unobservable in this solution. Although some approaches try to solve this problem, such as [7], they are unable to guarantee the security of smart-grid systems, since they are limited to traditional power grid scenarios, where adversaries can only assess physical meters. Therefore, attacks through cyber infrastructures can directly compromise *state estimations* without being detected by the detection approaches in traditional power-grid systems [2].

Anomaly Detection for FDI Attacks

Since FDI attacks would result in abnormal state estimations, we can apply anomaly detection techniques [5] to smart grids, which has been studied in various areas such as cyber intrusion, sensor networks, and image processing [5]. In current research on intrusion detection, rule-based mechanisms are the simplest solutions. They typically exploit static thresholds to identify anomalies. When the value of a raw data input exceeds these thresholds, this value is regarded as anomalous. These schemes introduce very low overhead to the system, but they fail to adapt to valid changes in the environment. Furthermore, a high level of professionalism is required to manually maintain detection thresholds.

To cope with the dynamics and heterogeneities in environments, machine-learning-based solutions are more often exploited [5]. Statistical models (e.g. Bayesian, k-nearest neighbor) and artificial intelligence (e.g. neural networks) are typical approaches utilized to distinguish anomalies from norms. However, these solutions are usually complex, which limits their scalability when applied to large complicated networks, such as smart grids. Therefore, to achieve a more balanced solution, further assumptions are usually made to reduce computational complexity. What we exploit here the most is one simple general assumption: *spatiotemporal correlation*.

Spatio and Temporal Correlations in Smart Grids

Spatiotemporal correlation is a natural property found in various physical phenomena including human behaviors, since they are typically continuous over both the time and spatial domains. Since cyber states are mappings to these phenomena, the estimations of these states within a correlated sphere should also be spatiotemporally correlated. In general, the following correlations exist in smart grid systems.

Traditional power suppliers usually generate energy based on the current needs of power consumers: Typically, to reduce the extra overhead caused by energy storage, traditional power suppliers (such as fossil-fuel power stations) are inclined to dynamically adapt their power generation to the requirements of consumers. Therefore, these power suppliers, in the same area, can be spatially temporally correlated as they should all reflect the current state of power consumption nearby.

Renewable-energy suppliers are typically correlated to nearby suppliers: As renewable energy comes from natural resources (such as solar and wind power), energy generators, for example, wind farms and hydroelectricity plants, deployed in a nearby area should be able to simultaneously reveal the current state of the environment therein. That is, they are spatially correlated. Also, these spatial correlations should be similar to those that have occurred in the past, that is, these energy supplies are temporally correlated, as natural resources are continuous and should therefore result in interrelated effects to these power suppliers.

Energy consumers within the same area should behave in similar patterns: In general, power consumption in a related area should simultaneously reflect the current state of this region, such as weather, activities, and government policies. For example, in winter, power consumers of nearby cities tend to expend more power supplies than in summer due the use of heaters, even when each individual has its own patterns regarding power consumption. Therefore, the total node behavior can still be consistent and correlated to other consumer nodes at higher abstractions (such as districts of a city).

Real-Time Correlation-Based FDI Detection

In order to minimize the damage of FDIs in smart grid systems, we propose an effective real-time mechanism to detect FDI attacks against state estimation in smart grid systems, based on the spatiotemporal cyber-state correlations discussed previously. We call strongly spatiotemporally correlated smart components as neighbors. Potential anomalies can be detected by monitoring the temporal consistencies of the spatial correlations between state estimations. This detection mechanism can be divided into the following three phases, discussed below.

Spatial-Pattern Recognition and Temporal-Pattern-Consistencies Evaluation

This phase models the current states of the smart grid as its normal-state references (i.e. the state without FDI attacks). It is a semi-supervised process, similar to many other machine-learning mechanisms [5]. In our approach, correlation patterns are recognized between each pair of state estimations within the same correlation sphere defined by smart-grid operators, rather than the whole network. A correlation sphere typically contains several smart meters, while a smart meter may simultaneously belong to multiple correlation spheres. The defining of correlation spheres can guarantee that all the smart components are spatially and temporally correlated, which is required by our approach. This design also avoids the high computational complexity when analyzing data over high-dimensional spaces [5].

We assume that genuine changes and errors in state estimations follow Gaussian distributions, which is a common assumption [5]. Let $s_i(t)$ denote the state estimation of a smart component i (e.g. a photoelectric energy harvester) at current time t. The sequence of previous T state estimations for a smart component i before current time t can be represented as

$$S_i(t, T) = (s_i(t - T), s_i(t - T + 1), \cdots, s_i(t - 1)) \quad (1)$$

We consider each pair of smart components i and j in a correlation sphere G, such as a set of correlated photoelectric-energy harvesters in a *town*. Based on their previous state estimations $S_i(t, T)$ and $S_j(t, T)$, we can compute their spatial-correlation consistency region (SCCR), which represents the set of all possible potentially correct estimation pairs of $s_i(t)$ and $s_j(t)$ at current time t. If the current estimation pair $(s_i(t), s_j(t))$ belongs to the set defined by the SCCR, we say the current estimations $s_i(t)$ and $s_j(t)$ are *consistent*; otherwise, they are *inconsistent*.

Geometrically, SCCR can be approximated by a rotated ellipse, as illustrated in Fig. 2, Here, the center of the SCCR ellipse, $(\overline{s}_i(t), \overline{s}_j(t))$, is computed by using the exponential weighted moving average (EWMA) for the previous estimations $S_i(t, T)$ and $S_j(t, T)$. The major and minor axes of the SCCR ellipse are computed by using singular value decomposition (SVD), based on the previous estimations $S_i(t, T)$ and $S_j(t, T)$. In our approach, SVD is a mathematical procedure to convert observations of multiple observers into two orthogonal principal components $\vec{a}$ and $\vec{b}$ (which define the rotation angle θ) and their associated variances σ_a^2 and σ_b^2. We set the lengths of these axes as the three deviations $3\sigma_a$ and $3\sigma_b$, which cover 99.46 percent normal observations, respectively.

Trust-Based Voting

After the correlation mapping consistency between each pair of state estimations is obtained, trust-based voting is exploited to identify the *occurrences* of anomalies. This voting is divided into two different rounds. In the first round, reliable state estimations are selected with the votes from the state estimations of their correlation neighbors. For a smart component i in a correlation sphere G, define its correlation neighbor set $N_i \subset G$ as the set of all devices in G excluding i; and define its consistent neighbor set $N_i^c \subseteq N_i$ as a set of devices, where the state estimation $s_j(t)$ of each device j in N_i^c are *consistent* with the state estimation $s_i(t)$. Take Fig. 3 for instance: $N_A = \{B, C, D, E\}$, $N_E = \{A, B, C, D\}$, $N_A^c = \{B, C, D, E\}$, and $N_E^c = \{A\}$.

Let $|N_i|$ and $|N_i^c|$ represent the sizes of sets N_i and N_i^c, respectively. We say that component i and its estimation $s_i(t)$ are *likely* to be *reliable* (LR) at time t, if $|N_i^c|/|N_i| \geq$ 50 percent. Otherwise, we say that the smart component i and its estimation $s_i(t)$ are *likely* to be *anomalous* (LA) at time t. After the first round voting, smart component E and its current state estimation are LA while others are LR, as shown in Fig. 3.

Spatiotemporal correlation is a natural property found in various physical phenomena including human behaviors, since they are typically continuous over both the time and spatial domains. Since cyber states are mappings to these phenomena, the estimations of these states within a correlated sphere should also be spatiotemporally correlated.

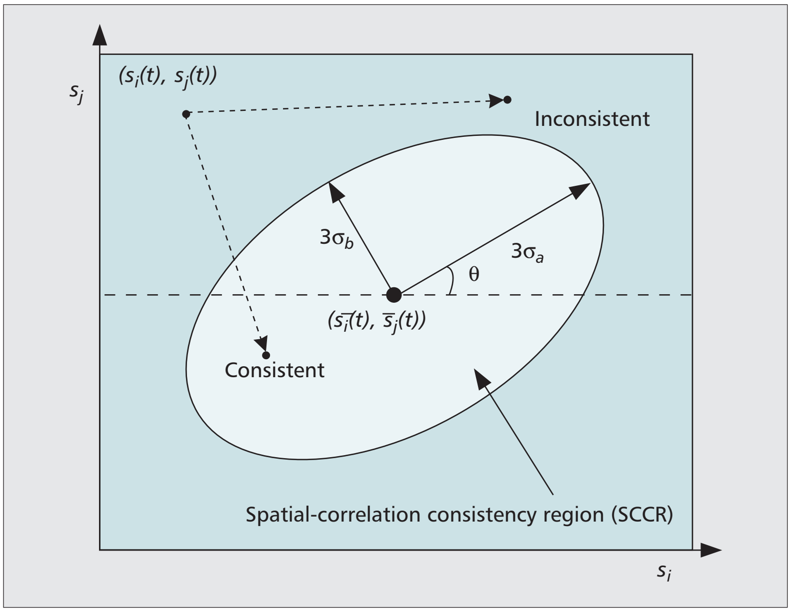


Figure 2. Illustration of a spatial correlation consistency region.

After the first round voting, only LR components and their state estimations are qualified to be involved in the second round voting, such as A, B, C, D in Fig. 3. For a smart component i, define $N_i^{LR} \subseteq N_i$ as the set of all LR components in the correlation sphere G excluding i; and $N_i^R \subseteq N_i^{LR}$ as the set of all LR components that are *consistent* with i. For instance, $N_E^{LR} = \{A, B, C, D\}$ and $N_E^R = \{\text{A}\}$, shown in Fig. 3. The second round voting for state estimation $s_i(t)$ has one of the following three outcomes:

- **Good**, if $|N_i^R|/|N_i^{LR}| \geq 50$ percent and $|N_i^{LR}|/|N_i| \geq 50$ percent.
- **Abnormal**, if $|N_i^R|/|N_i^{LR}| < 50$ percent and $|N_i^{LR}|/|N_i| \geq 50$ percent.
- **Unknown**, otherwise.

It can be seen that we can determine whether $s_i(t)$ is **Good** or **Abnormal**, if the majority of i's neighbors are labelled as LR at slot t; otherwise, the reliability of $s_i(t)$ is unable to be determined (i.e. **Unknown**), due to the lack of reliable references.

