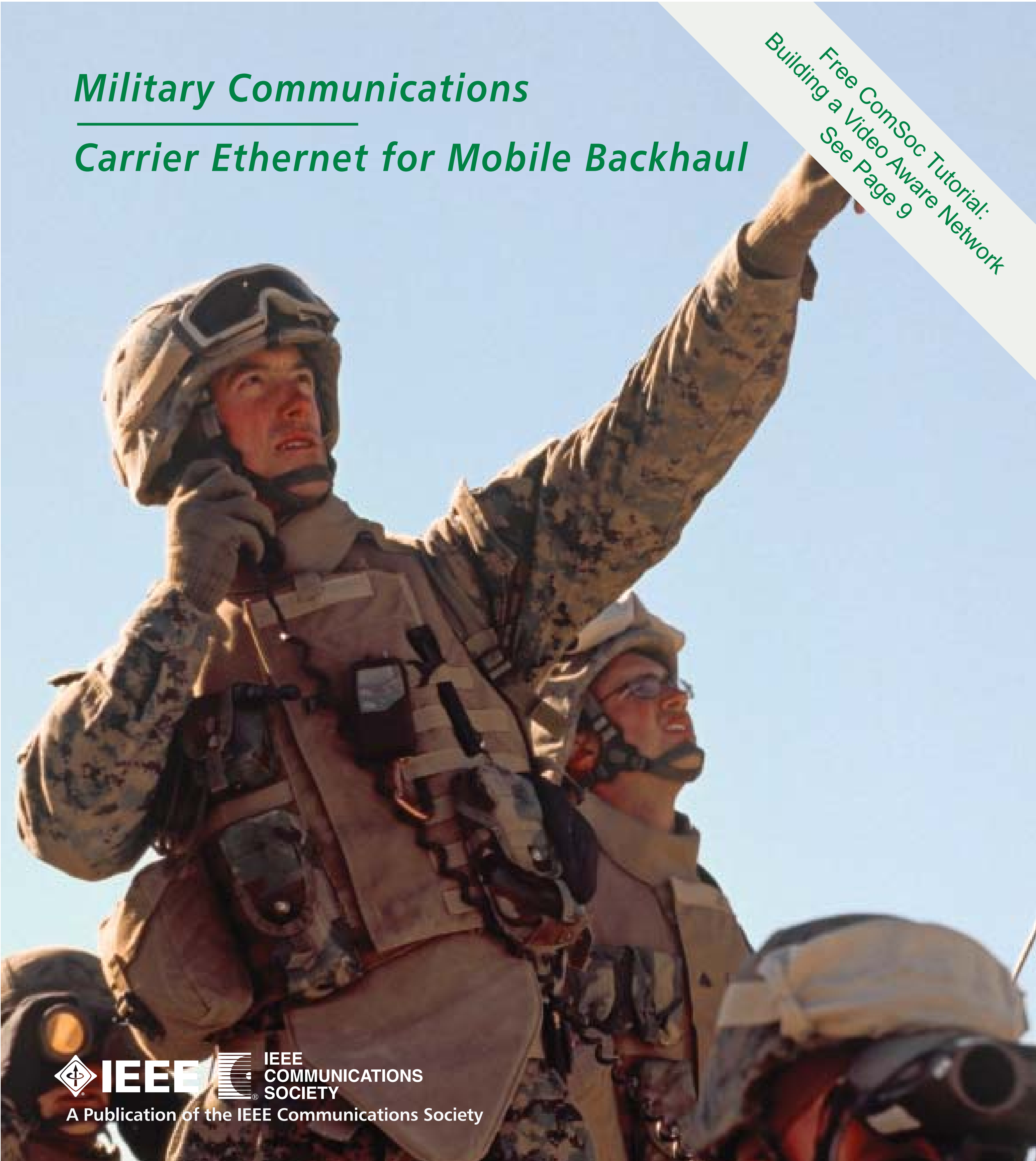
THIS MONTH'S DIGITAL DELIVERY OF IEEE COMMUNICATIONS MAGAZINE

SUPPORTED BY:

October 2010, Vol. 48, No. 10

IEEE Communications MAGAZINE

www.comsoc.org



IEEE Communications MAGAZINE

October 2010, Vol. 48, No. 10
www.comsoc.org/~ci

HISTORY OF COMMUNICATIONS

EDITED BY MISCHA SCHWARTZ

MILITARY COMMUNICATIONS

GUEST EDITORS: TORLEIV MASENG, RANDALL LANDRY, AND KENNETH YOUNG

CARRIER ETHERNET FOR MOBILE BACKHAUL

GUEST EDITORS: DAVID HUNTER, ALAN MCGUIRE, AND GLENN PARSONS

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/ci. 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, Steve Gorshe, PMC-Sierra, Inc., 10565 S.W. Nimbus Avenue, Portland, OR 97223; tel: +(503) 431-7440, e-mail: steve_gorshe@pmc-sierra.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.

ADVERTISING: Advertising is accepted at the discretion of the publisher. Address correspondence to: Advertising Manager, *IEEE Communications Magazine*, 3 Park Avenue, 17th Floor, New York, NY 10016.

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. Instructions can be found at the following: http://dl.comsoc.org/livepubs/ci1/info/sub_guidelines.html. For further information contact Sean Moore, Associate Editor-in-Chief (smoore-phd@ieee.org). All submissions will be peer reviewed.

BPA INTERNATIONAL

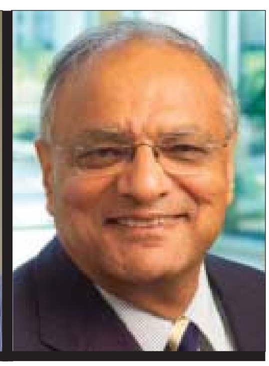
INDUSTRY: A KEY COMSOC CONSTITUENCY

BYEONG GI LEE

The IEEE Communications Society (ComSoc) has two major goals prescribed in its Constitution: 'the scientific and educational activities' for advancing the theory, practice and application of communications engineering and related arts and sciences in serving the needs of the industry and our constituency; and 'the professional activities' for promoting high professional standards, development of competency, and advancement of its members. What is significant about these Society goals is that we advance communications technology and the profession so that affordable communications networks and services may become available to the people of all nations. If these goals are successfully achieved, they benefit humanity by enabling people to connect with each other and to reach out for knowledge and services that will enhance their quality of life and in turn make this a better world.

In practice, it is the Information and Communications Technology (ICT) industry that provides the functional means for realizing the goals of ComSoc. That is why industry is a key ComSoc constituency. The ICT industry, through its manufacturing and operating functions, deploys communication networks and provides communications services to serve people throughout the world. The ICT industry has a wide impact, from the way people run the most sophisticated and complex enterprises to the hope of eliminating isolation and poverty from the most underserved citizens of the world. The breakthroughs and progress in communications technology combined with the globalization of the ICT industry has led to profound changes that have improved and will continue to improve the lives of billions of people on our planet.

ADAM DROBOT



SHRI GOYAL

STAN MOYER

HARVEY FREEMAN

Therefore it is fundamentally important for ComSoc to maintain a close collaboration with the ICT industry in all its activities. We need to exert the effort to move ComSoc closer to industry and vice versa, thereby securing industry as a vibrant participant in ComSoc. Historically, ComSoc had operated with active involvement and cooperation of industry in terms of volunteers, technical committees, publications, and conferences. This was the case until we encountered major changes in the transformation of the ICT industry following the divestiture and privatization of telecommunication operators and the rise of new manufacturers. The present ecology in the ICT industry is dramatically different from the 'good old days' and as a consequence the interest and the needs of industry in its relationship with ComSoc are also significantly different.

This President's Page, therefore, examines issues of ComSoc's industry relations and discusses what ComSoc has been doing and plans to do in the future to better promote industry relations and to provide unique value to industry. We strive to resume closely coupled ComSoc-industry relations, enabling us to achieve the two major goals of ComSoc in the new mobile convergence era that we are now experiencing.

COMSOC'S TEAMS FOR INDUSTRY RELATED MATTERS

ComSoc has four different teams that deal with industry related matters. Two are led by Directors, Marketing and Industry Relations, and Membership Programs Development; and two by ad hoc committee Chairs, Industry Promotion, and Industry Service.

First, the Director of Marketing and Industry Relations (MIR), Stan Moyer, deals with general industry matters in addition to performing the Society's marketing function. The Marketing and Industry Relation Board (MIRB) consists of Stan (Chair), Doug Zuckerman (Ex-Officio), Shri Goyal (liaison-MPDB), Tariq Durrani, Elena Neira, Bob Shapiro, Heiner Stuttgen, Tomohiko Taniguchi, Heather Yu, Angela Zhang, John Pape (staff) and Eric Levine (staff).

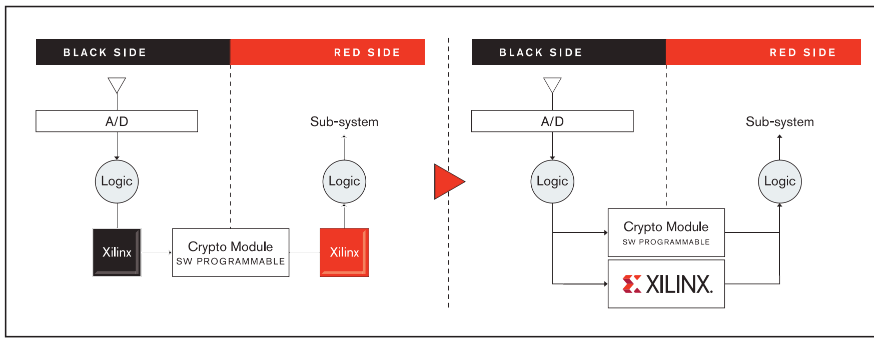
Second, the Director of Membership Programs Development (MPD), Shri Goyal, deals with membership programs that serve industry members in addition to other members, membership programs, and Chapter-related services. The Membership Programs Development Board (MPDB) consists of Shri (Chair), Stan Moyer (liaison-MIRB), Yigang Cai, Qian Zhang, Jingxian Wu, Ashutosh Dutta, Gim Wan Soon, Ashok Jagatia, Roberto Sarraco, Naoaki Yamanaka (Director-Asia Pacific), Tariq Durrani (Director-Europe, Middle East and Africa), Jose-Devid Cely (Director-Latin America), Gabe Jakobson (Director-North America), and John Pape (staff).

Third, the Ad Hoc Industry Promotion Committee (IPC) deals with industry promotion and operation of the Corporate Patron Program (CPP) and Industry Now Program (INP). The committee consists of Adam Drobot (Chair), Gong Ke (co-Chair), Stan Moyer (Director-MIR), Shri Goyal (Director-MPD), Madhu Pitke, Tomonori Aoyama, and a few more

(Continued on page 8)

Secure SWaP-C: Xilinx Extends Battery Life, Lowers Costs, AND Increases Security for MILCOM

By Colby Hoffman



Xilinx Single-chip Cryptography enables the highest level of information assurance with support for Type-1 systems, while optimizing SWaP-C.

THE NEXT DECADE of Military Communications demands secure communication with an ever-increasing sensitivity to size, weight, power, and cost (SWaP-C) concerns. Battery life will be expected to be days and not just hours, and more functionality must be built into increasingly smaller and less expensive hand-held devices. Xilinx defense-grade FPGA solutions help overcome these challenges, by enabling single-chip designs with improved system security *and* SWaP-C (Secure SWaP-C).

Military communications and other defense applications are not new to Xilinx. For the past 20+ years, Xilinx has advanced cost-effective solutions and design platforms tailored for this industry. The latest Xilinx® Spartan®-6Q FPGA family, for example, features long product lifecycles, high reliability, unique manufacturing flows, specialized design services, and advanced security solutions for high-assurance applications.

Xilinx has earned a clear lead in market share by delivering the highest level of integration of advanced capabilities such as single-chip cryptography (SCC) certifiable to Type-1 requirements, for accelerating development and optimizing Secure SWaP-C. Xilinx pioneered and first introduced the SCC methodology through collaborations with leading defense solution developers and key government agencies.