System Condition Inference

We further infer the phenomena that the three outcomes represent. First, **Good** indicates that current state estimation $s_i(t)$ is reliable, since it is highly correlated to other state estimations that are in the same correlation sphere G. This usually suggests that the state estimation is in a stable state and probably normal. On the contrary, **Abnormal** indicates the value of $s_i(t)$ deviates from the majority of state estimations in the same correlation group G. This typically suggests an abnormality in $s_i(t)$, which can either be due to failures or an FDI attack. Finally, the **Unknown** occurs when there are not enough state estimations that can be regarded as reliable references when we compute the reliabilities. This happens when there is an extensive change among the state estimations in G due to genuine occasions that dramatically change the behavior of the smart grid or large-scale attacks, which should be regarded as an indispensable event that requires further inspection from system administrators.

Practical Issues

Except for assuming spatiotemporal relationships and that errors follow Gaussian distributions, we do not require any further assumptions. Therefore, the proposed FDI detection mechanism is quite general and can be used for heterogeneous types of correlated components in the smart grid systems. When we apply the proposed detection mechanism to practical smart grid systems, the window size of previous state estimations, T, should be considered. This is a user defined parameter, which decides the computation of the SCCR (i.e. the ellipse shown in Fig. 2). Consequently, users should carefully assign T according to their practical scenarios to assure that the entire state estimations that lie in this interval are temporally correlated, while avoiding unnecessary extra storage and computational overheads.

A Case Study

To demonstrate how FDI attacks would impact the energy distribution of a smart grid system, we conducted an evaluation based on a simplified version of the US smart grid consisting of 48 states, as shown in Fig. 4 (http://www.oe.energy.gov/smartgrid.htm.). Furthermore, to simulate local energy generation and consumption of each state, we decomposed each state into 10 energy suppliers and 10 energy consumers. Each of these suppliers/consumers contains 365 daily energy generation/consumption profiles, which is correlated to other suppliers/consumers in the same state, to simulate the annual behavior of the smart-grid system. All these profiles are based on the 2009 US Energy Information Administration State Electricity Profiles (available on http://www.eia.gov/).

In this simulation, we study two different types of FDI attacks. The first FDI attack aims to increase the *total energy transmission cost* [4]. In this scenario, FDI attacks were randomly inserted to increase the advertised supply-demand values of suppliers and consumers by 15 percent and then 30 percent , that is, 0.15× false and 0.3× false shown in Fig. 5a. The total energy transmission cost $Cost_{total}$ is defined as

$$Cost_{total} = \Sigma_{L_{ij} \in \boldsymbol{L}} E_{ij} \cdot Cost_{ij}, \quad (2)$$

where $\boldsymbol{L}$ denotes the set of all power lines in a given smart-grid system; L_{ij} denotes the power line between supplier i and consumer j; E_{ij} denotes the amount of energy transmitted through power line L_{ij}; and $Cost_{ij}$ denotes the cost of power transmission per unit through power line L_{ij}.

The second FDI attack aims to promote significant power-supply outage. The demands of energy consumers were falsely increased by 1 and 2 times, that is, 1× false and 2× false shown in Fig. 5b, while the advertised supplies were falsely decreased by 10 percent and 20 percent. The user outage rate R_{out} is defined as:

$$R_{out} = \frac{\text{total number of outage nodes}}{\text{total number of demanding nodes}}$$

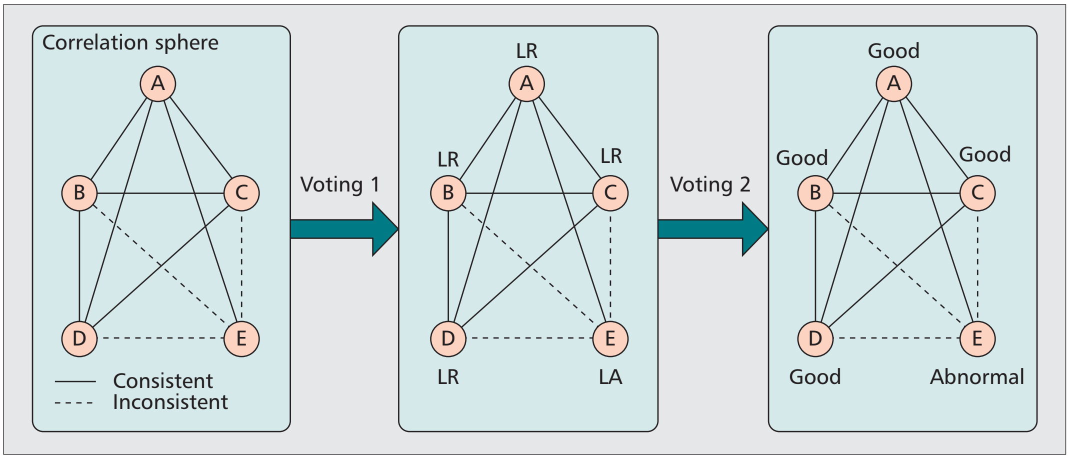


Figure 3. Visualization of a trust-based voting example.

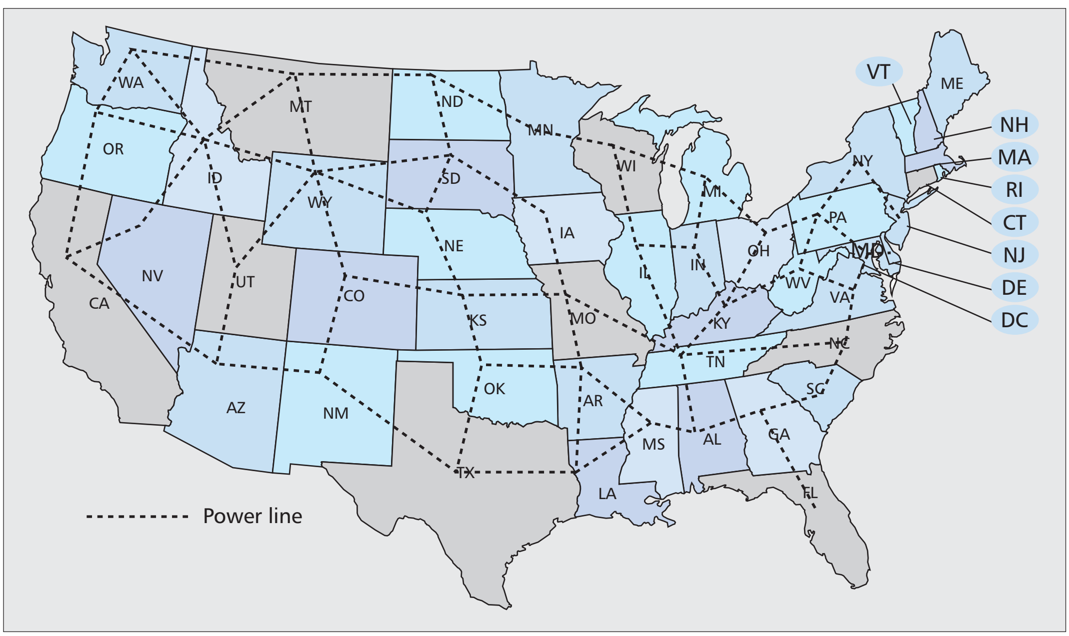


Figure 4. A simplified US smart grid topology.

Except for assuming spatiotemporal relationships and that errors follow Gaussian distributions, we do not require any further assumptions. Therefore, the proposed FDI detection mechanism is quite general and can be used for heterogeneous types of correlated components in the smart grid systems.

Figure 5 shows the simulation results when applying our FDI-detection approach and the state-of-the-art distanced-based (DB) solution proposed in 2013 [9, 10]. Both approaches referenced the past 30 days data during evaluation (i.e. $T = 30$). Also, we set the user-defined parameters of the distanced-based solution as $r = 0.5$ and $D = 0.5$ [9]. Note that since there were no false-positive detections caused by our approach in the cost-increasing scenario, it cannot be plotted under log scale in Fig. 5c.

As can be seen in Figs. 5a and 5b, both FDI-detection mechanisms can significantly reduce the adverse impacts caused by FDIs. In the cost-increasing experiment, our solution reduced up to 1.14 million US dollars without any false-positive detection. Furthermore, our approach can achieve about 55.5 percent user outage reduction with only 0.43 percent false-positives.

In contrast, although the DB-detection scheme achieves comparable performance in detection FDIs, it suffers from extremely high false-positive detection rates, which means that it misreports correct state estimations when FDI attacks with a high probability. As can be observed in Fig. 5c, false-positive detections are increasing as the FDI-attack rate increases for both schemes; however, compared with our approach, DB detection incurs nearly two-orders of magnitude more false-positive results. Such a high false-positive detection rate of the DB-detection scheme would result in significant confusion in decision making and would incur additional costs, such as extra labour to understand the problem where the integrity of the information is required to be checked manually, that is, interpreted by human beings.

Furthermore, as shown in Fig. 5d, the computational overhead of our approach is extremely low, that is, nearly two-orders of magnitude less than that of DB-detection. This demonstrates that our approach has a great potential to be applied for real-time FDI attack detection in smart grid systems.

Conclusion

Compared to traditional power grids, smart grids are predicted to be more reliable and effective energy-distribution solutions that can cope with complicated power supply and demand scenarios. However, they have the potential to be more vulnerable to cyber-attacks such as from malware or malicious cyber-intrusions. Among all these attacks, the most critical ones are false-data injection (FDI). Adversaries can launch these attacks to compromise *raw meter measurements* or *state estimations*. This can significantly disrupt the effectiveness of the energy distribution in smart grids, which results in additional costs or some hazards that would have larger impacts, such as blackouts or enormous costs.

Recent approaches to identifying FDIs are less

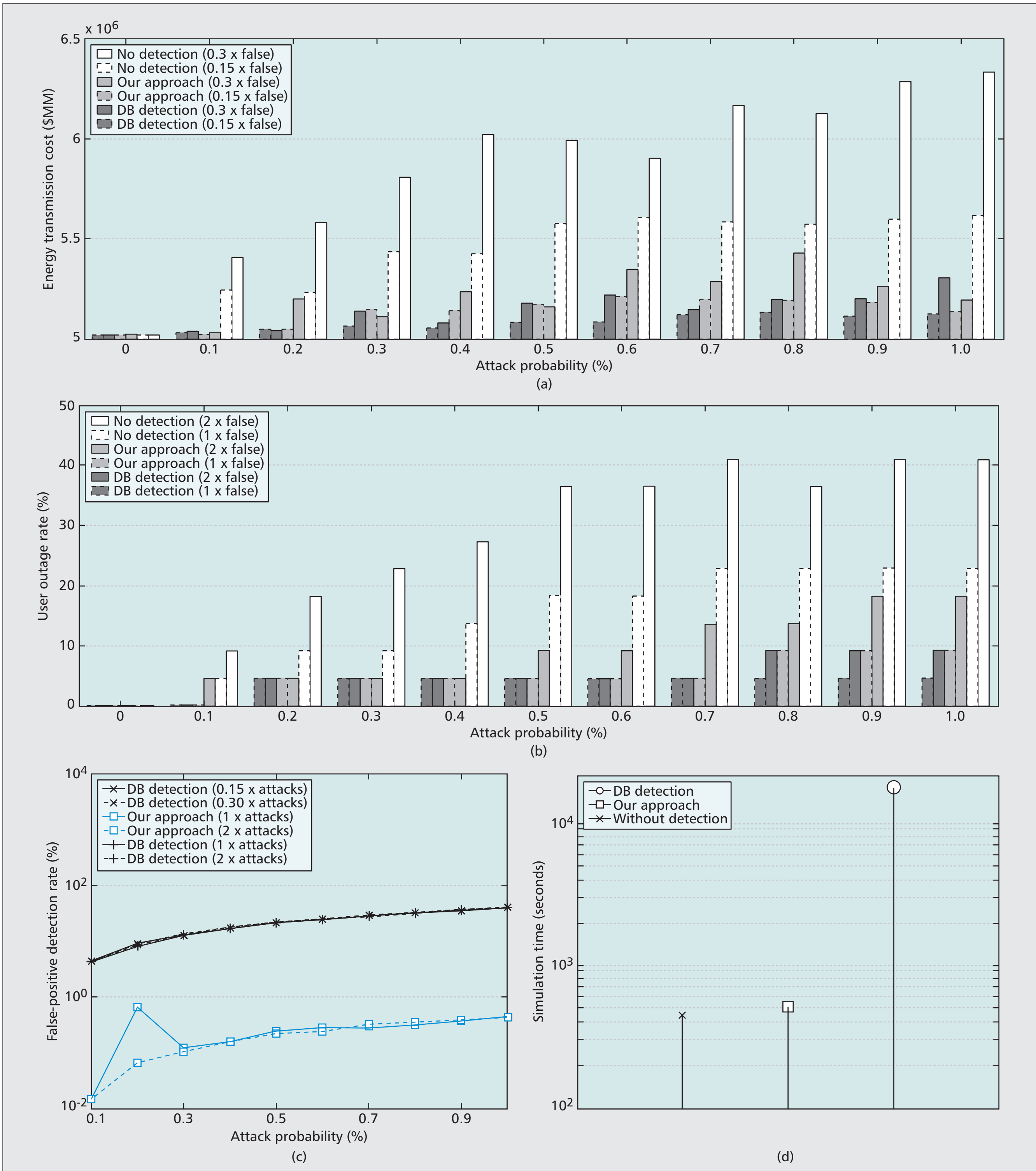


Figure 5. The correlation between (a) energy transmission cost (b) user outage rate (c) false-positive detection rate, and attack probability for different attack criteria; and (d) required simulation time.

effective, as they assume the technical architecture of traditional power grids. In this article we propose a simple and robust solution to detecting false data injection in smart grid systems that consist of many differing ICT components. This solution exploits spatiotemporal cyber-state correlations and trust-based voting to evaluate the reliabilities of *state estimations*. FDIs can be detected once those unreliable state estimations are identified. A case study is presented to demonstrate the impacts of FDIs with our approach versus the nearest state of the art approach. The simulation results show that the proposed solution can effectively and efficiently detect malicious FDIs and prevent potential extra costs and security threats.

Even with our solution, smart-grid systems remain vulnerable to some FDIs such as those that slowly evolve to prevent detection, as well as other sorts of attacks both physical and computational. Therefore, more powerful countermeasures are required as these attacks can become potential threats to national security and citizen wellbeing.

Even with our solution, smart-grid systems remain vulnerable to some FDIs such as those that slowly evolve to prevent detection, as well as other sorts of attacks both physical and computational. Therefore, more powerful countermeasures are required as these attacks can become potential threats to national security and citizen wellbeing.

Acknowledgments

This work is sponsored by the Intel Collaborative Research Institute (ICRI) for Sustainable Connected Cites, and the Natural Science Foundation of China (NSFC), 61373115 and 61402356.

References

[1] E. A. Lee, "Cyber-Physical Systems—Are Computing Foundations Adequate," position paper for NSF Workshop on Cyber-Physical Systems: Research Motivation, Techniques and Roadmap, vol. 2, 2006.
[2] M. Yilin, *et al.*, "Cyber-Physical Security of a Smart Grid Infrastructure," *Proc. IEEE*, vol. 100, 2012, pp. 195–209.
[3] Y. Liu, P. Ning, and M. K. Reiter, "False Data Injection Attacks Against State Estimation in Electric Power Grids," *ACM Trans. Inf. Syst. Secur.*, vol. 14, 2011, pp. 1–33.
[4] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly Detection: A Survey," *ACM Comput. Surv.*, vol. 41, 2009, pp. 1–58.
[5] M. Baghaie, S. Moeller, and B. Krishnamachari, "Energy Routing on the Future Grid: A Stochastic Network Optimization Approach," *Proc. IEEE POWERCON*, 2010, pp. 1–8.
[6] L. Jie *et al.*, "On False Data Injection Attacks Against Distributed Energy Routing in Smart Grid," *Proc. IEEE/ACM ICCPS*, 2012, pp. 183–92.
[7] R. B. Bobba *et al.*, "Detecting False Data Injection Attacks on DC State Estimation," preprints of the First Workshop on Secure Control Systems, CPSWEEK, 2010.
[8] O. Kosut *et al.*, "Malicious Data Attacks on the Smart Grid," *IEEE Trans. Smart Grid*, vol. 2, 2011, pp. 645–58.
[9] Y. Lei and L. Fengjun, "Detecting False Data Injection in Smart Grid In-Network Aggregation," *Proc. IEEE SmartGridComm*, 2013, pp. 408–13.
[10] S. Subramaniam *et al.*, "Online Outlier Detection in Sensor Data Using Non-Parametric Models," *Proc. VLDB*, 2006, pp. 187–98.

Biographies

Po-Yu Chen received a B.S. degree from National Cheng Kung University, Tainan, Taiwan, and an M.S. degree from National Taiwan University, Taipei, Taiwan. He is currently working toward a Ph.D. degree in the Department of Computing, Imperial College London, London, U.K. His research interest is sensor-data mining, including data-correlation discovery and anomaly detection, in wireless networks, networked sensing systems, and distributed systems.

Shusen Yang received a Ph.D. degree from Imperial College London, London, U.K., in 2013. He is currently a research associate with Imperial College London, where he is also with the Intel Collaborative Research Institute for Sustainable Connected Cities. His research interests are wireless networks, networked sensing systems, and distributed systems, including stochastic network optimization, energy harvesting networks, mobile crowd sensing, in-network processing, and network reliability. He is a member of the Association for Computing Machinery.

Julie A. McCann is a professor of computer systems at Imperial College London, where she leads the Adaptive Embedded Systems Engineering Research Group and the Intel Collaborative Research Institute for Sustainable Cities, and works with NEC and others on substantive smart city projects. Her research centers on highly decentralized and self-organizing scalable embedded systems. She is a Fellow of the British Computer Society.

Jie Lin received his Ph.D. degree from the Department of Computer Science and Technology at Xi'an Jiaotong University in 2013. He is currently a lecturer in the Department of Computer Science and Technology at the Xi'an Jiaotong University. His main research interests include wireless sensor networks security and cyber space security in cyber-physical systems.

Xinyu Yang received a Ph.D. degree in computer science from Xi'an Jiaotong University in 2001. He is currently a professor in the Department of Computer Science and Technology at the Xi'an Jiaotong University. His main research interests are in the areas of wireless communication, mobile ad hoc networks, network security, and cyber-physical systems security.

SMART for Mobile Health: A Study of Scheduling Algorithms in Full-IP Mobile Networks

Anpeng Huang and Linzhen Xie

The authors are with Peking University.

This study is supported in part by the Beijing Nature Science Foundation (No. 7152087), in part by the National Nature Science Foundation of China (NSFC) (No. 81471184), in part by the National Science and Technology Major Projects in Mobile Healthcare (No. 2012ZX03005013, No. 2013ZX03005008-003), in part by the National Key Technologies R&D Program (No. 2013BAI05B07), in part by the National Basic Research Program of China (973 Program) (No. 2010CB328201, No. 2013CB329205), in part by the National High-tech R&D Program of China (863 Program) (No. 2014AA01A704), in part by the National 985 Program Fund from Peking University, and in part by the Seeding Grant for Medicine and Information Sciences of Peking University (No. 2014-MI-02).

ABSTRACT

As a joint initiative of WHO and ITU, mHealth is becoming a feasible tool to offer data-intensive healthcare services online. In general, each kind of mHealth application has user-specific service-quality requirements that could be satisfied in the cost of bandwidth overconsumption. In any mobile network, bandwidth overconsumption is selfishly competing with limited radio resources. This phenomenon is further magnified by the "best-effort" service pattern inherited from IP networking, and the influence from interference, fading, and other mobility effects on unreliable radio channels. To enrich mHealth applications, we propose a Service-Mapping Adaptive Radio Transmission (SMART) scheduling algorithm, in which system parameters (QCI, CQI, MCS in AMC, etc.) are employed to enable "better than best effort" packet-switching services in full-IP mobile networks (e.g. LTE/LTE-Advanced). In this proposal, the designed scheduling metric is a self-adaptive mechanism that can combine mobile service-quality inspection and transmission pattern recognition dynamically. System-level simulation experiments demonstrate that the proposal can obtain approximately 31.25 percent gain on average, in contrast to existing solutions. Additionally, the runtime complexity of this solution is lowered to $O(K \log K)$, where K is the number of mHealth users within a cell.