Today this Xilinx innovation is still the world's only single chip FPGA solution in production for Type-1 systems. Combined with its extensive development ecosystem, Xilinx SCC technology and security IP help shorten design cycles and reduce project risks. With Xilinx Spartan-6Q FPGAs, systems can accomplish what used to be unobtainable. These defense-grade Xilinx FPGAs are ideal for secure hand-held radios and other communications solutions. Unlike the alternatives, FPGAs offer a single-chip solution that enables reprogrammability with a click of a button. Devices become more cost-effective with the ability to support multiple protocols and formats, creating better user experiences even as communication requirements become increasingly sophisticated.

Besides offering unique Secure SWaP-C benefits, the Spartan-6Q FPGA family is part of a broad range of commercial, defense and space-grade devices. Off-the-shelf, ready-to-order Spartan-6Q FPGAs are rated to handle the operating temperatures of industrial (-40 to +100°C) and extended (-40 to +125°C) specifications. Standard lead content eliminates tin-whiskering concerns.

Time to market, reduced cost and risk mitigation are key to a programs overall success—making Xilinx defense-grade solutions the obvious choice. Developers can start out with Xilinx commercial-grade devices and later switch to 100% pin-compatible defense-grade devices with a seamless transition from prototyping to the low-rate initial production (LRIP) phase. And the inherent reprogrammability and functional flexibility of these FPGAs allow easier and faster design changes at any time.

Xilinx defense-grade solutions also include extensive development tools and support from industry-experienced operations and support teams. Aerospace and Defense is a top-tier market for Xilinx. Priority resources are dedicated to meeting the needs of next-generation applications. Come see Xilinx at MILCOM 2010, booth #1413 to learn more about Xilinx solutions for Military Communications or go to www.xilinx.com/defense.

About the Author: Colby Hoffman is the Senior Military Communications Architect, Xilinx Inc. (San Jose, Calif.). Contact him at more_info@xilinx.com

THE PRESIDENT'S PAGE

(Continued from page 6)

members to be named shortly.

Fourth, The Ad Hoc Industry Service Committee (ISC) deals with a diverse set of services focused on industry members and industry practitioners. The committee consists of Harvey Freeman (Chair), Fred Bauer, Mark Karol, Alex Gelman, Stan Moyer, and additional participants to be chosen from ComSoc membership, who have graduated in the last ten years and who are currently working in industry.

This President's Page is shared with the four volunteer leaders of the above four member-relations teams. Following this introduction, short biographies precede their thoughts and proposals on Industry as a key ComSoc constituency.

Adam Drobot joined 2M in July 2010 as Managing Partner and CTO. He currently serves as CEO of WhiteSpaceDB, a 2M Company. He is also the Chairman (non-executive) of Cebatech. Previously, from 2002-2010, Adam was the President of Advanced Technology Solutions and CTO of Telcordia Technologies. In that position he managed the Applied Research and Government and Public Sector Business Units. From 1975-2002 he was with Science Applications International Corporation, and among other assignments served as the Senior VP for Technology and Manager of the Advanced Technology Group He has published over 100 journal articles and holds 19 patents. Dr. Drobot serves on the boards of: the Telecommunications Industry Association, the American Occupational Therapy Foundation, OpenTechWorks Inc., and Advanced Green Computing Machines Inc. He is also on the Advisory Boards for the University of Michigan Transportation Research Institute, and the DOT Intelligent Transportation System. Adam is the 2007 recipient of the IEEE Managerial Excellence Award and the 2009 Chairman's Award for Communications Quality and Reliability. He is a Senior Member of the IEEE. Adam was granted a BA in Engineering Physics from Cornell University in 1968 and a Ph.D. in Plasma Physics from the University of Texas at Austin in 1975.

Shri Goyal received his Ph.D. in Electrical Engineering from North Carolina State University. Since 2002 he has been the Dean, College of Technology & Management at St. Petersburg College, Florida. Previously he was Director, Technology Programs in Operations Systems Technologies of Verizon Laboratories, where he developed and deployed intelligent operation support systems. Shri has 30 plus years experience spanning R&D, management, operations and engineering of wire-line and wireless networks/services. Shri has organized numerous global symposia/conferences, including NOMS and IM, and launched EntNet@Supercomm. He has published extensively and was Guest Editor of several IEEE publications. As an IEEE Fellow, he has been active in many ComSoc professional and leadership roles. He was Chair of CNOM and Enterprise Networking Technical Committees (1994-98). Within ComSoc, he served on the Board of Governors as the Director-Meetings and Conferences 2004-5 and is currently serving (2006-2011) as Director-Membership Program Development. He was the recipient of the IEEE Communications Society Donald W. McClellan Meritorious Service Award in 2008. He holds two patents.

Stan Moyer is an executive director and strategic research program manager in the Applied Research area of Telcordia Technologies, where he has worked since 1990. Currently, he is leading a business development effort for end-user information privacy protection for mobile services. In the past he led research and business development activities related to digital content services and home networking. Stan has been a frequent speaker at events such as IEEE CCNC, the Internet Engineering Task Force, the Broadband Home conference, IEEE ICC, and other technical workshops. Prior to that, he worked on ATM switch hardware, broadband network architectures and protocols, middleware, Internet network and application security, Internet QoS, and voice over IP. He is currently President of the OSGi Alliance. He served as a member of the IEEE Technical Activities Board Finance Committee. Within ComSoc, he is currently serving as Director-Marketing and Industry Relations and Vice-Chair of the IEEE CCNC steering committee. Stan has a ME in Electrical Engineering from Stevens Institute of Technology and an MBA in Technology Management from the University of Phoenix.

Harvey Freeman is the President of HAF Consulting, Inc, a Minneapolis-based technology-consulting and project-management company. With over 35 years of experience in the communications field, he specializes in designing/managing innovative technology-based solutions. Harvey received his B.S.E.E. degree from the University of Pennsylvania and Ph.D. in electrical engineering from the University of Illinois. Harvey is a long-term, enthusiastic volunteer in ComSoc. He served as Chief Information Officer, Vice President of Membership Services and Vice President of Technical Activities. He currently serves as Treasurer, a job he filled for four years previously. Harvey made numerous significant contributions to ComSoc, including first as Editor-in-Chief/founder of *IEEE Network Magazine*, founder/first General Chair and current Standing Committee Chair of IEEE INFOCOM, and founder of IEEE SECON. He served on the Meetings & Conferences Board, Strategic Planning Committee, Awards Committee, Staff & Facilities Committee, and Nominations Committee. Currently he is the IEEE TAB representative to the IEEE's Educational Activities Board. He was honored with ComSoc's Meetings & Conferences Exemplary Service Award in 2001 and the Society's Donald W. McLellan Meritorious Service Award in 2004.

INDUSTRY RELATIONS ISSUES

Where Does the Global ICT Industry Stand?

Drastic changes in the communications marketplace started with the divestiture of AT&T in 1984 and were followed by the privatization of national communications providers around the world. This was succeeded by the privatization and competition in the voice telephone market, the convergence of voice and data services, and the subsequent competition between telephone and Internet companies. The 'first digital convergence war' occurred during the wireline voice-data convergence process, initiated by ISDN in the 1980s and followed by ATM technology in the 1990s, where circuit-mode-based telecom equipment providers lost market share to packet-mode-based equipment providers centered on the Internet. As a result, an abrupt downturn in the fortunes of switching companies and a steep rise of router companies began.

The circuit-mode based traditional communications industry has had an increasingly hard time as the Internet-based industry has dominated the mainstream wireline market. A turnaround occurred when the traditional operators switched gears and went into the wireless business, resulting in substantial growth and global expansion. This trend has persisted through the first three wireless generations. However, packet-mode based Internet technology has rapidly gained power and expanded to wireless access networks, first through WiFi and then WiMAX, strongly challenging the circuit-mode based access technology. As a result, Internet technology has

(Continued on page 10)

(Continued from page 8)

become pervasive in the 'second digital convergence war' between the conventional telecom-based Long-Term Evolution and Internet-based Mobile WiMAX technologies. Looking ahead to fourth generation (4G) mobile communications, all-IP based Internet technology will completely replace the long dominance of the circuit-mode technology, and we can expect an additional battle for dominance between traditional operators and new content providers.

More recently, as the ICT infrastructure matures and as our knowledge-based society stabilizes, a paradigm change is emerging in the information and communications fields. This change is firmly rooted in computer software technology and encompasses not only communications, broadcasting, and Internet, but also old media such as music, newspapers, books and magazines. It invites computer hardware, software, and Internet service companies to compete with the communications industry. It establishes PC-like platforms based on mobile devices and offers new and old media services and applications. It creates a new business model that puts application content in the center and de-emphasizes the role of communication services. It segregates content providers independent from communications operators and pushes communications operators out to the periphery, while maintaining users in the central position. This 'third digital convergence war' threatens the traditional communications industry in every aspect.

The ordeal of repeated transformation and intensified competition forced communications companies to concentrate on survival and to shed many of the functions that they carried. Communications service providers greatly reduced or totally ceased investing in research. Communications manufacturers concentrated on developing components and systems for deployment. They focused on 'tomorrow' and often disregarded fundamental research for 'the day after tomorrow.' As a consequence, few communications companies, whether operators or manufacturers, invested in the long-term fundamental research that would benefit society in the long run. In the same context, with some notable exceptions such as Telcordia, which developed an independent research business, the communications industry appears to have de-emphasized the investment in the long-term technical expertise of their employees and the support for involvement with colleagues through ComSoc activities.

In relation to ComSoc, the transformation of the ICT industry reveals the following important observations. First, while undergoing a period of transformation and competition, many communications companies, especially operators, seem to have lost touch with professional and learned communications societies such as ComSoc. Second, a main thrust of the communications industry, especially manufacturers, has shifted to new geographies where the presence and importance of ComSoc has not been well publicized. Examples of this are India, one of the largest markets for wireless and a source of software enterprises, and China, which is an increasingly important manufacturer and also represents one of the largest service provider markets. Third, ComSoc's fields of interest as well as its coverage of industry need to expand in accordance with the broadening of the ICT industry. These include newly established fields such as software, storage and computing platforms, user interfaces, applications, and content.

What is the Status of the ICT Industry's Participation in ComSoc?

There are many ways for the ICT industry to participate in ComSoc. First, ICT industry employees can participate in ComSoc's technical activities as individual members. They can participate in Technical Committee (TC) activities as a contributing member or a leader of a TC. They can write papers and articles and publish in ComSoc's journals and magazines, or present their papers in ComSoc conferences. They can participate more actively as a member of a publication's editorial committee or as an organizing member or Technical Program Committee (TPC) member of a conference. They can participate in ComSoc's standards activities or can serve as a ComSoc Distinguished Lecturer in their fields of expertise. Through participation, they can interact and network with other experts in their own fields of interests, further developing their technical capabilities. Such employee level participation in ComSoc is beneficial not only to the individual employee but also to the industry as well because the strengthened employee's technical competency itself enhances the competitiveness of the company.

Second, the ICT industry can sponsor ComSoc conferences and participate in exhibitions or other events. Also, industry can join a patron program or other ComSoc package programs. Such programs include multiple products and services that ComSoc has developed, including publicity of the company by different types of advertising, speaking opportunities at conferences for a company's top-level management, package memberships at affordable rates, participation in Webinar programs, and participation in certification and training programs. Such company-level participation returns high-level rewards to the company in terms of image building for all participants and especially for student members.

Industrial individual member-level participation in ComSoc activities can be measured by the number of ComSoc members from industry. Until recently, the largest fraction of ComSoc members was affiliated with private industry, while the portion of those in educational institutions was comparatively small. As a specific example, 59 percent of ComSoc members were in industry and 11 percent in academia (excluding students) in 1999. However, by 2009 industry membership had dropped to 45 percent while academic membership had increased to 19 percent. This is indicative of how much impact the ICT industry transformation, experienced during the past 10 years, has affected ComSoc's industry membership. The contribution to papers and presentations at conferences by members from industry has had an even faster fall off. At the same time, it testifies to how much the ICT industry has retrenched from involvement with the technical community and from networking with the technical community today.

The big shift of membership distribution from industry toward academia presents a big challenge to ComSoc. Though industry membership is still a major portion of ComSoc's constituency, the rate of decrease is substantial. The reasons for the decrease in industry membership in the past 10 years may be identified as: 1) the downturn in the telecomm industry and the economy, which means there are fewer people in industry to draw from; 2) a reduced emphasis on basic and fundamental research in industry, which means there is less incentive for industry to publish papers and present at conferences; 3) an increase in R&D funding for academic and government institutions, with an incentive to publish results, leading to greater participation by this constituency; and 4) less relevancy of ComSoc activities to the technical interests and carrier benefits of industrial members.

Apart from membership, the participation of industry members in our major conferences and publications has also been decreasing significantly, with the vacancies being filled

(Continued on page 12)

(Continued from page 10)

by academic attendees. This shift has the danger of making the content of our conferences, journals and magazines more academic oriented, which, if it actually happens, would possibly make ComSoc's technical activities even less attractive to industry members. Industry, due to its need for research, generates a "pull," and sets an important direction for research. We can fix the problems by inviting more industry members to the editorial boards of publications and the TPCs of conferences. As the ICT industry in now more interested in standards than before, we may further reinforce support for standards activities. We expect that our recently launched VP-Standards Activities will be a big help in this regard. Further, we need to pay attention to the needs and interests of the engineers in fast-growing countries such as China and India, by developing new programs, services, and products, specially designed to serve their needs.

What Industry Members Can Contribute to ComSoc?

Industry members can participate in technical, publication, conference and other activities as any other members in other sectors do in general. In those activities they can bring their technical findings in an industrial environment and thereby contribute to furthering technical knowledge. They can make unique contributions that the members in other sectors cannot make by participating in industry-specific activities, as:

- Actively be involved in ComSoc's governance by supporting staff members who want to participate and help support and promote their involvement.
- Take advisory roles in the formulation of ComSoc's range of programs, making sure that there is relevance for industry.
- Get involved in cooperative internship programs at industry affiliates for ComSoc's junior and senior membership.
- Host Distinguished Lecturer Tours and other ComSoc events at industry locations.
- Encourage participation of industry's senior leadership in ComSoc held events and activities.
- Publicize and expose important problems facing industry and participate in helping to energize the community in attacking the problems.
- Get involved in various recognition programs and publicize their value to industry.
- Make available 'real' information and data for important problems, without compromising competitiveness.
- Sponsor contests and events that inspire the best young people to enter the profession.

What ComSoc Can Offer to Industry Members?

In the past ComSoc's mainstream products and services revolved around conferences and publications, which primarily serve the more research-oriented professionals. Now that fewer people from industry are involved in research, ComSoc needs to provide other products and services to better target this audience. Recognizing such a situation, ComSoc developed the following initiatives to better serve industrial members:

- Established Wireless Communications Engineering Technologies (WCET) Certification program and training programs in support of WCET (already established).
- Improved the selection of education and training content available to members in the form of online tutorials and webinar broadcasts (offered free to members).
- Created the ComSoc VP-Standards Activities in order to foster the creation and adoption of communications related standards that are of interest to industry (being done).
- Established new package programs that encourage greater support for/from industry in the form of patronage/sponsorship and membership. Corporate Patron Program (CPP) and Industry Now Program (INP) are typical examples (being done).

What are the Issues in Industry's Participation in ComSoc?

There are many issues related to industry's participation in ComSoc's activities. To identify the specific issues we first need to determine various detailed statistics, for example, regional and country based industry membership distribution; industry members' participation in conferences, tutorials, and workshops; industry members' participation in ComSoc's governance; industry members' contribution to publications and so on.

Second, we need to better understand common practice in industry, for example, how much industry protects IP for reasons of competitiveness; how much industry limits information sharing; what is the trust level in information sharing among industry, government, and academic institutions; and what is the financial pressure on industry in supporting employees' participation in ComSoc activities.

Third, we need to examine the relevance of topics covered in ComSoc activities and publications with industry's interest; contrast of views of ComSoc and industry on professional development; participation of different sectors of industry (e.g., manufacturers vs. operators vs. content providers) in IEEE standards activities.

Fourth, we need to understand how industry perceives support and sponsorship of events; how industry's contributions are recognized by members or event participants; how ComSoc members recognize what is important from industry's viewpoint; how industry knows what's happening on the technology front; how beneficial the certification and training programs are to industry; how industry can grow future industry leaders, and so on.

We leave the four different sets of issues addressed above as the subjects for further study by ComSoc's industry relations teams.

EMPHASIS ON INDUSTRY PROMOTION

As a task force team to deal with industry related issues focusing on industry promotion, we organized an Ad hoc Industry Promotion Committee (IPC) in December 2009. The main tasks of the committee are two-fold: one is to define and promote the Corporate Patron Program (CPP); the other is to refine and promote the Industry Now Program (INP). The committee is led by Adam Drobot, the Chair, and Gong Ke, the co-Chair, and includes the two Directors, Stan Moyer and Shri Goyal, as well as Tomonori Aoyama, Madhu Pitke, and some yet to be filled spots.

Corporate Patron Program (CPP)

The Corporate Patron Program is a package of existing ComSoc products and services that can be specifically customized and bundled to meet the needs of each individual participant company. The program was developed to provide an opportunity for industry leaders to reach out to ComSoc's influential members through exposure across the ComSoc web site, publications, and conferences. This program enables companies to leverage ComSoc products and services through package discounts. Components of the package include *IEEE Communications Magazine* advertising, webinar sponsorship, conference patronage, discounts on wireless training and certification, on-line tutorial sponsorship, and more. Several levels

(Continued on page 14)

THE PRESIDENT'S PAGE

(Continued from page 12)

of opportunity exist, such as Platinum, Gold, Silver, and Bronze, enabling the participating company to tailor the level of products and services to meet specific needs.

The IEEE Communications Society Corporate Patron Program currently has five participant Patrons — Samsung, LG Electronics, Cisco, SK Telecom, and Korea Telecom — and we expect that participation will significantly increas in the future.

In this early stage of the CPP, it is most important to broadly publicize the program so that potential patrons receive the information. ComSoc officers are dividing the job of publicizing and contacting potential patrons. In support of this, Eric Levine, the ComSoc staff person in charge of the CPP, has generated an explanatory brochure (please visit http://www.comsoc.org/corporate-supporters). ComSoc members reading this *IEEE Communications Magazine* can also help publicize the CPP and identify contacts of the potential patrons. You may simply provide the contact information to Eric Levine at e.levine@comsoc.org, or Stan Moyer at s.moyer@ieee.org.

Industry Now Program (INP)

The Industry Now Program is designed to promote industry participation in ComSoc around the world by offering companies and their employees the opportunity to use the values that ComSoc creates by working with professionals around the world. It is specifically tailored for industry organizations in the fast developing regions of the world. The program offers the option of customizing packages to address both geographic and company-specific needs. Packages utilize IEEE and ComSoc resources to access the latest information on technology and trends through tutorials, Distinguished Lecture Tours (DLTs), conference participation, the Wireless Communication Engineering Technologies (WCET) certification program, and participation in technical and standard committees. The package also offers options for connectivity with world leading technology organizations and visibility worldwide. In addition, through our Sister and Related Society program ComSoc also offers networking and collaboration with other IEEE Societies, local/national societies in global regions, and professional organizations.

The INP packages were initially offered in India and are now being tailored for China. Many companies and organizations have shown interest in the INP. Through a two-way dialog several organizations have helped in fine tuning INP packages to suit their needs. We have offered flexibility to create membership packages for even 25 members for the smaller and upcoming companies in India. Industry Days, a one day event in partnership with the IEEE Global Business Development organization, are being planned in India (February 2011) and China (later in 2011) wherein communication industries will participate and share information. In contrast with the CPP, the main feature of the INP is ComSoc membership for employees. Through this membership, the goal is for employees of these companies to participate in ComSoc activities for self development, enhancing the company's visibility, and adding value to the overall ComSoc community. For an introduction of potential INP participants or for more information on the INP, please contact Shri Goyal at shrigoyal@Comsoc.org.

SERVING SPECIAL INDUSTRY NEEDS

The need for increased ComSoc services to our existing industry members and potential new members has been apparent for some time. Academia is comparatively well served by our numerous publications and conferences. Currently an equivalent focus on our industry community and its practitioners is just getting started. Over the past few years we have been heavily involved in wireless standards and certification. However, members and non-members alike have been asking for more and offering many suggestions.

As another task force team to deal with industry related issues focusing on the industry service side, we organized an Ad hoc Industry Services Committee (ISC) in September 2010. The main tasks of the committee are to identify service items to effectively satisfy the need of industry members and propose fast-track implementations for the most promising items among them. The ISC is chaired by Harvey Freeman and has, as members, Fred Bauer, Mark Karol, Alex Gelman, Stan Moyer, and a few more still to be filled.

Examples of the types of new services that we can offer to our industry members are the following:

Provide more networking advice, services, and information: Small to medium size companies need this more than ever. ComSoc can provide access to relevant technical information, opportunities for technical debates, sounding boards for new ideas, news and business information and the like.

Be an anchor for members throughout their careers: In the past many ComSoc members would work for a corporation like Bell Labs, AT&T, or IBM for their entire careers. Nowadays most will change employers many times as well as work for small companies and startups. ComSoc could be the anchor that these people keep along for their entire working life.

Filter useful information for the practical engineers: ComSoc is an excellent place to find out about the latest R&D ideas, but it is not most peoples' favorite place to get information about all of the other information needed in their jobs. For its practitioner members ComSoc could do the reading between the lines of industry newsletters, forums, associations, etc., and present them in a concise and reliable way.

Provide product information: Engineers learn a lot through attending vendor presentations, especially during RFP processes. In smaller companies it is much harder to gain access to such information. ComSoc could play a role here in making a large amount of filtered product information available to the people in these companies.

Provide business information: Much of the information coming from ComSoc is about research or 'how to do' something. A lot of work by practitioners is about telling people 'what to do.' Currently very little comes from ComSoc about the trends that are out there and how they are perceived by industry. ComSoc could provide access to information that helps in understanding the direction that industry is taking. Examples are providing economic statistics, marketing analyses, market penetration data, and the like.

Assist in job searches: ComSoc could offer information on trends in careers in the telecommunication industry to help our members find employment. We can invite employers seeking potential employees to have hospitality suites at our conferences to meet our interested attendees.

We are striving to serve the needs of our industry constituents. These potential programs will be evaluated, with the most effective ones implemented. For identification of new service needs or suggestions to offer, please contact Harvey Freeman at h.freeman@ieee.org.

MOVING FORWARD

As ComSoc moves forward we would like to see better linkages between and integration of the newer programs and

(Continued on page 16)

(Continued from page 14)

efforts with the 'traditional' ComSoc services and products such as conferences and publications. For example, holding standards education and training sessions at ComSoc conferences and/or creating new ComSoc conferences around new and emerging standards. Incorporating training and education for the WCET certification in the form of intensive tutorials and training sessions that could be held at conferences and even offered online is another example of this integration. Additionally, recognizing that it is difficult for members to travel to conferences, ComSoc has started offering recorded programs of selected conference sessions made available via the web, enabling industry members to realize the benefits of ComSoc conferences even if they are not able to attend in person.

ComSoc is a vibrant society, with the tradition of over half a century of excellence in serving the telecommunications community around the world. Its strength lies in the community formed by the thousands of volunteers, from industry and academia, forming a nice complimentary team. Industry participation is crucial in defining the values and creating the needs for innovation, thereby enabling ComSoc to achieve its goals through industry. We will continue to develop special programs to attract industry through the Corporate Patron Programs and the customized Industry Now Program packages for enrolling members from industry and serving their needs.

We will also start new services aimed at our industry practitioners that will help members in career development, to excel in their current jobs, and to look to ComSoc throughout their careers. Our partnerships with industry, continued constructive initiatives, and energetic leadership will propel us to our goals. The result will be a tighter link with industry and a reduction or elimination of the mismatch between what ComSoc offers and what industry needs. This closely linked industry relationship will provide a sound basis for the future growth of the Society.

To do so we would like to start with a study of the issues of the ICT industry and some of the major changes and directions that are emerging. Next we will address how ComSoc can serve its industrial partners, what works well and what does not. We will then examine what industry can do that would benefit both industry and the broad membership of ComSoc. Finally, we will take a look at how ComSoc can increase industry involvement through offerings that are valuable to industry and at the same time in synchronization with the values of the Society. None of these issues are simple but they are important and we hope that we can investigate them in as forthright a way as possible. If you, the readership, would like to comment, please do so as we would very much like to hear from you.