MOBILE HEALTH: A PROMISING TOOL FOR ONLINE DATA-INTENSIVE HEALTHCARE

As emphasized in the World Health Organization (WHO) report [1], the cost of healthcare is becoming the unaffordable socio-economic problem in the world. Let us take noncommunicable diseases (NCDs) (chronic diseases) as an example. According to the definition by WHO, NCDs are of long-duration with slow-progression symptoms [1]. These include cardiovascular diseases, cancer, chronic respiratory diseases, diabetes, strokes, and so on. It is reported that NCDs are the leading cause of death, accounting for 63 percent of the annual deaths in the world. In China, there are 260 million chronic disease patients, according to Chinese government statistics [2]. In other words, one in every five Chinese person is affected with at least one kind of chronic disease. Most importantly, major organs, e.g. brain, heart, and kidneys, are subject to malfunction and even impairments from NCDs. Moreover, NCDs are seen as the major cause of physical disabilities, which affect one in every three working-age adults (aged 18-65 years) to have work limitations in the United States. After more than 20 years of steady increases, healthcare expenses now represent 17.6 percent of GDP, nearly US$ 600 billion in the US [3]. Only considering the main three types of NCDs in China, heart disease, stroke, and diabetes, the national GDP lost US$ 550 billion during the past 10 years [2]. In addition, NCDs also lead to many kinds of mental illnesses, for example, stress, anxiety, and depression.

To solve this huge socio-economic problem, a trend of data-intensive applications is under way in healthcare (a.k.a. big-data healthcare), including data from clinical trials, inpatients, and outpatients. To help NCD-affected people, their health data must be collected in a timely manner and delivered to a professional healthcare center online, without unexpected disruption and distortion. Furthermore, this health data can help doctors to make right clinical decisions quickly, and their diagnosis/treatment knowledge and information can be shared and used to develop healthcare self-learning systems for the public. Driven by these multiple incentives, online data-intensive healthcare services are becoming more and more necessary. To convey this health data to their users, wireless mobile communications is the natural choice, as it can allow users to access health data services without time and location limitations [4, 5]. This rising wave of healthcare services over mobile communications is known as mobile health (mHealth), which is expected to change healthcare delivery today and is at the core of responsive health systems [6–9]. To ride

 0163-6804/15/$25.00

this wave, WHO and ITU have founded a joint mHealth working group for improving public healthcare in the world [9], and they have a common expectation that mHealth will become a useful tool of online big-data health services, with the proliferation of the global medical wearable electronics (with a market of more than $2.8 billion in revenue in 2014, expected to pass $8.3 billion in 2019, and growing at a healthy compound annual rate of 17.7 percent from 2014 to 2019) [10]. When healthcare services are carried in a mobile network, their data transmission quality suffers from interference, fading, path loss, shadow, and other negative effects. Since wireless bandwidths are always limited, an mHealth system operation may be confronted by an issue from data-intensive transmission or system integration [4]. On the other hand, when mobile networks are evolving into full-Internet Protocol (IP) packet-switching (e.g. Long-Term Evolution (LTE)/LTE-Advanced), the existing "best effort" service pattern born from IP networking is unsuitable for service-sensitive applications such as mHealth [11, 12].

To provide data-intensive healthcare services in full-IP mobile networks, a scheduling strategy plays a special role [13, 14]. In this study, we focus on this topic,[1] and study how to conceive a scheduling strategy that can intelligently offer "better than best effort" full-IP services for mHealth applications.

QCI index	Resource type	Priority level	Packet delay	Packet loss	Typical examples from mHealth applications
1*	GBR	2	100 ms	10^{-2}	Emergency VoIP call
2		4	150 ms	10^{-3}	Consultation video call
3		3	50 ms	10^{-3}	Patient tracking in remote video
4		5	300 ms	10^{-6}	Daily health monitoring
5*	Non-GBR	1	100 ms	10^{-6}	Tele-medicine and consulting video
6		6	300 ms	10^{-6}	Medical data transmission with TCP
7		7	100 ms	10^{-3}	Healthcare self-learning systems
8		8	300 ms	10^{-6}	Daily health condition notices
9*[1]		9	300 ms	10^{-6}	Medical image download, etc.

[1] Three asterisked QCI levels are cited as typical examples in this study.

Table 1. QCIs for mHealth in full-IP mobile networks [15].

Service Model for mHealth Applications in Full-IP Mobile Networks

Based on system specifications from the 3rd Generation Partnership Project (3GPP) [14], the full-IP pattern is the irreversible trend for mobile networks, e.g. LTE/LTE-Advanced. To study how to carry mHealth services in all-IP mobile networks, it is necessary to develop a service model first.

QCI Standardization for mHealth Service Model

Typically, each type of application may have particular Quality of Service (QoS) requirements in mobile networks (QoS is the kernel of Quality of User Experience (QoE)). For example, an mHealth user can be engaged in an emergency VoIP call with more stringent requirements for QoS in terms of delay and delay jitter than health condition daily notices and medical image download, while the latter requires a much lower packet loss rate. In order to support multiple QoS requirements in LTE-Advanced networks, different QoS class identifier (QCI) levels are standardized by the 3GPP organization [15]. Each QCI is characterized by a priority (the priority is ranked in terms of acceptable packet delay budgets and packet loss rates), in which these regulated budgets consider the composite effect contributed from both user and control planes in full-IP mobile networks. The packet delay budgets and the acceptable packet loss rates from the QCI regulations determine how an evolved node base station (eNodeB) scheduler handles packets sent over wireless radio resources. For instance, a packet with a higher priority can be scheduled before a packet with a lower priority across the radio air-interface.

In Table 1, QCI No. 1 has a stringent delay budget and a relaxed packet loss rate requirement, and QCI No. 9 requires a small packet loss rate with a loose delay budget. But in 3GPP specifications, QCI No. 5 is ranked as the top priority because it can take account of both packet delay budget and packet loss rate requirements. In fact, there is the trade-off between delay budgets and packet loss rates. This is because these two metrics often interact in a contradictory fashion. This fact has forced the 3GPP organization to make compromises to its standards. This is also why the packet delay budget in QCI No. 5 is set at 100 ms with the smallest packet loss rate, rather than targeting the smallest 50 ms that LTE-Advanced mobile networks can offer. In the most advanced mobile network, QCI No. 5 is the top-priority service that is only dedicated for control information delivery (e.g. system signaling). Even using the

[1]To turn mHealth from a concept to real applications, related technical advances have made the same significant contributions as mobile networks, such as data mining, personalized medicine, pervasive computing, Internet of Things/Webs, Machine-to-Machine communications, and so on, in which two or more of them are working together to collect and analyze information from various sources. This interdisciplinary research for mHealth is shown in the WE-CARE system (which is the first mHealth system issued the government clinical licence in the world) [4, 8].

Channel bandwidth (MHz)	1.4	3	5	10	15	20
Number of available resource blocks (N_{RB})	6	15	25	50	75	100

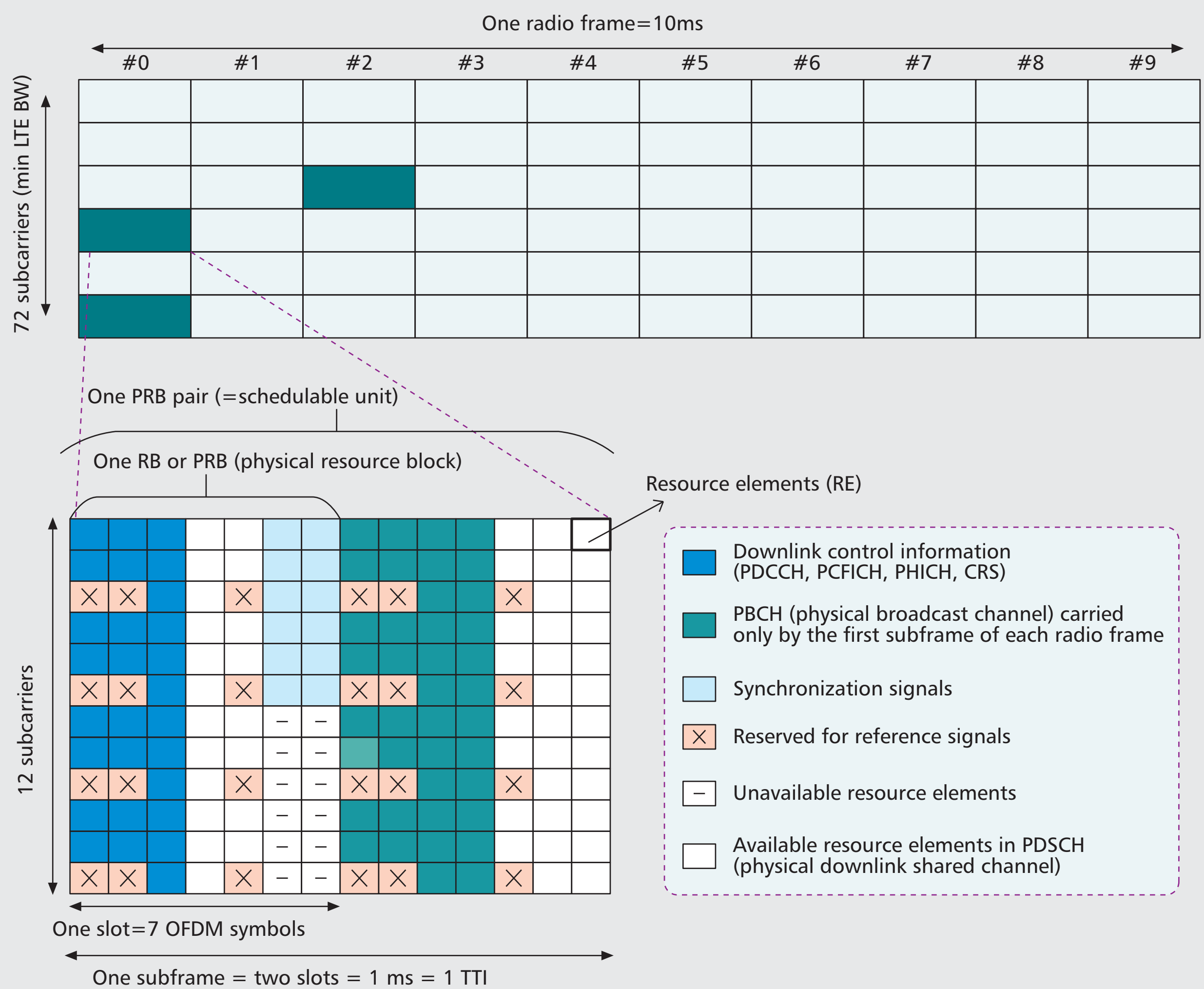


Figure 1. Downlink radio frame structure with four antenna ports in LTE/LTE-Advanced (reproduced from 3GPP specifications [14]).

top-priority of QCI No. 5, some critical mHealth applications are not yet qualified, for instance, the real-time operation video (e.g. a robot-assisted surgical system) that has to wait for new mobile advances available in the future. For typical mHealth applications, each type of service can be mapped with a QCI level correspondingly, as shown in Table 1.

"Better than Best Effort" for mHealth Services

In fact, the QCI issue was originated from the "best effort" pattern in computer IP networks. To backtrack the design intention of IP networking in history, a free and open interconnection mode was deployed for the purpose of "self-healing" that could tolerate any kind of routing failures. Unfortunately, hardware conditions in the 1960s did not allow IP designers to consider a scheduling logic to guarantee service quality in IP networks (at that time, the CPU was upper bounded around 108 KHz) [16]. This unfinished design remaining in IP networks has far-reaching impact on today's service-sensitive IP applications. When mobile networks are evolving into the full-IP pattern (a.k.a. the integration trend of telecommunication and computing), the original "best effort" pattern in IP networks conflicts with the native requirements in mHealth applications. Furthermore, an mHealth user is becoming more sensitive to QoS performance because a radio channel is full of uncertainties/variations in comparison to a wired channel (e.g. cable or fiber). As aforementioned, in terms of the most advanced wireless technologies today, the QCI No. 5 level in Table 1 is the highest service priority that a mobile network can currently offer. Nevertheless, the top-priority service of QCI No. 5 is obtained at the cost of reduced resource utility (i.e. occupying more wireless bandwidth at a lower MCS order). Therefor, radio resources are always limited, which explains why the QCI No. 5 service is only dedicated to control-level information for guaranteeing system-signaling delivery in LTE-A mobile networks. Thus, it is the right time to consider a scheduling function to enable a novel pattern for IP-origin services, with the support of modern hardware advances. Since the "best effort" pattern is still and forever a basis for any new possibilities in full-IP networks, the desired new mode is called "better than best effort." In this article we study how to offer the "better than best effort" services for mHealth (and also applied to other service-sensitive full-IP applications), without degrading resource efficiency in mobile networks.

Full-IP Network Model for mHealth Applications

For investigating how to deliver "better than best effort" services to mHealth, a network model is configured by the most related features of the newly standardized full-IP mobile networks as LTE/LTE-Advanced.

Radio Frame Specification in the Physical Layer

In the physical layer, the radio resource frame format is introduced below. Radio resources are measured in time and frequency dimensions, as shown in Fig. 1. In LTE-Advanced mobile networks, the basic unit of radio resources is a resource element (RE) that consists of one 15 KHz subcarrier and an OFDM symbol, in which an RE is also an intact modulation symbol.[2] One resource block (RB) consists of seven OFDM symbols (= 0.5 ms, namely, one slot) and 12 subcarriers (= 180 KHz), which contain 84 REs. One radio subframe (= two slots, and 10 radio subframes = 1 radio frame) with 12 subcarriers is formed into an RB pair [14], as shown in Fig. 1. In LTE-Advanced networks, the basic scheduling unit is an RB pair (i.e. one transmission time interval (TTI = 1 ms) in the time domain).[3]

In Fig. 1 we observe that some REs are reserved for system information, for example, physical downlink control channel (PDCCH). According to LTE-Advanced standards, only the physical downlink shared channel (PDSCH) in each RB pair is used to carry the downlink full-IP data load. Note that except for the first subframe (i.e. the first RB pair) of each frame (in which 64 REs are used for the physical broadcast channel), the PDSCH in all other subframes has 144 REs available for carrying the full-IP traffic load (see Fig. 1).

According to 3GPP specifications, carrier frequency is around 2.0 GHz with a flexible bandwidth between 1.4 MHz and 20 MHz in LTE-Advanced mobile networks (here we do not consider carrier aggregation). In any setting, some margin of bandwidth must be reserved for the band guard. Because of the band-guard overhead, 5 MHz of bandwidth has only 25 RBs available. In LTE-Advanced networks, RBs should be shared by all users dynamically, including mHealth users.

AMC Mechanism: Self-Adaptation to Stable Service Requirements from mHealth Applications

Most importantly, mHealth services should be guaranteed at a stability threshold, irrespective of variations and uncertainties in volatile radio channels. To handle unreliable channels in mobile networks with limited radio resources, the adaptive modulation and coding (AMC) mechanism is very useful, which can avoid dealing with various channel effects (e.g. interference, pathloss, fading, and other mobility effects) individually.

In LTE-Advanced mobile networks, the AMC mechanism dynamically adjusts a modulation and coding scheme (MCS) level for an RB pair according to CQI feedback. In terms of the CQI feedback mechanism, each RB pair is given a particular CQI level by measuring signal to noise ratio (SNR) of the sounding reference signal (SRS) signals in wideband and/or subband, and a same RB pair may have different CQI levels for different users in a cell.[4] In LTE-Advanced mobile networks, radio resources are usually scheduled in

It is the right time to consider a scheduling function to enable a novel pattern for IP-origin services, with the support of modern hardware advances. Since the "best effort" pattern is still and forever a basis for any new possibilities in full-IP networks, the desired new mode is called "better than best effort."

[2] The OFDM symbol is a notion in the time domain, and an OFDM symbol is around 67 μs. The modulation symbol is a concept in information theory that carries bit flows. For example, a QPSK modulation symbol can carry 2 bits.

[3] Transmission Time Interval (TTI) is a parameter in LTE-Advanced mobile networks (and other digital telecommunication networks) related to the encapsulation of data from higher layers into data frames. In fact, a data frame (different from a radio frame concept in time domain) is a transport block that contains a MAC Protocol Data Unit (PDU). A transport block appending with CRC bits forms a code block, or is divided into two or more code blocks (a code block must be between 40 and 6144 bits long), in which each code block is encoded in turbo or convolution codes with rate matching. After the rate-matching, each individually processed code block has to be concatenated into a single codeword. A codeword is scrambled, and converted into a modulation symbol. These well-encapsulated modulation symbols are mapped into layers to be sent over the air within each TTI. In fact, a TTI refers to the length of an independently decodable transmission on the radio channel. In LTE-Advanced mobile networks, one TTI (1 ms) is also equal to the length of one radio subframe.

[4] Please refer to [14] for detailed descriptions about Reference Signal (RS) measurement for enabling CQI feedback mechanisms in LTE-Advanced networks. A CQI level reflects the overall situation, incorporating the effects that a radio channel is suffering from, for example, fading, interference, pathloss, and other mobility effects. In mobile networks, the CQI feedback may not represent exact radio channel conditions because of its SNR measuring method limitations and/or feedback delay. Unfortunately, this limitation is not avoided in any telecommunication systems. Thus far, CQI measurement is the only method used for weighting radio quality in deployed mobile networks.

I_{mcs}	Q_m	CR × 1024	I_{tbs}	SE	I_{cqi}
0	2 (QPSK) SNR threshold = 13.6 dB	33	0	0.0644	LIM
1		78	1	0.1523	1
2		99	2	0.1933	LIM
3		120	3	0.2344	2
4		157	4	0.3066	LIM
5		193	5	0.3770	3
6		251	6	0.4903	LIM
7		308	7	0.6016	4
8		449	8	0.8770	5
9		602	9	1.1758	6
10	4 (16QAM) SNR threshold = 20.6 dB	301	9	1.1758	LIM
11		340	10	1.3281	LIM
12		378	11	1.4766	7
13		434	12	1.6954	LIM
14		490	13	1.9141	8
15		553	14	2.1602	LIM
16		616	15	2.4063	9
17	6 (64QAM) SNR threshold = 26.8 dB	411	15	2.4063	LIM
18		466	16	2.7305	10
19		517	17	3.0293	LIM
20		567	18	3.3223	11
21		617	19	3.6153	LIM
22		666	20	3.9023	12
23		719	21	4.2128	LIM
24		772	22	4.5234	13
25		823	23	4.8222	LIM
26		873	24	5.1152	14
27		911	25	5.3379	LIM
28		948	26	5.5547	15
29	2 (QPSK)	Reserved[1]			
30	4 (16WAM)				
31	6 (64QAM)				

[1] Notations: CQI index = I_{cqi}; MCS index = I_{mcs}; transport block size (TBS) index = I_{tbs}, modulation orders = Q_m; code rate = CR; spectrum efficiency = SE (bits/symbol) (where the symbol refers to a modulation symbol); resource element (RE) = carrying an intact modulation symbol; SNR threshold (in dB); Linear Interpolation Method = LIM.

Table 2. A mapping solution between CQI levels, MCS selection, and spectrum efficiency per resource element (RE) (this suggestion is coherent with LTE-Advanced standards [14]).

the unit of RB pairs. The physical size of an RB pair is fixed in the time and frequency domains, but its load-carrying capacity is calculated by its selected MCS level, and an MCS selection is directly determined from the QCI level by the AMC mechanism (see Table 2). In LTE-Advanced networks, the AMC mechanism is dedicated to connecting the CQI feedback with the adaptive selection of MCS levels. However, there are only 15 CQI standard levels in LTE/LTE-Advanced. How can these 15 CQI levels be mapped with 29 MCS criterion options in practice?