CONFERENCE PREVIEW

MILCOM 2010: "THE NEXT DECADE OF MILITARY COMMUNICATIONS"

Military communicators will converge on San Jose, California from October 31 to November 3 for MILCOM 2010. Now in its 29th year, MILCOM is one of the largest government/industry conferences in the world, and the premier technical communications, networking and information-sharing event of its kind. More than 5,000 people from 30 countries are expected to attend this annual gathering, which draws professionals from government, scientific, academic and engineering communities, contractors, allies, and top international educational institutions.

"New directions and innovations in space technology are key to strategic military success, and we are prepared. You will be too when you attend this important conference with a collection of world class thought leaders such as our excellent panel of former assistant secretaries of defense for command, control, communications and intelligence," said Rick Skinner, MILCOM 2010 executive committee co-chair.

The conference will include more than 400 technical papers in addition to tutorials, workshops and panel discussions. Presented by recognized authorities and technical experts, the classified (U.S. Secret and below) and unclassified sessions will cover topics including:

- Waveforms and Signal Processing
- Networking Protocols and Performance
- Cyber Security and Network Management
- Systems Perspectives

Some of the world's brightest technical experts and government/industry leaders that will speak at MILCOM include:

- LTG Dennis Via, USA, J6, The Joint Staff
- Mr. Steve Wozniak, Apple Computer, Inc. Co-Founder and Fusion-io Chief Scientist
- Mr. Russ Daniels, Chief Technology Officer, HP Enterprise Services, moderating the industry panel "Looking to the Future with Technology"
- Dr. Robert Hermann, Mr. Thomas Reed and Mr. John Stenbit, leading the "Command, Control, Communications and Intelligence Panel"

Panel sessions include:

- Post TCA – What's Next?
- Protecting and Defending Internet Traffic for the .Gov Network
- SCA Waveform Portability
- Unmanned Aerial Systems to Provide Enhanced C4ISR Capabilities to the Warfighter
- MILSATCOM Terminal
- Networking the Battlefield to Connect Every Warfighter
- Protecting DoD Network Infrastructure Today and Into the Next Decade
- Key Technologies that Enable Cognitive Capabilities in Radio Systems

In addition to the technical program, MILCOM 2010 will include exhibits from more than 200 of the world's leading providers of information technology, communications and defense technologies.

MILCOM 2010 will take place at the San Jose Convention Center. San Jose has a campus-like feel with the proximity of the hotels, restaurants, and attractions including museums and performing arts venues. All are within walking area, and the conference hotel, the Fairmont, is practically next door.

There are two social gatherings that will provide attendees with networking opportunities.

On Monday, November 1 the conference kicks off with an evening at the San Jose Tech Museum. Designed to inspire the innovator in everyone, this San Jose landmark features hands-on and interactive exhibits divided among themed galleries that offer visitors the unique, comprehensive Silicon Valley experience.

The annual Chairman's Banquet will be held Tuesday evening, November 2 in the Convention Center. Distinguished military representatives will join the much anticipated celebration to recognize this year's award recipients for outstanding achievement and technology advancement in the field of military communications. The evening will also feature video tributes and lively entertainment.

There are also recreational offerings unique to the region including golf at Cinnabar Hills Golf Club, Winchester Mystery House flashlight tour and specialized guest trips to San Francisco and Monterey/Carmel.

MILCOM is co-sponsored by the Armed Forces Communications and Electronics Association (AFCEA) and the Institute of Electrical and Electronics Engineers (IEEE) Communications Society. Lockheed Martin and The Aerospace Corporation are this year's corporate hosts. The conference's Department of Defense advisor is Air Force Space and Missile Systems Center.

Conference details and registration are available at http://www.milcom.org.

Some of the world's brightest technical experts and government/industry leaders speak at MILCOM

MILCOM attendees will dialogue with the experts in panel sessions on unmanned aerial systems, protecting internet traffic, DoD network infrastructure, and radio systems, and more.

History of Communications

Edited by Mischa Schwartz

Significant Contribution to the Development of Wireless Communication by Professor Alexander Popov

Orest Vendik

The physical and engineering foundations of wireless communications were laid by giants of science on the frontier between the 19th and 20th centuries:

James C. Maxwell (1831–1879)
in theoretical physics
Heinrich R. Hertz (1857–1894)
in experimental physics
Alexander S. Popov (1859–1906)
in applied physics
Guglielmo Marconi (1897–1937)
in engineering and business

This article explains the work of Alexander Popov (Fig. 1), who was professor of physics and for a short period the director of the Electrical Engineering Institute, which is now Saint Petersburg Electrotechnical University.

Figure 1. *Professor Alexander Stepanovich Popov (1859–1906).*

This university has links with the IEEE and has a particularly good relationship with the IEEE History Committee. Figure 2 shows the IEEE Milestone tablet, which was dedicated in 2005 in the Memorial Laboratory of Professor Popov at the university. The author of this article has investigated the life and scientific activity of Alexander Popov [1–5].

Alexander Popov was born in 1859 in the village Turinskiye Rudniki (Fig. 3) in the Ural Mountains. He became interested in natural sciences early in his youth. His father ensured that Alexander received a good education at the seminary in the city of Perm. He later studied physics at St. Petersburg University. After graduation in 1882 he began work as a laboratory assistant at the university.

Alexander Popov was a physicist and an electrical engineer with a wide range of interests [2–5]. His paper "Conditions of the Most Beneficial Operation of a Dynamo-Electrical Machine" was published in the authoritative Russian journal *Electrichestvo* ("Electricity") in 1883. He took part in a scientific expedition to Siberia for observation of the solar eclipse in 1887. He studied X-rays, constructing an X-ray tube and displaying the X-ray phenomenon before an open audience just after Roentgen's discovery in 1895. During the summer of 1896 he worked as an electrical engineer in charge of the power plant at the Annual Fair in Nizhnii Novgorod. He was interested not only in fundamental physics, but also in different kinds of practical applications of electricity.

In 1901 Alexander Popov was appointed professor of physics of the Electrical Engineering Institute and for a short period served as the director of the Electrical Engineering Institute.

Alexander Popov took some of the earliest steps in the development of radio communications, documented as follows:

1. On 7 May 1895 he demonstrated a Hertzian wave receiver at a meeting of the Russian Physical Chemical Society (RPCS) in St. Petersburg. In March 1896 Popov demonstrated his detector to an audience of scientists from the RPCS. Although there is no published evidence that he demonstrated transmission of intelligent signals, his full paper [7] in the 1896 volume of the *Journal of the Russian Physical-Chemical Society* (JRPCS) describes both controllable sources of electromagnetic radiation used in his experiments, such as "a large Hertzian vibrator," and the hope for future application of his apparatus in the transmission of signals over a large distance. It is not unreasonable to presume that he had already transmitted information, including at this demonstration, but was unable to say so in his paper because of military secrecy, as further described below. These papers were cited in Marconi's patent specification.

2. An English-language letter to the editor by Popov [8] was published in *Electrician*, a British magazine.

3. We will make some observations about the personal relations between Popov and Marconi.

4. There is documentation for the cooperative work of Popov with the Russian Navy. A practical application of Popov's wireless communication system was used in January 1900 to provide an urgent communication link to a Russian cruiser, which was punctured and ran aground in the Finnish Gulf [9]. The radio communication link was established by Popov and his assistant, Pyotr Rybkin. This incident has been commemorated, as described below.

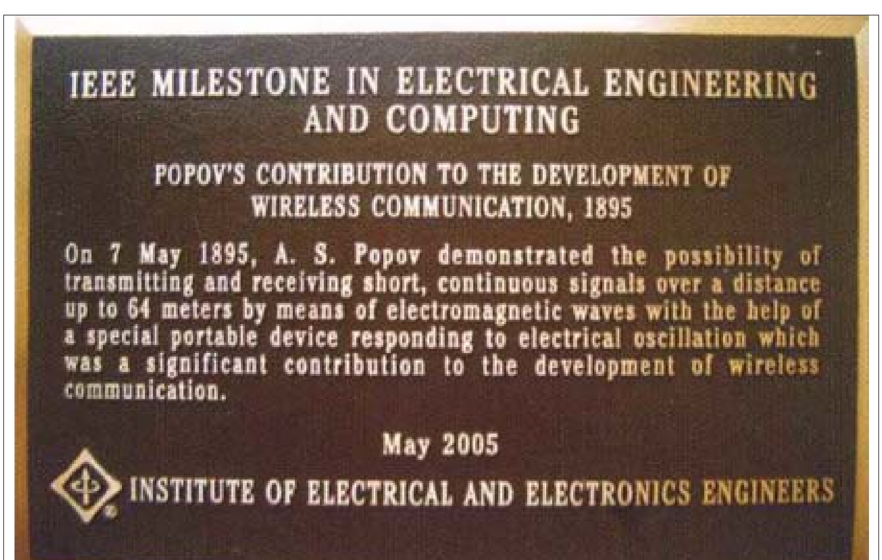


Figure 2. *The IEEE Milestone tablet, displayed in the Memorial Laboratory of Professor Popov at Saint Petersburg Electrotechnical University.*

(Continued on page 22)

0163-6804/10/$25.00 © 2010 IEEE

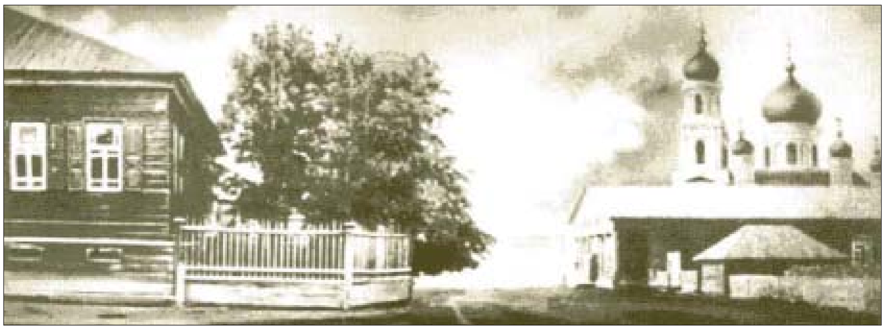


Figure 3. *The house in the village Turinskiye Rudniki where Alexander Popov spent his childhood.*

(Continued from page 20)

WHY THERE IS NO SPECIFIC MENTION OF THE TRANSMISSION OF INTELLIGENT SIGNALS IN POPOV'S PAPERS

The description of the presentation done by Popov in March 1896 in front of the audience of scientists from the RPCS was very brief and actually did not contain new scientific information. In order to explain this, one should take into account that Popov was on the staff of the Russian Navy Torpedo School in the Kronstadt naval base. The Navy Department of the Russian Empire would not permit an open publication describing transmission of intelligent signals. The officers of the Memorial Popov Museum at this university have found a document, signed by Alexander Popov in 1890, confirming that he had sworn to serve Emperor Alexander III and to maintain the secrecy of his State service [4, 5]. His service for the Russian Navy continued many years and was rather active.

ABOUT POPOV'S LETTER TO THE EDITOR OF *THE ELECTRICIAN* (DECEMBER 10, 1897)

Popov's letter begins as follows: *"Sir: The attention which you gave to the coherer in your issue of November 12 leads me to trust that you will consider my little work with this instrument described in the* Journal of the Russian Physical and Chemical Society, *Jan. 1896.* [10; O. J. Lodge, *Electrician*, vol. 40, Nov. 12, 1897, pp. 87–91]. Thus, we can understand that Popov's letter was inspired by Lodge's letter. Popov continues: *"The contents of my article were communicated to a meeting of the Physical Section of our Society in April, 1895 [6]. I translate, with abbreviation, some extracts of it."*

The most important paragraph of Popov's letter to the Editor [8] is the following:

"In conclusion, I can express my hope that my apparatus (when further perfected) will be applied for signaling on great distances by electric vibrations of high frequency, as soon as there will be invented a more powerful generator of such vibrations." This text is the exact translation of the conclusion of the paper [7] published in Russian, but the words in parentheses, which were in the Russian text, were omitted in the English-language version published in *The Electrician*. In *The Electrician* the figure showing the apparatus, which responded to electromagnetic waves, is slightly changed from the figure presented in [7]. The famous Professor Augusto Righi cited in his book [10] Popov's apparatus with the figure presented in [8].