To the best of our knowledge no public solution exists for the mapping between the AMC mechanism and CQI levels. Because the mapping is necessary for any kind of eNodeB scheduler, we suggest a mapping solution in Table 2 using the linear interpolation method (LIM) following 3GPP LTE-Advanced specifications [14] (the mapping is also applied in our simulation tests later). When an MCS is selected for a user through the AMC mechanism, carrying capacity per resource unit (RB or RE) is determined, because the bit rate of the radio signal emitted from a transmitter is set up by the modulation format and channel encoding rate. In order to guarantee the QoS of different services with distinct traffic patterns and requirements, it is worthwhile incorporating AMC into the scheduling process so that radio resource utilization can be maximized. For example, as a user moves from a cell edge to its center, it is reasonable to adjust its $I_{mcs} = 0$ (with spectrum efficiency 0.064 b/symbol) to $I_{mcs} = 28$ (with spectrum efficiency 5.55 b/symbol) to adapt to CQI variation. This example reveals that an MCS selection in AMC of an RB pair implies CQI information because they are mapped with each other (see Table 2).

SMART Algorithm to Enable mHealth Applications in Full-IP Mobile Networks

To maintain "better than best effort" services for mHealth applications in different scenarios, a scheduling strategy should be robust and agile to handle variations and uncertainties over unreliable radio channels. Moreover, in any mobile network, wireless radio resources are always limited. Hence, it is significant to exploit the potential of a scheduling algorithm, so as to offer just-enough radio resources for user-specific service requirements as mHealth. To design such a scheduling strategy, let us first briefly investigate the existing scheduling algorithms.

Overview of Conventional Scheduling Strategies

In terms of the scheduling metric definition, these conventional scheduling algorithms can be classified into two types: CQI-based scheduling or QCI-based scheduling.

CQI-Based Scheduling — The goal of channel quality indicator (CQI)-based scheduling is to maximize "system throughput" by fully utilizing CQI information, for example, in the Max C/I (carrier-to-interference ratio) scheduling algorithm [17]. For a user in an LTE-Advanced mobile network, wideband and subband CQI information are dynamically reported to the eNodeB scheduler for properly responding to variations in radio channels. Then CQI information is implanted into the scheduling process of a CQI-based scheduling algorithm [18]. CQI-based scheduling delivers a fairness issue to co-existing users because it allows higher-priority users to preempt resources. To handle the issue of fairness, the authors in [19] investigated proportional fair scheduling and opportunistic fair scheduling, which consider historical scheduling information (e.g. average throughput). There also exist studies on how to improve CQI measurements for better scheduling performance (see [20] for an investigation of CQI-related interference and power control issues). In reality, the effectiveness of a scheduling strategy is greatly dependent on how channel quality indicator (CQI) information is used. Obviously, CQI information should also be embedded into a scheduling strategy to support diversities in mHealth applications. Most importantly, mHealth services should be guaranteed at a stability threshold, regardless of variations and uncertainties in volatile radio channels. Consequently, CQI information usability in a scheduling scheme should be self-adaptive to maintain stable services for mHealth applications.

QCI-Based Scheduling — The goal of QoS class identifier (QCI)-based scheduling is to intelligently select a user's IP packets to make better use of available radio resources. In LTE-Advanced networks, each type of service is regulated with a specific QCI level under two basic requirements: packet delay budgets and packet error loss rates (Table 1). In LTE-Advanced networks, all service traffic is carried over a radio air-interface in original IP packets. This IP delivery pattern raises two challenges. One is the mentioned conflict between packet delay and packet loss. Usually the design strategy of a scheduling algorithm is dedicated to one of the requirements, at the cost of the other, e.g. in the deadline-driven scheduling algorithm [21]. This phenomenon is also considered in the 3GPP standards (see Table 1). As mentioned already, an mHealth user is service-sensitive, requiring health information to be "accurate" and "real-time" [7, 8]. The other challenge is the "best effort" IP pattern in QCI-based scheduling. So far there is no solution to guarantee service quality will match QCI requirements exactly in a scheduling process. To date, QCI parameters could not take effective control of scheduling, which is different from a CQI factor in a CQI-based scheduling algorithm [22]. In fact, QCI-based scheduling algorithms try to approach QoS requirements in "best effort," and there are no QoS-guaranteed measures available now. For mission-critical applications such as mHealth, it is a must to find a solution that can offer services better than the "best effort" pattern, a solution called "better than best effort" in this article.

As observed above, it is evident that a CQI-based scheduling algorithm is designed to maximize utilization of telecommunication systems, while a QCI-based scheduling algorithm is to designed to enhance service quality that is a concern originally from computer IP networks. Thus, scheduling algorithm design in full-IP mobile networks is a cross-discipline topic between telecommunication engineering and computer science. Obviously, if an mHealth user needs the top-priority service in terms of both packet loss rate and packet delay budget, what does this mean to coexisting users? This case reveals that scheduling design is also a multiuser optimization problem in a mobile system. Thus, it is necessary to minimize or remove the discrepancy between CQI-based scheduling and QCI-based scheduling objectives. Based on in-depth investigation of the topic, we propose a "Service-Mapping Adaptive Radio Transmission (SMART)" scheduling algorithm, in which the conceived self-adaptive mechanism can seamlessly combine aspects of telecom related CQI and IP related QCI. In this proposal, radio resources can be scheduled to offer "better than best effort" services for an mHealth user, without adversely affecting other co-existing mHealth/non-mHealth users. In other words, it is designed to achieve the best-fit mapping between radio resource utilization and full-IP packet-switching services, so as to enrich mHealth applications. We now describe our proposal.

In any mobile network, wireless radio resources are always limited. Hence, it is significant to exploit the potential of a scheduling algorithm, so as to offer just-enough radio resources for user-specific service requirements as mHealth.

Service-Mapping Adaptive Radio Transmission (SMART) Scheduling Algorithm

In addition to combining aspects of QCI and CQI, the proposed scheduling algorithm should have low computational complexity, under the constraint of the time-sensitive requirement in mHealth applications. For offering "better than best effort" services for mHealth applications, a scheduling metric plays a key role. To design an effective scheduling metric, the QCI requirements in Table 1 (i.e. the packet delay budget and the packet loss rate limitations) should be taken into account. Furthermore, it should be noted that various effects in a user radio channel, e.g. fading, interference, pathloss, and so on, can be overall quantified by a CQI level [14]. In LTE-Advanced networks, CQI results directly determine which MCS level should be selected through the AMC mechanism [14]. If the MCS level information is self-aware of the scheduling procedure, the corresponding CQI information is also implicitly applied into radio resource assignment. As we investigated earlier, the AMC mechanism is a link adaptation function that chooses a suitable MCS level in accordance with a reported CQI level. This AMC mechanism can be used for this desired function. Of course, if we can get a scheduling metric as below, it is approaching this objective:

$$U = QCI(delay + loss) + CQI(through\ AMC), \quad (1)$$

where the valid range of U may be normalized in Eq. 2. In reality, the scheduling metric is available when properly using factors in Tables 1 and 2 (related mathematical analysis is skipped for better readability):

$$0 \leq U \leq 1. \quad (2)$$

1. Launch an mHealth service request from a user UE_i from K users.
2. Become aware of QoS-benchmarked requirements of this service (i.e. parameterize QCI requirements).
3. Drop IP packets that are missing QoS-benchmarked requirements to save valid radio resources as much as possible.
4. Capture the MCS information of available RB pairs of the user UE_i.
5. Calculate and rank the value of U for IP packets over available RB pairs within the scheduling period.
6. While in the *valid range* [0, 1] of scheduling metric U.
    6.1 Schedule packet delivery with the highest U first that stands for the preferable mapping between an IP packet's QCI requirements and an RB pair's MCS selection.
    6.2 Repeat until either all IP packets are uploaded, or all RB pairs occupied, whichever condition is met first.
7. Release RB pairs whose packets are discarded.
8. Released RB pairs may be shared by newly-arrived IP packets.

Algorithm 1. SMART scheduling to enable "better than best effort" for mHealth in full-IP mobile networks.

In LTE-Advanced networks, radio resources are scheduled in the unit of RB pairs. As shown in Table 2, the load-carrying capacity of an RB pair is dynamic to its selected MCS level.[5] Clearly, U in Eq. 1 enables a scheduling strategy to be dynamically adapted to variations in radio resources in relation to QoS requirements. In general, a scheduling metric is an interface between the scheduling objective and the scheduling process. The quality of a scheduling metric has direct impact on scheduling effectiveness and computational complexity. Compared with scheduling metrics in the literature [14], we observed that the scheduling metric U has the following favorable characteristics.

CQI and QCI Awareness: U is a self-adaptive scheduling metric that consists of QCI and MCS parameters (even though CQI information is implied in MCS, but its effectiveness is almost the same as the CQI feedback report in this proposal). The self-adaptive scheduling metric makes the proposal more effective and valid for mHealth applications, because it can connect resource assignments with QoS-benchmarked requirements, and find a best-fit match between them. Based on this feature, the conceived scheduling algorithm is a cross-layer scheme as well.

Normalized Valid Range: The valid range of the self-adaptive scheduling metric is normalized to [0, 1] for lower computational complexity and higher scheduling efficiency. In addition, the weight addition can contribute lower scheduling computational complexity. This feature is more significant to time-sensitive mHealth applications.

Standard Compliance: In terms of the scheduling metric, parameters are abstracted from industry standards as shown in the QCI values in Table 1, and the mapping between MCS indices and CQI levels in Table 2. This feature can enrich mHealth services in future mobile ecosystems.

Using the self-adaptive scheduling metric U in Eq. 1, a novel scheduling algorithm is constructed for mHealth services in Algorithm 1.

In contrast to existing solutions, the major contribution of this scheduling algorithm is its ability to function the mapping between CQI information and QCI requirements, and schedule appropriate radio resources to provide the desired "better than best effort" for mHealth applications. Furthermore, the self-adaptive scheduling does not involve unacceptable signaling overhead. This is because both QCI parameters (packet delay budgets and packet loss rates) of each type of service and an MCS-level selection of each RB pair of a user are self-aware to the eNodeB scheduler in the LTE-Advanced system. The additional merit of the proposal is its lower runtime complexity $O(K \log K)$.