About the omission in the English text submitted to *The Electrician*, the author of this paper would like to say that he does not know how experienced in the English language Popov was. However, in 1895–1897 Alexander Popov, as a professor of Kronstadt Torpedo School, was in good relations with Vladimir Skobeltsyn (Fig. 4), who was an assistant professor of the Physics Department of the Electrical Engineering Institute in St. Petersburg. Vladimir Skobeltsyn descended from an old aristocratic family and from early childhood had been instructed in the main European languages. He was well informed about the investigations of Popov. In April 1896 he delivered a lecture at the St. Petersburg Electrical Engineering Institute, *"Apparatus by A. S. Popov for Registration of Electrical Oscillations"* [11].

Figure 4. *Vladimir Vladimirovich Skobeltsyn (1863–1947). Assistant Professor of Physics Department of the Electrical Engineering Institute in St-Petersburg.*

Certainly, Skobeltsyn could have helped Popov to correct the English of the letter to the Editor [8]. One should take into account that the letter was written, translated into English, and posted from St. Petersburg to London in a very short time interval. This circumstance suggests that the words *"when further perfected"* were omitted in [8] because Popov's apparatus was already adequate for application in a signaling system.

ABOUT PERSONAL RELATIONS BETWEEN POPOV AND MARCONI

In many published materials the information about personal relations between Popov and Marconi goes back to the book *My Father Marconi* by Degna Marconi (McGraw-Hill, 1962).

There are two events described by Degna Marconi concerning Popov and Marconi:

1. The author writes about the visit of the Italian cruiser *Carlo Alberto* to Kronstadt, the Russian naval base, on July 12–21, 1902. On board the *Carlo Alberto* was Marconi's first floating wireless laboratory. On page 132 one reads: *"One day a Russian caller arrived at the foot of Carlo Alberto's gangway and said to an Italian sailor who helped him aboard: 'I want to pay my respect to Marconi, the father of wireless.' That caller was Alexander Stepanovich Popoff, the Russian scientist …"*

2 The author writes about Marconi's wedding in the spring of 1905. On the page169 one reads: *"Presents poured in: silver, jewels, linens and laces, plates enough for banquets, and glasses enough for routs. Popoff sent a sealskin coat and a silver samovar from Russia."*

The historical facts related to this are limited. It is known that from 1888 to 1901 Popov worked at the Kronstadt Torpedo School and lived in the town of Kronstadt. In 1901 he was elected Professor of Physics at the Electrical Engineering Institute in St. Petersburg, and

(Continued on page 24)

Figure 5. *Admiral of the Russian Navy Stepan Osipovich Makarov (1848–1904).*

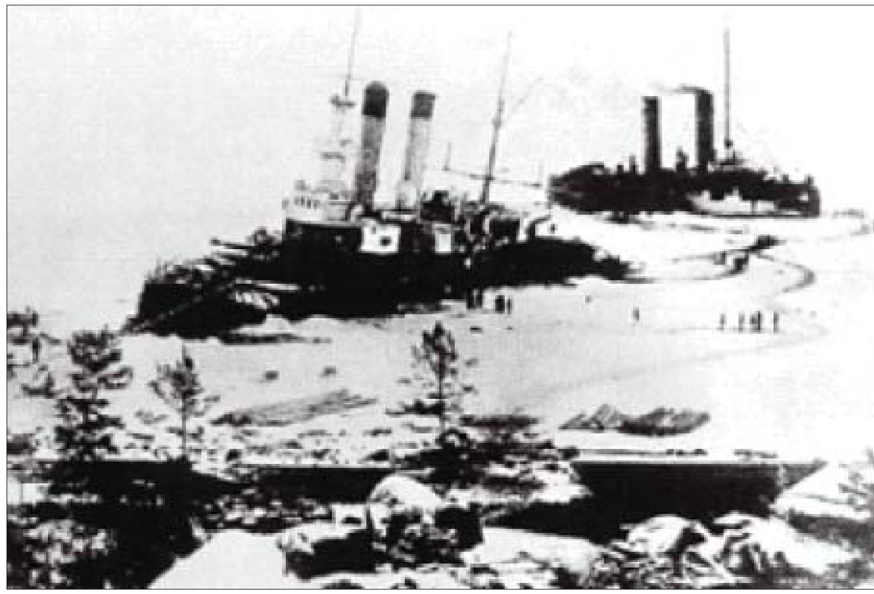


Figure 6. *Armored cruiser* Admiral Apraksin *and icebreaker* Ermak *in salvage operations near Gogland Island.*

(Continued from page 22)

on 25 September 1901 Popov moved to his new flat in St. Petersburg. There is no evidence that Popov visited Kronstadt in July 1902. Moreover, the Memorial Popov Museum at the Electrotechnical University is in possession of documents stating that the cruiser *Carlo Alberto* was visited by Admiral of the Russian Navy Stepan Osipovich Makarov (Fig. 5). The Admiral visited Marconi's laboratory on board the cruiser and had a conversation with Marconi about naval applications of wireless. One may suppose that somehow the two typical Russian names Makarov and Popov were confused. This may have resulted in the legend about that personal contact between Popov and Marconi.

At the same time it should be stressed that Professor Popov had a very good relationship with Admiral Makarov. The Admiral made active efforts to encourage development of wireless communications in the Russian Navy.

Being interested in the history of radio, the author of this article became acquainted with the historian Dr. Rolf Barrett who worked in Great Britain and delivered lectures on the history of radio. In 1996 Dr. Barrett wrote that he had a friendly correspondence with Degna Marconi (1908–1998). Degna Marconi was a lady of advanced years, but she was very active in correspondence with her friends.

The author of this article wrote to Dr. Barrett claiming that the question about personal contacts between Popov and Marconi is a very important point in the history of radio and asked him to use his correspondence with Degna Marconi to investigate whether the silver samovar is still in her possession, and what kind of settlement there was concerning the samovar. The letter from Dr. Barrett was received on 21 August 1996. The letter contained the reply of Degna Marconi to the question about the samovar [3]. She wrote: *"The Popov who gave the samovar and fur coat was not Popov, the scientist. It seems that this other Popov was a rich Russian industrialist, which my parents met, I am not sure where."*

In Degna Marconi's book, the paragraph about the wedding present to her father does not contain any words about Popov the scientist. The legend about friendship between Popov and Marconi seems to be an identification error.

Figure 7. *The map of Eastern part of the Finnish Gulf marking the position of the Island Gogland and the small Finnish town of Kotka.*

Commemoration of the Organization by A. Popov of the Radio Communication Link in the Finnish Gulf in January 1900

In Fig. 6 one can see the armored cruiser *Admiral Apraksin*, which was punctured and ran aground in the Finnish Gulf. The icebreaker *Ermak* took part in salvage operations near Gogland Island. The salvage operations were complicated by the absence of communication between the cruiser in distress and the command center. In this case wireless communication became indispensable.

Figure 7 shows a piece of the map of the eastern part of the Finish Gulf. The map helps us imagine the position of Gogland Island and the small Finnish town of Kotka, which at that time had telephone and telegraph connection with St. Petersburg, the capital of the Russian Empire. A practical application of Popov's wireless communication system was realized in January 1900 to provide an urgent communication link between Gogland Island and the town of Kotka.

The radio communication link was established by Alexander Popov and his assistant Pyotr Rybkin over a distance of more than 30 miles. The line operated reliably for a few months under severe Russian winter conditions.

For this work Professor Popov was encouraged with a message of thanks from the Emperor of Russia Nikolai II and awarded a significant sum of money. Popov used the money for acquisition of

(Continued on page 26)

Figure 8. *Monument to one of the first practical applications of wireless communication.*

Figure 9. *Monument to Alexander Stepanovich Popov at the cemetery (photo 1972).*

(Continued from page 24)

an estate in the Tver' region.

In Fig. 8 one can see the monument to one of the first practical applications of wireless communication, placed in 1968 on Gogland Island in the Finish Gulf by a group of radio engineers from Leningrad.

CONCLUSION

In conclusion, it is worth citing opinions of recognized scientists expressed during the lifetime of Alexander Popov.

In 1903 Popov was a participant in the first International Conference on Wireless Telegraphy in Berlin. In the opening speech at the conference, the State Secretary of the Ministry of the Post of Germany, Mr. Kraetke [12],said:

"In 1895 Popoff, in connection with his research intended to investigate perturbation of atmospheric electricity, suggested to produce telegraph signals by means of Hertzian waves: it is to him we are greatly indebted for the first apparatus recording radio signals. Marconi, with the first application of antenna as part of a transmitter, opened a new way for practical use of wireless communications. At the same time a number of highly appreciated inventors did their best to improve the new means of communications. The names of Braun, Ducretet, De Forest, Fessenden, Righi, Slaby, Arco, and Tesla are world-wide known. It is impossible to list all the names completely."

American scientist C. H. Sewall wrote in his book *Wireless Telegraphy* [13], published in 1903: *"A. Popov discovered new properties of the coherer, which opened up new possibilities of using the coherer or combination of steel needles and carbon plates as a detector in the front end of a radio receiver. Such a receiver was used in May 1900 to receive radio messages with head phones."*

It should be noted that the use of a *"combination of steel needles and carbon plates as a detector"* describes a nonlinear element. It was the first step in the study of nonlinear electric circuits under a high frequency current, one of the starting points of modern radio physics.

The year 1905 was the year of the First Russian Revolution. The St. Petersburg Institute of Electrical Engineering could not stand aside from political events of that time. In December 1905 Professor Popov was ordered by the Governor of St. Petersburg to take repressive measures against student political disturbances. He refused, and these events severely affected his health. He died soon afterward [14].

In Fig. 9 one can see the monument to Alexander Stepanovich Popov at the cemetery. At right is his daughter Ekaterina K'yandskaya (Popova) (1899–1976). She served as the director of the Memorial Museum of A. S. Popov in the St. Petersburg Electrical Engineering Institute from the foundation of the Museum in 1948.

ACKNOWLEDGEMENT

The author is grateful to Prof. Mischa Schwartz (Columbia University) for his interest in this topic and for fruitful discussions.

REFERENCES

[1] O. G. Vendik, "Popov, Marconi, Radio," *Nature*, vol. 374, 1995, p. 672.
[2] O. G. Vendik, "Contribution of Prof. Alexander S. Popov to the Development of Wireless Communications," *Proc. 25th Euro. Microwave Conf.*, Bologna, 1995, vol. 2, NEXUS, pp. 895–902.
[3] O. G. Vendik, "Did Professor Popov Send a Samovar as a Wedding Gift to G. Marconi?," *POLHEM Tidskrift för Teknikhistoria*, Gothenburg, 1997, vol. 15, pp. 207–09.
[4] O. G. Vendik, "Professor Alexander S. Popov: First Steps in History of Wireless Communcations," *Proc. St. Petersburg IEEE Chapter*, vol. 1, 2005, pp. 10–13.
[5] O. G. Vendik, "Additional Findings Concerning Activity of Professor Popov in the Development of Wireless Communication," *EUROCON '09*, devoted to the 150th anniversary of Alexander S. Popov, May 18–23, 2009, pp. 1096–99.
[6] A. S. Popov, "On the Relation of Metallic Powders to Electrical Oscillations" (in Russian), *Zh. Russ. Fiz.-Khim. Obshchestva (Physics, Pt. I)*, vol. 27, 1895, pp. 259–60.
[7] A. S. Popov, "Apparatus for the Detection and Recording of Electrical Oscillations" (in Russian), *Zh. Russ. Fiz.-Khim. Obshchestva (Physics, Pt. I)*, vol. 28, 1896, pp. I–14.
[8] A. S. Popov, "To the Editor," *The Electrician*, vol. 40, Dec. 1897, p. 235.
[9] G. I. Golovin, "A. S. Popov, Inventor of Radio, Life and Work" (in Russian), *Gosudarstvennoe Izdatel'stvo Literatury*, Moscow, USSR, 1945.
[10] A. Righi and B. Dessau, *Die Telegraphy ohne Draht* (in German), F. Vieweg und Sohn, 1907.
[11] V. V. Skobeltsin, "Apparatus by A. S. Popov for Registration of Electrical Oscillations" (in Russian), *Pocht.-Telegraf. Zh.*, 1896, pp. 547–49.
[12] M. Kraetke, "Le Discours d'Ouverture" (in French), *Conférence Préliminaire Concernant la Télégraphie sans Fil*, Berlin, Aug. 3–14, 1903.
[13] C. H. Sewall, *Wireless Telegraphy*, Van Nostrand, 1903.
[14] *Dictionary of Scientific Biographies*, v. 11, C. C. Gillispic, Ed., Scribner, 1981, pp. 93–94.

COMMUNISM'S DEFEAT SPAWNS POLAND'S GREATEST IT ENTREPRENEUR

BY RICHARD STEVENSON

Janusz Filipiak was one of the pioneer volunteers in IEEE Communications Society from Central and Eastern Europe. Along with publishing in *IEEE Transactions on Communications*, *IEEE Journal on Selected Areas in Communications*, and *IEEE Communications Magazine*, he was one of the first from this region to attend Globecoms and ICCs. He also organized several successful conferences in the area of network and service management. His example attracted many new members to our Society, and he supported them in their activities.

The article about Janusz Filipiak, written by Richard Stevenson, within the Technology Leaders' Forum, differs somewhat from those published earlier in this column by presenting also the personal side of this outstanding entrepreneur, engineer, scientist, and educator, as well as by showing the political background of Poland in the years of transition from the communist-style to marked-oriented economy. I do believe that the readers of the *Communications Magazine* will find this article interesting and stimulating.

Andrzej Jajszczyk

Janusz Filipiak has masterminded Comarch progression from a spin-off of AGH University of Science and Technology to one of the biggest software houses in the world.

It was 1993. Shattered remains of the Berlin wall hung as framed mementos in the homes of a recently re-united Germany, and over the border in Poland millions of out-of-work citizens grabbed their newfound freedom, launching their own businesses.

Janusz Filipiak, then a professor of electrical engineering at AGH University, Krakow, joined the startup fray. But unlike most of these entrepreneurs whose initial forays into capitalism failed, Filipiak succeeded. Today, Comarch, the firm he founded to build a database for Poland's leading telecommunication company, is huge, generating sales of over $200 million a year and employing 3500 people in 14 countries.

"I think we're bigger than HP or IBM if you ignore sales of computer equipment," says Filipiak.

And in his homeland, he certainly has the prominence of a Bill Hewlett or Dave Packard. His 42 percent stake in Comarch put him on Forbes' list of the hundred richest men in Poland, and his name regularly crops up in the business section of national newspapers. Turn over a few pages and you might also see him mentioned in connection with his family's luxurious nine-room restaurant that's welcomed celebrities as famous as Kate Moss, Steven Spielberg and Robert De Niro. And he also appears in the sport pages: he's president of the nation's oldest soccer club, Cracovia, Krakow.

Fame, of course, is a mixed blessing. Step out of line and everyone hears about it, like the time Filipiak was controversially arrested following accusations of withholding wages. But at least his public apology from Minister for Justice made the headlines too.

These days Filipiak takes this level of fame in his stride. But it wasn't something he anticipated as a child growing up in Bydgoszcz, a city in the north of the country. Back then he dreamt of being an academic.

He reasoned that if he could make his dream come true, he could bask in the joys of learning in the sole haven of meritocracy. But if he failed, his career options would depend on his connections, and his father, a middle level manager, and his mother, a primary school teacher, would not be able to offer much help.

The majority of Comarch's staff are based in a light, airy complex built on the outskirts of Krakow. These buildings are bedecked with paintings from one of Poland's finest modern artists, Rafal Olbinski.

Filipiak's parents did their best to equip their offspring with strong, well-rounded educations, forking out for private English and music lessons. "My parents forced me to play an accordion," says Filipiak. "The piano was too expensive for them."

But neither language nor music enraptured the child. Then, when he was in his mid-teens, an inspirational teacher turned him on to theoretical physics. He turned to the city's public library to feed this new passion. Russian texts dominated the physics section, but that didn't matter, because everyone had to learn that language. Soon, Filipiak's knowledge of physics surpassed that of his teachers, and he began giving local schoolteachers presentations on theoretical physics.

At age 19 and in his last year of schooling, Filipiak applied to study physics at one of Poland's premier universities, AGH University of Science and Technology in Krakow. He chose this five-year, masters course because it matched his interests: constructing theoretical models of real life.

But he didn't get accepted. He aced the entrance exam, but when he got there the Dean of the Faculty of Engineering told him that physics department was already full, and he should study electrical engineering instead.

"It was a tragedy for me," says Filipiak. "I blamed the whole world that it was unfair."

Filipiak had very few options at his disposal. Getting onto another physics course at a different university would take another year, and he couldn't stomach the wait. So he complied with the Dean's wishes.

He was pleasantly surprised. "I noticed that I could construct mathematical models in the electrical engineering area." Within six months his engineering teachers, spotting his prowess in the parts of the course that touched on physics, encouraged him to switch subjects. But engineering had hooked him.

(Continued on page 30)

(Continued from page 28)

Looking back, it was the right choice. He still cherishes theoretical physics, but feels it has made little headway in the last 30 years. "There are no new facts or knowledge to push theoretical physics forward."

In his fourth year, Filipiak took a trip that would change his outlook on life forever. By winning a competition for a three-month internship in Japan, he not only got the chance to get his hands on state of the art IT technology at Nippon Steel Works, but the opportunity to see a completely different way of life. He loved living and working in a different country, and returned to Poland eager to repeat that experience.

In 1976 he finished his electrical engineering degree, won best graduate of the year at AGH University, and started a PhD there, a highly theoretical research project looking into ways of improving the efficiency of traffic flow through telecommunication networks.

But his life was not just filled with complex equations describing traffic flow and the computer code that represented them. He had fallen for Elizabeth Staszak, a materials engineering student at the university. They married in 1977, and had the first of their three children, Anna, the year after.

In 1979, PhD in hand, Filipiak became an associate professor, fullfilling his dream.

He taught and brought in research contracts. His first, for Huta Katowice, a company owning a steel processing plant in south-west Poland, provided access to cutting–edge French and US technology. Built to showcase modern Poland, this plant housed some of the nation's fastest computers, and Filipiak exploited them to optimize the flow of steel through the mills.

For another contract, he designed a database for recording every citizen's personal details, such as their name, and address. Filipiak's role was purely theoretical; he never set eyes on the personal data itself. Poland's current national database incorporates his work.

Filipiak also landed a contract for improving coal miner safety. By tracking every miner's location, it is possible to shut the doors at the ends of empty corridors, stemming the flow of poisonous gases. "The main idea of the project, which Janusz came up with, was to put a sensor emitting a code on each miner," recollects Zdzislaw Papir, a co-worker on the project who is now an electrical engineering professor at AGH. The sensor network aced field trials involving a score of miners running in different directions. But today Filipiak and his colleagues on the project do not know if coal mines ever implemented their technology.

In the early 1980s, Filipiak moonlighted as a security guard. It's not that his academic career didn't keep him busy, or pay him a decent salary. It's because his young family desperately needed more living space. One way to increase the chances of getting an apartment was to help out some way in its construction, so Filipiak agreed to keep an eye on the building site. His efforts paid off, and in 1985 his family moved out of their two-room dwelling in a student hostel into one of the brand new apartments.

The security guard job wasn't particularly demanding; Filipiak spent much of the time writing papers. He craved international recognition for his work, and knew this would only happen if he wrote papers in English, which was practically unheard of in Poland at that time. It didn't come easy; he would draft the paper in Polish, then scan through English papers in journals, find appropriate sentences, and swap them with his Polish ones. It worked well enough. He first had a paper accepted by the UK journal Large Scale Systems, and others followed. His confidence blossomed, and he then drafted a book detailing his take on optimizing data transfer rates in high-speed networks. One of his happiest memories is the day he received a letter from the large German publisher, Springer, saying that they would print it. In 1988 Modelling and control of dynamic flows in communication networks hit the academic book market.

Thanks to this growing international reputation, in 1984 Prosper Chemouil from France Telecom offered him a one-year contract to work in Paris.

Convincing the University to keep his job open was easy. But he still needed a passport, and refusing to bribe the officials in this department didn't go down well. They slowed down the application process, and he left for France two months late, on 1 October 1984.

The French government wasn't much easier on him than the Polish government had been. It labeled him a political refugee and barred his wife from visiting, fearing she would immigrate. However, thanks to good friends of his who personally invited her, Elizabeth made two brief trips to Paris, visiting her husband in the middle and near the end of his time there.

Filipiak says that there are two approaches to combating the stresses of work: drinking and exercise. He's backing the latter, and through his company he is sponsoring professional sport, both in Poland and overseas. Recently Comarch has started to support the German football club, TSV 1860 Munich, which is also known as "The Lions."

Still, Filipiak had a good time, reveling in the opportunity to direct all of his energy at his research, free from teaching duties and shifts as a security guard.

He returned to Poland in June, 1985, but headed out of the country again two years later for a three-year stint at the University of Adelaide, Australia, working on research contracts for the country's domestic and international telecom providers, Telecom Australia and OTC.

This time he took his wife and family with him. Nearly for good, as he was offered a professorship there.

"If I accept this chair we immigrate from Poland," he reasoned. "If I don't accept it, we have to go back."

He opted for the latter. Australia just wasn't home. "Australia is far away and very different," he says. "How can you get involved in cricket or Aussie rules?"

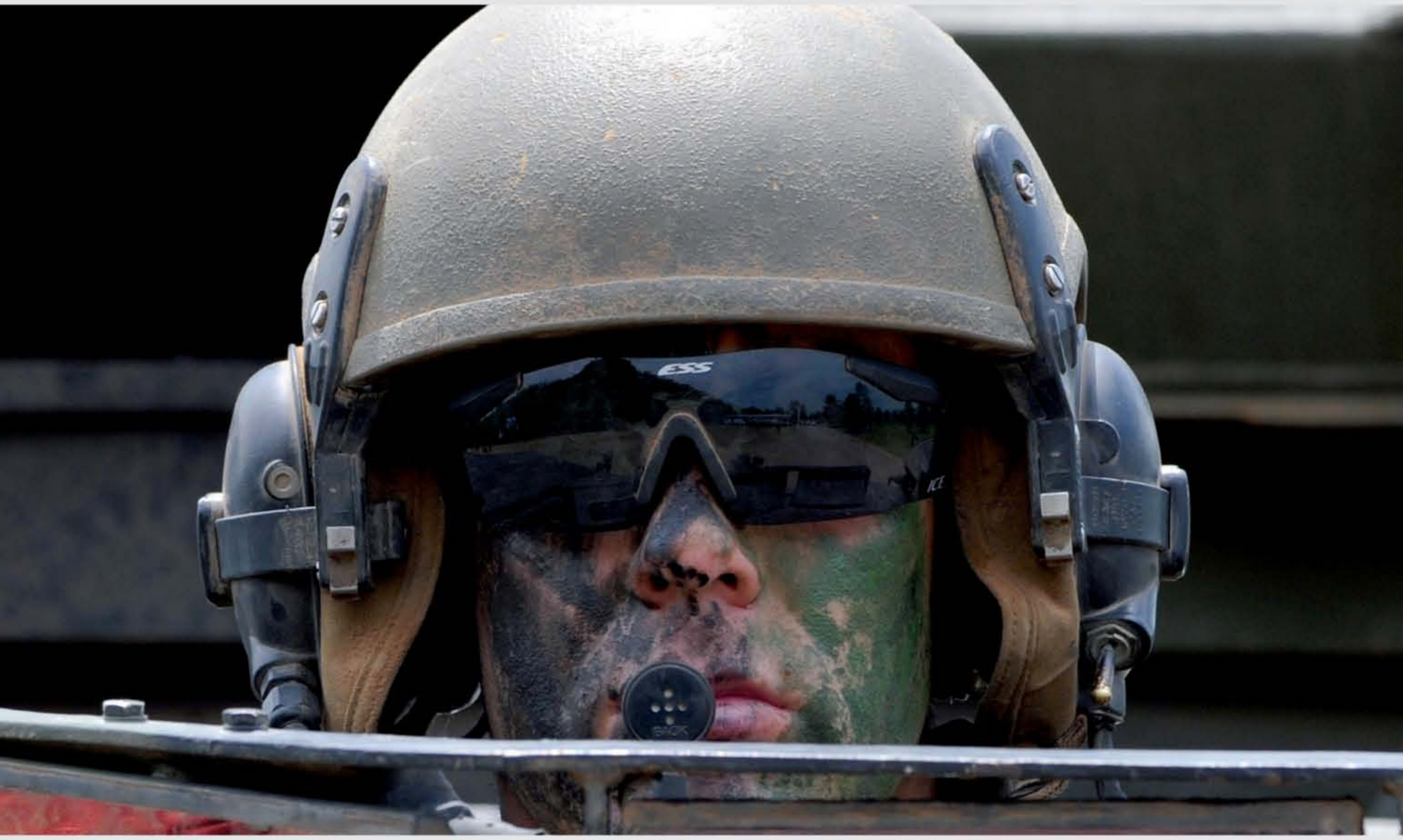
Back in Poland the political landscape was changing fast. Civil unrest in the spring of 1988 gave way to strikes in the summer, followed by talks in the Fall between the government and the independent trade union Solidarity. Finally, in April 1989, the voters in national elections rejected Communist candidates, and Poland started transitioning to a free market economy. By the time Filipiak and his family came home, just before Christmas 1989, the foundations for democracy were falling into place.