Parameters	Value
Cellular layout hexagonal grid	3-sector sites
Carrier frequency	2.0 GHz
Bandwidth	5 MHz
Sampling rate	30.72 [MHz]
eNodeB height	30 m
Terminal height	1 m
eNodeB power	46 dBm
Terminal power	23 dBm
Noise power	–164 dBm/Hz
Number of coexisting users	100
User mobility	0, 30 Km/h
Wireless channel	Lognormal channel

Table 3. Parameters in system-level simulation experiments.

PERFORMANCE EVALUATION

To evaluate the proposal for mHealth applications, simulation experiments were conducted in a system-level C-programming simulation platform by employing multi-core parallel-computing technologies (which were developed to assist Chinese

[5]An MCS level selection depends on radio signal quality. A higher-order MCS format could be deployed for a higher bit rate in robust channel conditions, while a low-order modulation format (e.g., BPSK or QPSK) is applied to strengthen traffic transmission. For instance, in terms of lower-order BPSK and QPSK in LTE/LTE-Advanced systems, each RE is an independent modulation symbol that occupies a single quadrant in the I-Q plane. This is why a lower-order MCS format is used to strengthen radio transmission (e.g., for carrying control messages) and/or against poor radio channel conditions, since it can avoid causing any confusion to demodulation constellation. In other words, the larger distances between constellation points are, the higher possibilities of the radio signals are for demodulation and decoding.

deployment field tests).[6] To demonstrate the effectiveness of the proposal, it was compared with three alternatives: Ref. 1 for considering delay first [21]; Ref. 2 for citing proportional fairness scheduling [19]; and Ref. 3 for considering traffic behavior [22]. We chose three typical mHealth services for experiments in the system-level simulation platform: emergency VoIP call in QCI No. 1 at 12.2 kb/s adaptive multi rate (AMR) with 40 percent voice activity factor in the full buffer model; telemedicine consultation in QCI No. 5 at the guaranteed bit rate (GBR) of 64 kb/s in the real-time video traffic model; and medical image download in QCI No. 9 at the peak bit rate (PBR) of 384 kb/s in the HTTP model [4], in which the traffic model of each service can be customized in real deployment scenarios. In these experiments, user distribution is implemented by using the conventional random dropping approach. In these simulation experiments, each user is only assigned one type of service for a straightforward presentation of our proposal (leaving single-user multi-service scenarios for future study). The parameters of the simulation experiments are listed in Table 3.

Figure 2 shows that the proposal can exceed the alternatives in terms of system throughput. On average, the SMART algorithm can contribute 31.25 percent system throughput gain in contrast with the alternates, because the SMART algorithm also gives priority to system throughput (functioned as a precondition). In our proposal, CQI information can be inserted into the scheduling process by the AMC mechanism, while trying to guarantee QCI requirements, meaning that mHealth services can get the "better than best effort" services according to QoS requirements, in contrast to alternative scheduling strategies. This performance gain is a benefit of the unique self-adaptive design from the seamless combination of QCI and CQI information.

Regarding user experience, we mainly investigate service satisfaction and user fairness. Usually, service satisfaction of a single user may have immediate impact on the service quality of coexisting users [14]. Figure 3 shows a comparison of the "user fairness" and "service satisfaction" between our proposal and the alternatives. To be specific, "user fairness" focuses on how to allocate radio resources so as to maintain fairness among users, but "service satisfaction" aims to guarantee benchmarked QoS for a single user, which may intentionally sacrifice user fairness to serve the particular mHealth application. We define two performance criteria for their comparison: the degree of user fairness is the probability of accessing radio resources, and the degree of service satisfaction is the percentage of QoS-benchmarked services. Compared with "user fairness," "service satisfaction" is new but vital for mHealth and other full-IP mobile services, which is the major index to indicate the performance of the "better than best effort" service style.

The results in Fig. 3 demonstrate that the SMART algorithm can achieve *service satisfaction* for mHealth services and maximal *efficacy* of radio resources with acceptable relaxation of "user fairness." This performance contribution results from three favorable factors: first, high-quality radio resources are made available to high-priority IP packets first; second, the physical layer CQI information is also emphasized adequately by embedding the AMC mechanism into the self-adaptive scheduling metric; third, packets with larger delay budgets can be scheduled on idle radio resources by taking advantages of delay toleration characteristics. This proposal can also be applied to other mission-critical applications, such as mobile banking.

[6]In mobile networks, most radio technologies can be tested in a link-level (physical layer) simulation. To test and optimize scheduling algorithms, simulation experiments have to be performed at the system level (i.e., network context).

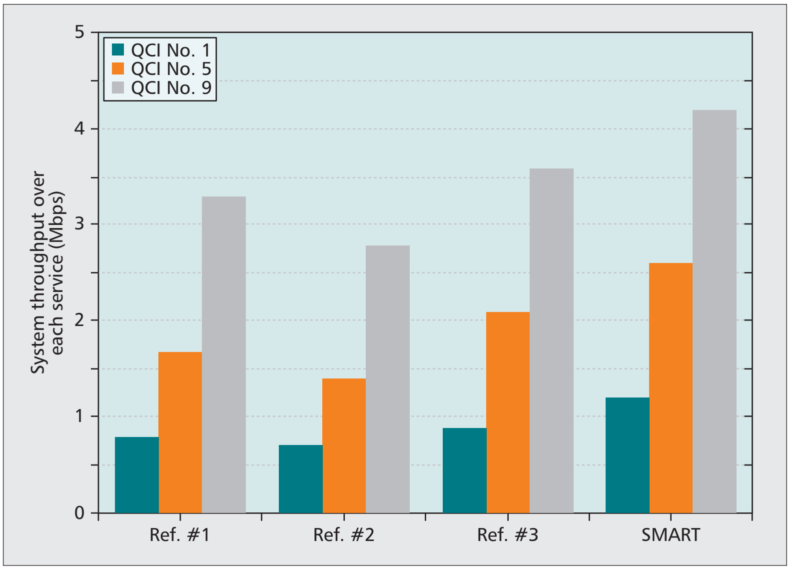


Figure 2. System throughput comparison.

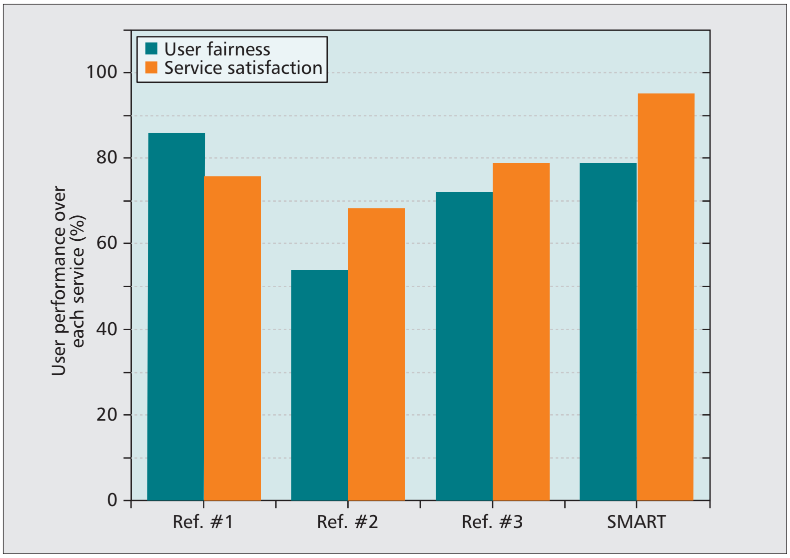


Figure 3. User experience comparison.

Discussion and Conclusion

In this article we focused on how to adaptively utilize radio resources in order to enable "better than best effort" services. This study is part of

With the proliferation of mission-critical applications such as mHealth, how to build an E2E QoS-guaranteed connection is a hot research topic. Hence, there is extensive attention paid to studying system protocols and routing algorithms in the fiber-based mobile backbone, and scheduling designs in radio access.

offering end-to-end (E2E) QoS-guaranteed services for mHealth applications in all-IP mobile networks. In general, a mobile network consists of two major parts: access and the backbone. When setting up an E2E-QoS connection, it has to cross over the fiber-based mobile backbone and the radio access. In today's technical advances, the challenge presented by an E2E QoS-guaranteed connection is still unsolvable, even in a pure wired network with the help of the traditional circuit switch. Without any doubt, this challenge becomes more critical in an all-IP mobile network with a hybrid of fixed backbone and radio access, under constraints of the packet-switching pattern and unreliable radio channels. With the proliferation of mission-critical applications such as mHealth, how to build an E2E QoS-guaranteed connection is a hot research topic. Hence, there is extensive attention paid to studying system protocols and routing algorithms in the fiber-based mobile backbone, and scheduling designs in radio access [13]. Furthermore, it is worthwhile exploring the "better than best effort" pattern when the packet-switching full-IP applications are ubiquitous.

Driven by these multiple incentives, we presented a comprehensive study of scheduling design for "better than best effort" services in radio access. Our simulation experiments demonstrated that the devised SMART scheduling algorithm is an effective and efficient candidate to provide the "better than best effort" services to mHealth applications (as well as other types of full-IP mobile applications) under multiple constraints, e.g. limited radio resources, volatile radio channels, mobility effect, pure IP packets, multiuser scenarios, and so on. However, there are some system-level factors that might affect the proposal, e.g. ARQ in RLC and HARQ (Hybrid ARQ) in MAC, TCP multipath over mobile, traffic/channel profiles, and so on, which are left for future study.

References

[1] World Health Organization, "Noncommunicable Diseases Country Profiles 2011," WHO Report, Sep., 2011.
[2] Ministry of Health of China, *2013 China Health Statistics Yearbook*, Beijing, China, Apr. 2014.
[3] B. Kayyali, D. Knott, and S. V. Kuiken, "The Big-Data Revolution in US Health Care: Accelerating Value and Innovation," Tech. Report Mckinsey, Apr. 2013.
[4] A. Huang *et al*., "WE-CARE: An Intelligent Mobile Telecardiology System to Enable mHealth Applications," *IEEE J. Biomed. Health Informatics*, vol. 18, no. 2, Mar. 2014, pp. 693–702.
[5] S. Adibi, "Biomedical Sensing Analyzer (BSA) for Mobile-Health (mHealth)-LTE," *IEEE J. Biomed. Health Informatics*, vol. 18, no. 1, Jan. 2014, pp. 345–51.
[6] F. Collins, "The Real Promise of Mobile Health Apps: Mobile Devices Have the Potential to Become Powerful Medical Tools," *Scientific American Mag*., vol. 307, no. 1, July 2012.
[7] S. Adibi, A. Mobasher, and T. Tofigh, "LTE Networking: Extending the Reach for Sensors in mHealth Applications," *Wiley Trans. Emerg. Telecom. Tech*., vol. 25, no. 7, July 2014, pp. 692–706.
[8] A. Huang *et al.*, "System Light-Loading Technology for mHealth: Manifold-Learning-based Medical Data Cleansing and Clinical Trials in WE-CARE Project," *IEEE J. Biomed. Health Informatics*, vol. 18, no. 5, Sept. 2014, pp. 1581–89.
[9] WHO and ITU, *mHEALTH for Non-Communicable Diseases (NCDs) Initiative*, Apr. 2014.
[10] Mordor Intelligence Corp., *Global Wearable Medical Device Market — Growth, Trends and Forecasts (2014-2019), Market. Analy. Report*, Aug. 2014.
[11] Z. Wang *et al*., "Modeling and Analysis of QoS Class Mapping for Hybrid QoS Domains using Flow Aggregate," *Proc. IEEE 9th Int'l. Conf. Wireless Commun. Mobile Computing*, Sardinia, Italy, July 1-5, 2013, pp. 503–08.
[12] R. S. H. Istepanian, A. Alinejad, and N. Y. Philip, "Medical Quality of Service (m-QoS) and Quality of Experience (m-QoE) for 4G-Health Systems," *Multimedia Networking and Coding*, co-authored by Reuben A. Farrugia and Carl J. Debono, IGI Global, Dec. 2012, pp. 359–76.
[13] 3GPP TS 23.207, *3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; End-to-End Quality of Service (QoS) Concept and Architecture*, http://www.3gpp.org/DynaReport/23207.htm, Release 11, accessed Oct. 30, 2014.
[14] 3GPP TS 36.321, *3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) Protocol Specification*, http://www.3gpp.org/DynaReport/36321.htm, Release 12, accessed Oct. 30, 2014.
[15] 3GPP TS 23.203, *3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Policy and Charging Control Architecture*, http://www.3gpp.org/DynaReport/23203.htm, Release 12, accessed Oct. 30, 2014.
[16] L. G. Roberts, "The Evolution of Packet Switching," *Proc. IEEE*, vol. 66, no. 11, Nov. 1978, pp. 1307–13.
[17] L. C. Wang and W. J. Lin, "Throughput and Fairness Enhancement for OFDMA Broadband Wireless Access Systems Using the Maximum C/I Scheduling," *Proc. IEEE 60th Conf. Veh. Technol.*, Los Angeles, USA, vol. 7, Sept. 14-17, 2004, pp. 4696–4700.
[18] M. A. Awal and Lila Boukhatem, "Dynamic CQI Resource Allocation for OFDMA Systems," *Proc. IEEE Conf. Wireless Commun. Netw*., Quintana-Roo, Mexico, March 28–31, 2011, pp. 19–24.
[19] I. G. Fraimis and S. A. Kotsopoulos, "QoS-Based Proportional Fair Allocation Algorithm for OFDMA Wireless Cellular Systems," *IEEE Commun. Lett.*, vol. 15, no. 10, Oct. 2011, pp. 1091–93.
[20] R. Aggarwal *et al*., "Joint Scheduling and Resource Allocation in the OFDMA Downlink: Utility Maximization Under Imperfect Channel-State Information," *IEEE Trans. Signal Process*., vol. 59, no. 11, Nov. 2011, pp. 5589–5604.
[21] J. Lee *et al.*, "Dynamic Parameter Adaptation for M-LWDF/M-LWWF Scheduling," *IEEE Trans. Wireless Commun*., vol. 11, no. 3, Mar. 2012, pp. 927–37.
[22] F. Capozzi *et al*., "Downlink Packet Scheduling in LTE Cellular Networks: Key Design Issues and a Survey," *IEEE Commun. Surv. Tutor*., vol. 15, no. 2, Second Quarter 2013, pp. 678–700.

Biographies

ANPENG HUANG [M'05–SM'14] received his Ph.D. degree from Peking University, China, in June 2003. Since November 2007 he has been an associate professor at Peking University (PKU), China. His research interests include mobile health, mobile networks, and optical networks. He has published more than 60 journal and conference papers and is the holder of 36 patents and US pending patents (with PCT application). He was the winner of 42nd Geneva Invention Award in 2014, and the founder of the Mobile Health Lab at Peking University.

LINZHEN XIE [M'13] has been a professor at Peking University in China since 1978. He received his BS degree from Peking University in 1963, and was a visiting scholar at Berkeley EECS, University of California, USA, from 1980 to 1982. He is the founder of the State Key Laboratory of Advanced Optical Communication Systems & Networks at Peking University, China. He has published more than 140 papers in journals and at conferences in these areas, and is the holder of 21 patents.

Call for Papers
IEEE Wireless Communications Magazine

Social Networks Meet Next Generation Mobile Multimedia Internet

Background

With ever growing popularity and widespread adoption of mobile social applications among users, the traffic handled by mobile networks and the Internet has grown significantly. While researchers have been making advances in the study of social networks and independently in the area of next generation wireless networks, very little attention has been given to the interplay between the two, and their impact on each other and society. The interplay between social networks and mobile networks is compounded by the fact that advances in smart handheld devices and those in wireless technologies have paved the way for increasing bandwidth catering to very high data rates. It is entirely likely that such advances in turn could lead to novel social applications not yet thought of. For example, new social applications could emerge in the area of social health, social games, location-based social applications, real-time social collaboration applications, or real-time massively multi-player, multimedia, 3-D, role playing games, and m-commerce. Although there is no standard definition of what constitutes a mobile multimedia Internet, for the purposes of this proposal, we stipulate that the mobile multimedia Internet is one that facilitates the convergence of various wireless networks and protocols, including 3G, 4G, and next generation 5G wireless networks and beyond, wireless LANs, voice over IP (VoIP), cognitive radio, software-defined networking, and cloud networks and protocols.

The emergence of research in the area of Big Data and its applicability to social networking can hardly be overemphasized. Together with the evolution and deployment of cloud RANs, cloud-based and software-defined networks could lead to new forms of Big Data, for example, involving user locations, usage patterns, user mobility, and other user-specific behaviors. Also, the ability on the part of next generation networks to facilitate peer-to-peer and ad hoc networking paves the way for new forms of interactions with mobile social networks and applications. New application areas such as location-based social networking applications (LBSNAs) and Internet of Things (IoT)-based social networking (IoTBSNA) are maturing and gaining ground.

Through this Feature Topic we intend to create a venue to bring together researchers and practitioners from different disciplines, especially computer and information sciences and next generation mobile/wireless multimedia Internet/networks as well as other related disciplines to share, exchange, learn, and develop preliminary results, new concepts, ideas, principles, and methodologies, aiming to advance mobile social networks in the new generation of Information and Communication Technologies enabled by Web 3.0, also referred to as next generation intelligent web.

Scope of Contributions

Original papers are solicited that highlight new advances and directions in cross-domain convergence and blending of mobile networks and social networks to address challenges and develop opportunities including, but not limited to, the following topics:

- Mining Big Data generated by mobile social networks to understand and formulate models social theories.
- Impact of social network growth models on the mobile multimedia Internet.
- New social applications capable of exploiting the new features that next generation networks will bring about.
- Infrastructure to facilitate awareness of the capabilities of social networks within mobile networks.
- Mobile network aware social media service architecture, middleware, and framework.
- The role of IP Multimedia Subsystem in the interplay between social networks and mobile networks;
- Resilient mobile networks leveraging information diffusion, containing contagion, and disincentivization concepts.
- Network-aware search, ranking, and recommendation.
- Mobile social network data collection, analysis, trends, tools, and applications.
- Privacy/security in mobile social applications.
- Location-aware mobile social computing.
- Citizen sensing applications — mobile sensing for community actions.
- Mobile social computing applications in crisis management.

Submissions

This Feature Topic solicits original work that must not be under consideration for publication in other venues. Authors should refer to *IEEE Communications Magazine* submission guidelines at http://www.comsoc.org/commag/paper-submission-guidelines for information about content and formatting of submissions. Manuscripts must be written in English and contain substantial tutorial content and be readable by a broad general audience working in other fields. All articles must be submitted through IEEE Manuscript Central (http://mc.manuscriptcentral.com/commag-ieee) by the submission deadline. Submit manuscripts to the category "October 2015: Social Networks Meet Next Generation Internet."

Schedule for Submissions

- Manuscript Submission Deadline: May 15, 2015
- Author Notification: July 10, 2015
- Final Manuscript Due: August 1, 2015
- Publication Date: October 2015

Seshadri Mohan
University of Arkansas at Little Rock
sxmohan@ualr.edu

Nitin Agarwal
University of Arkansas at Little Rock
nxagarwal@ualr.edu

Ashutosh Dutta
AT&T
ad5939@att.com

Advertisers' Index

ADVERTISING SALES OFFICES

Closing date for space reservation: 15th of the month prior to date of issue

Currently Scheduled Topics

Topic	Issue Date	Manuscript Due Date
Software Defined Radio: 20 Years Later	September 2015	March 1, 2015
Future Railway Communications	October 2015	April 1, 2015
Social Networks Meet Next Generation Mobile Multimedia Internet	October 2015	May 15, 2015
Software Defined Wireless Networks (SDWN)	November 2015	April 1, 2015

www.comsoc.org/commag/call-for-papers

Fraunhofer
Heinrich Hertz Institute

We invite you to visit us at OFC 2015 German Pavilion booth 1219

70 GHz COHERENT RECEIVER FRONTEND

- Coherent optical receiver with up to 70 GHz bandwidth in a compact, 1RU 19''- chassis
- Test and measurement equipment for development of multi-terabit transmission systems and components
- Four optical extender heads for direct connection to high-bandwidth oscilloscopes
- O/E converter for detection of arbitrary optical waveforms
- C+L-band operating wavelength range
- High-resolution optical spectrum measurements

www.hhi.fraunhofer.de/corx