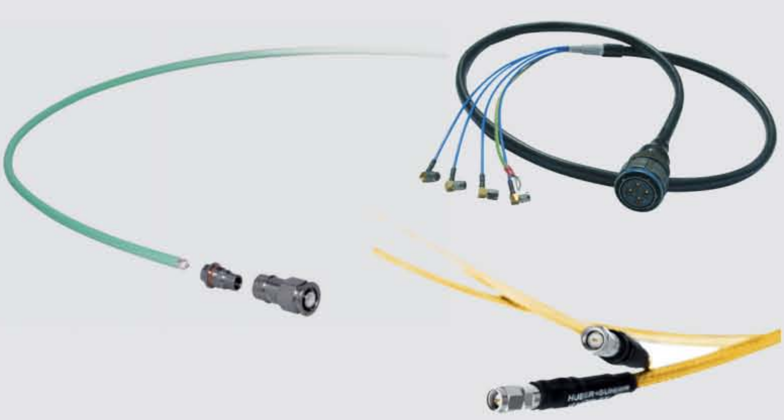
"In free Poland we could have a passport," he points out. "So I said to my wife, 'if something goes wrong we could always leave. We can always go to Australia or the US, because I could easily get a contract in one of those countries.'"

Back at the University, now a full

(Continued on page 32)

IEEE Communications | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | qMags





(Continued from page 30)

professor, a rare recognition in Poland, Filipiak launched a research contract sponsored by Telecom Poland to create a database inventorying its switches, hubs, transmission lines and other paraphernalia that made up its infrastructure, records then kept on bits of paper and random spreadsheets. Filipiak's team created the model for the database, and Telecom Poland tried to get a company to build it. But no one wanted the work—the database was just too big.

Filipiak wanted to try to build it. But to do so he'd need staff and resources. He thought he could make it work by establishing a research center within the university, something he had seen work successfully in the U.S.

But AGH University said no. "It was too innovative for them," says Filipiak. "Universities in Europe are very rigid, formal structures."

His wife had a solution: He should launch his own company. Poland, after all, had startup fever; Millions were already giving it a go, and Filipiak's contacts and knowledge gave him a great chance of success.

She won him over. But he didn't want to go it alone, so he joined forces with ten academic peers, launching a private company, Computer Communication Consultants. They leased a room in the engineering department, directly below Filipiak's offices.

"After several months I realized that I was the only one contributing to this company." Filipiak's peers lapped-up the kudos of co-founding a tech start-up, but were not going to let this venture distract them from their teaching and research.

This lack of commitment riled Filipiak. He bought out his colleagues and renamed the company Comarch, for computer architectures. He paid rent and wages for most of the first year, until the company got paid for building that database for Telecom Poland.

Filipiak's wife ran the company at first. "The [university] was my real life," says Filipiak, "and this was a shop on the side." Sticking with the university, it turned out, gave him an advantage—access to students, hungry for industrial experience and not put off by low wages.

Competitors failed to take Comarch seriously at first. "HP and IBM were laughing at me because I spoke about high-quality products meeting customer expectations," recollects Filipiak. It was the expectations part they didn't understand. Comarch's products weren't nearly as sophisticated as Western products, but Polish businesses didn't want cutting-edge software, preferring Comarch's simpler, easier-to-use offerings. "As Polish businesses became more sophisticated, we were in step with them, developing more sophisticated solutions," explains Filipiak.

Throughout the 1990s the company thrived, netting more contracts, ramping annual sales to $44 million, and building a talent base of former AGH University students. By 1998 head count reached several hundred and the company desperately needed a full-time leader. Filipiak resigned from the University.

Comarch grew up fast. In 1999 it launched on the Warsaw Stock Exchange, shedding its image of a company run by a bunch of students.

Today, Comarch's staff no longer cram into University offices; they now work in a custom-built North Krakow complex of light, airy buildings adorned in surrealist paintings from one of Poland's finest modern artists, Rafal Olbinski.

The dot-com bubble at the beginning of the decade took its toll on many Polish IT firms, including Comarch. Angel investors from the U.S. rode into town and began poaching local talent with offers of better wages and company cars. Filipiak could play that game. "I took a risk, bought one hundred cars and gave them to the best employees."

Staff retention issues vanished when the bubble burst. But Filipiak then had another headache to deal with: Comarch lost the Telecom Poland contract which, at the time, made up 40 percent of sales. He implemented a contingency plan that he had been tweaking since 1995—developing software for a variety of markets, without seeking contract support first. Products for banks and insurance companies came first, followed by software to help governments and utilities manage their inventories and bill customers.

Comarch also started to chase overseas contracts. Initially Filipiak traveled all over world trying to drum up business, but now he's put the scattergun away, focusing on a handful of neighboring countries, plus China.

When he's not out of the country, Filipiak begins work at home. He spends much of his time thinking about how his company runs, focusing on the challenges of optimizing the company's hierarchy of internal flow of information. His belief in the benefits of mathematical modeling remains, and he has even published papers showing that really good information flow is the key to making good decisions.

Just before noon he leaves home, climbs into his four-door Bentley Continental Flying Spur, and takes a short drive from the South of the city to his massive, luxurious office. He typically grabs lunch and then begins a series of afternoon meetings. Once they are over he heads back home, where he spends several hours answering emails and signing paperwork.

But Comarch's CEO is not all work and no play. "If you work under a lot of stress, the solution is to either to drink or do sports," says Filipiak. He wants more people to take the healthier option, and he's encouraging this by backing sport at all levels, from the building of a publicly available fitness center at the company's headquarters to sponsoring professional football and ice hockey.

Today, Filipiak is most famous for his football connections. He's the president of the football club Cracovia Krakow, and since Comarch's takeover of Germany's SoftM he's also involved with sponsorship of the German soccer team Munich Lions.

Some of Filipiak's football-related fame stems from his arrest two years ago, following a dispute over a player's wages—police charged Filipiak with abetting the back-dating of a former soccer player's employment contract and breaking employment law. In a very public scene at Balice Krakow airport on the afternoon of 12 April 2008, a team of 30 police officers escorted him from a plane coming in from Italy and handcuffed him. After questioning they let him out on bail the following morning.

The courts quickly vindicated Filipiak, but not before newspapers printed headlines like: "Comarch CEO charged over soccer contract fraud"; and "Janusz Filipiak spent the night in jail".

To this day, Filipiak does not know who was behind his run-in with the law. He claims to have few enemies, but like anyone in the public eye, there will always be somebody out to get him.

What don't people like about Filipiak? It might be his unflinching honesty, and an unwillingness to take the bribes that would make their lives easier. Or it could be his rebellious, anti-authority streak, evident years ago, when, as a student, he helped to build a snowman, which Poles consider to be a symbol of stupidity, outside the offices of the management of a leather shoe-making factory. Or it may simply be envy of Filipiak's tremendous wealth, or the fact that generally he thinks he knows best, a trait that can make him come across as slightly arrogant. But even if he's guilty of all these peccadilloes, they are still overshadowed by his absolutely tremendous achievements: An incredibly successful academic career, followed by an even more successful venture into the business world, when he has undoubtedly played a major role in putting Poland on the IT map of the world.

CERTIFICATION CORNER

THE 2011 CANDIDATE'S HANDBOOK

BY ROLF FRANTZ

A common theme that readers of this column may have recognized is the continuing effort to keep all aspects of the WCET program up to date – the Delineation, the exam, the bi-monthly e-newsletter, etc. Even the WCET website, www.ieee-wcet.org, recently was revamped to improve navigation, streamline the design, and add links to social media to enhance peer networking.

The *Candidate's Handbook* is another part of the program that continues to be updated. In fact, it could be argued that it has seen the most alterations, since the 2011 version, soon going to print, will be the third revision since the original 2008 issue. Admittedly, much of the updating in some versions has been minimal: correcting some typographical errors, ensuring that the correct application and testing dates are included, and similar modest changes. However, the 2011 *Handbook* has more than the usual number of changes, and some of them are significant.

IEEE Communications Society
Wireless Communication Engineering Technologies

As reported in previous columns, the Delineation of tasks and knowledge has recently been refreshed. Appendix A of the 2011 *Handbook* contains this updated Delineation, which forms the basis for the WCET certification exams in 2011 and beyond.

We have found that wireless professionals bring a wide variety of backgrounds, in both education and experience, to the exam. Consequently, the eligibility requirements are now recommendations. People with a bachelor's degree (or comparable education) and three or more years of professional experience have had greater success on the exam than those who are not as well prepared. However, individuals can and should assess for themselves whether they are ready to sit for the exam.

The online application will now allow you to apply for either of the next two exams. That is, since the application window has closed for the Fall 2010 exam, you will soon be able to apply for either the Spring 2011 exam or the Fall 2011 exam. This is more clearly spelled out in the *Handbook* and the table of application and testing dates.

There are more materials available now to help candidates prepare for the exam than was the case even just a year or two ago. The "Examination Preparation" section of the *Handbook* has been completely revised and now provides a detailed listing of the available resources, including ComSoc materials and how to locate commercial training courses.

The Glossary of common wireless terms, as well as constants, conversions, and equations, is available on line during the exam, and has been updated regularly to ensure that it is complete and consistent with the questions on the exam. The 2011 *Handbook* will include the most up-to-date version.

In addition to these changes, of course, the *Handbook* includes information specific to 2011: the dates for applying and taking the exam, the list of Prometric test sites at which the exam can be taken, updated contact information (where applicable), and similar items a candidate will need to know when applying for and taking the WCET certification exam in 2011.

The 2011 *Candidate's Handbook* will be an essential reference for anyone considering seeking WCET certification next year. Watch our website, www.ieee-wcet.org, for an announcement and a link to request your copy.

BOOK REVIEWS

EDITED BY ANDRZEJ JAJSZCZYK

ADVANCES IN NETWORK MANAGEMENT

BY JIANGUO DING, CRC PRESS (TAYLOR & FRANCIS GROUP), BOCA RATON, 2010, ISBN 978-1-4200-6452-0, XXV + 364 PAGES, HARDCOVER

REVIEWER: KRZYSZTOF WAJDA

The author of this compendium-like book has aimed to take concise stock of all aspects, motivations, architectures, protocols, and techniques in network management developed over the last 30 years with some visionary consultant-like descriptions of emerging concepts and services.

The book is split into introductory and background Chapters 2 and 3, describing a wide range of network evolution aspects and management systems' components, and then Chapters 4, 5, and 6, complementing one another, and containing the main theories and techniques for network management (Chapter 4) followed by an interesting overview of specific aspects of management for emerging networks and services (Chapter 5), and a brief overview of autonomic and self-organizing networks (Chapter 6).

The idea of the book is appealing and straightforward.

Chapter 1, "Introduction," is very short and contains only remarks describing the motivation of the book, and finishes with a rather formal presentation of the content of each chapter. Using just a few sentences, the author suggests the main message of the book: a jump from network history, from the main theoretical and practical concepts in management, toward recently emerging solutions and environments: flexible, autonomic, virtualized, "mass market."

Chapter 2 is a vast repository of definitions, concepts, observed laws in computer science and telecommunications, and historical facts. Most of this information can be found in other sources and is quite well known, but such a compilation is aimed at showing how complex contemporary networking is. Section 2.2.2, containing facts from the history of computer networks, mainly focused on the Internet (e.g., packet networks, email, TCP/IP, IETF, RIPE, dial-up services, VoIP, blogs, Web 2.0, notes about many real networks and companies), illustrates well the rapid and impressive network evolution in the last 50 years. This chapter is reasonably complemented by the earlier Foreword II, giving data about complexity and costs of management systems vs. hardware and software.

Fundamental concepts and protocols of management are described in Chapter 3. They comprise general definitions of object, agent, manager, protocol, architecture, and relation. This is a starting point to introduce complex solutions such as TMN, SNMP, and CMIP, with adequately chosen "building blocks" or "tools": MIB, ASN1, RMON, and so on. There is also a functional cross section of the fault, configuration, accounting, performance, and security (FCAPS) management system, giving a task-oriented view. The content of Chapter 3 is not sufficient to learn the concepts in detail, but it gives their brief description and also presents mutual relations among the solutions.

Chapter 4, with a rather formal and classical title ("Theories and Techniques"), aims at giving a description of broadly chosen concepts, not always fully standardized and deployed. By this approach we are able to figure out how impressive and sometimes even controversial concepts are considered for use in monitoring and management of systems. The general target is to collect credible data about a system's status, then to analyze it and compute messages, sending them back to managed systems in order to get required results. The goal is rather straightforward, but the proposed concepts cover a very broad scope: from policy-based network management (PBMN) via expert systems, machine learning, neural networks, formal mathematical methods like decision trees or probabilistic approaches, to agent techniques, active networks, and bio-inspired solutions, just to name some examples.

While Chapters 2, 3, and 4 are devoted to somewhat classical concepts and solutions developed and successfully implemented over the last 30 years, Chapters 5 and 6 aim at giving a fresh but consistent view of novel, not always sufficiently matured concepts, network architectures, or applications.

Chapter 5 is probably the main contribution of this book to analysis of novel management systems. It gives examples of the most interesting or appealing solutions in networking and services emerging approximately in the last 10 years. The manner of description is clear and unified: first, description of the concept; second, complementary description of management specificity for each concept. The choice of networks and services is indeed important, starting from NGN, then wireless and WANET networks, optical and overlay networks, grid, storage and satellite networks, supplemented by cognitive networks. Finally, there is a neat and comprehensive view of an extremely wide and loosely defined future Internet concept.

Finally, Chapter 6 gives an overview, in a less strict manner, of emerging autonomic computing and self-management of networks. This description differs from the rest of the book since instead of details, it merely suggests observed tendencies and directions of evolution.

There are four appendices, completing the content of the book in a natural way. Appendix A lists organizations involved in management standardization. Appendix B points out RFCs related to the recent SNMPv3, while Appendix C lists a fundamental set of ITU-T TMN M3000 Series Recommendations. Appendix D lists IEEE 802 Working Groups, active and inactive. It would also be beneficial to include in a next Appendix the IETF WGs dealing with management issues. The book concludes with abbreviations, a glossary, a detailed index, and a bibliography.

This book is written for network engineers and students of telecommunications. Researchers can also find a unified description of protocols and ideas necessary for understanding modern management issues. Specialists in the domain should be interested in the visionary description of recently proposed concepts.

DEPLOYING QOS FOR IP NEXT GENERATION NETWORKS: THE DEFINITIVE GUIDE

VINOD JOSEPH AND BRETT CHAPMAN, MORGAN KAUFMAN, ISBN: 978-0-12-374461-6, HARDCOVER, 512 PAGES, 2009

REVIEWER: RAFAL STANKIEWICZ

Deploying QoS for Cisco IP Next Generation Networks: The Definitive Guide by two distinguished network architects from Cisco Systems, Vinod Joseph and Brett Chapman, is an outstanding book covering up-to-date and novel quality of service (QoS) concepts and solutions. Although the Cisco approach and proprietary solutions are emphasized, the book provides a deep understanding of the QoS and IP NGN framework. The authors first present those issues as meant by standards and recommendations by IETF, 3GPP, ITU-T, and IEEE. It is shown that the Cisco approach, even if it deviates from some standards or extends standard solutions, is derived from them.

The book covers architectural and implementation aspects of QoS deployment for a variety of mutlimedia services (e.g., voice, video, IPTV). It provides insight into QoS provisioning for various services in heterogeneous networks and interoperability issues of distinct networking techniques. It is also shown how network and service convergence can be achieved under the IP NGN framework. Design and implementation details for such widely deployed Cisco platforms as the CRS-1, 12000, 7600, and 7200 series routers are also provided.

Book Reviews

The plot of the book is very clear. The authors start by introducing QoS building blocks, then show how complex and complete QoS architectures may be built out of them. The book is full of detailed case studies, configuration examples, and real-life deployment descriptions.

Chapters 1 to 6 provide both theoretical and practical backgrounds for QoS provisioning issues. The description is rooted in standards and provides basic common knowledge on QoS architectures, frameworks, and solutions, but the Cisco specific approach is also introduced. Chapter 1 presents a definition of QoS. It also gives a historic snapshot on evolution of communication systems starting from public switched telephone networks (PSTNs), through X.25, frame relay, ATM, and others. Finally, the NGN framework is presented. The authors also introduce the Cisco NGN framework, which does not fully conform to ITU-T standards but is a set of flexible concepts enabling building convergent networks supporting automated provisioning of services with quality guarantees. This approach is developed and presented in details throughout the book. At the end of the first chapter the reader will also find brief summaries of all other chapters. Chapter 2 presents QoS supporting technologies including the IntServ and DiffServ architectures, and discusses basic building blocks of the IP QoS architecture including packet classification and marking, traffic shaping and policing, congestion avoidance and management, and others. The authors present a standard based vision on those issues referring to related IEEE standards and IETF RFCs. They also present a Cisco implementation and approach to QoS provisioning that is slightly different. The following chapter discusses QoS requirements of multimedia services, classifies the applications, and introduces the notion of classes of service (CoSs). The set of 12 CoSs defined by IETF is presented and followed by a Cisco-modified CoS model. Mapping options of those classes to smaller service provider sets of classes are proposed. Additionally, mapping of quality requirements to a CoS is dealt with. The authors also discuss related concepts (e.g., the service level agreement). The next chapter provides a deeper discussion on configuration options for customer equipment in IP NGN networks. Chapter 5 is devoted to QoS for mobile networks and cooperation with NGN. The history of mobile networks from the first to the third generation is briefly presented. 3GPP releases from 99th to 6th and their convergence with IP networks, including IPv6 and the MPLS architecture, are discussed. The chapter deals with the QoS concept and architecture in 3G mobile networks. The four 3G traffic classes are described and related to the NGN framework. Possible mapping of 3G classes to IP CoSs is given. The authors also present Cisco solutions for mobile networks including recommendations for QoS deployment therein. Chapter 6 provides an overview of metrics and measures for QoS control and conformance with SLAs.

Chapters 7 to 11 are strongly related to Cisco products and proprietary solutions. Numerous case studies and configuration examples provide broad and deep practical knowledge for readers. Chapter 7 overviews the architecture of four prominent Cisco router series. The chassis, line card options, and architecture of those routers platforms are presented. QoS implementation is also addressed. The following chapters are devoted to Cisco IOS and IOS-XR QoS implementation. Chapter 8 deals with MPLS-based virtual private networks, while Chapter 9 addresses implementation for carrier Ethernet and virtual leased-line services. Native IP Internet services and multicast services are dealt with in Chapters 10 and 11, respectively.

In Chapters 12 and 13 the authors revisit issues related to monitoring of QoS characteristics in existing networks and verification of conformance of the achieved service quality to that promised by SLA. In the former testbed-based verification of QoS performance in a Cisco IP NGN network is presented. The behavior of traffic classes in both congested and uncongested networks is validated. Chapter 13 discusses measurement of network performance in the context of the SLA framework. It describes tools and procedures for QoS performance monitoring and measurement.

The book concludes with a discussion on possible evolution of IP networks and future directions in QoS deployment (Chapter 14).

The book is primarily targeted at professionals: network designers and architects, administrators, engineers, business managers, and all people involved in designing and deploying QoS frameworks. It will also be useful for practitioners using network equipment provided by other vendors, since it provides a general view on QoS frameworks as well. The book is also a valuable guide for participants of Cisco specialist certification programs (CCNA, CCIP, CCIE) as well as students of telecommunications and networking. Also, course trainers will find it interesting.

Summarizing, it is a book that cannot be missed.

HANDBOOK OF PEER-TO-PEER NETWORKING

XUEMIN SHEN, HEATHER YU, JOHN BUFORD, MURSALIN AKON (EDITORS), SPRINGER SCIENCE+BUSINESS MEDIA, LLC, 2010, ISBN 978-0-387-09750-3, HARDCOVER, 1454 PAGES

REVIEWERS: PIOTR WYDRYCH AND BARTOSZ POLACZYK

Peer-to-peer networks have become very popular since the introduction of Napster in 1999. The book edited by X. Shen *et al.*, with contributions from 101 people, may be a good starting point for researchers who want to be quickly introduced to both basic and advanced topics in peer-to-peer networking. Except for the Preface, Acknowledgments, and List of contributors, it consists of 12 parts.

In the beginning, readers are introduced to the history, classification, and basics of peer-to-peer systems. The authors discuss the legal, sociological, economic, and political aspects of peer-to-peer networking. Different overlay topologies and different peer-to-peer network usages are presented.

Part II concentrates on unstructured architectures, where special attention is paid to the neighbor selection problem. First, it introduces a global overview and general classification of first-generation peer-to-peer networks. Then it presents topology creation algorithms and an optimal query-based search process.

Part III concentrates on dynamic hash tables (i.e., structured overlays). The authors present the design and behavior of different overlays, and describe how to add structure to an unstructured peer-to-peer network, how to model DHTs, and how to take into account the underlying topology.

The next part, entitled "Search and Query Processing," presents many proposals for time and resource optimization related to content searching in peer-to-peer systems. Most of them touch on unstructured networks, so prior to reading this part, familiarity with the material presented in Part II is recommended.

The following part, "Incentive Mechanisms," relies on game theory to analyze the players involved in peer-to-peer network behaviors. Attention is particularly devoted to multicast streaming and file sharing, where bandwidth trading can be an incentive for cooperation between peers.

Part VI deals with trust, reputation, anonymity, and privacy in peer-to-peer networks. A number of mathematical models of reputation are presented. The authors provide a survey of a few approaches to anonymity and private, in the sense of permission to join the overlay, networks.

Part VII investigates multicast services using peer-to-peer networks. The authors put stress on structured networks, although the gossip-based broadcast technique, devoted to unstructured overlays, is also explained.

"Multimedia Content Delivery" is dedicated to video stream dissemination over peer-to-peer networks. Both the ideas of live streaming (i.e., application-level multicast) and video-on-demand systems are presented and illustrated with examples.

Peer-to-peer and mobile ad hoc networks (MANETs) share common characteristics like a dynamic decentralized environment. Thus, keeping those networks strongly connected can be beneficial. Part IX contains an exhaustive description of solutions integrating peer-to-peer networks and MANETs together. It supplies information devoted to both non-hierarchical networks and those containing an information storage point.

Part X's title, "Fault Tolerance in P2P Networks," is a little misleading. Although the first chapter discusses methods for merging partitioned structured networks, others do not concern fault tolerance issues. The second chapter focuses on techniques for load balancing, and the last one presents an existing peer-to-peer model, the acyclic preference-based model, with a mathematical and statistical overview.

Part XI is full of experimental results. First, the behavior of free-riders (i.e., users who intend to download only and not to share) is investigated. Then an analysis of the traffic generated by peer-to-peer systems is done.

The last part comprises a set of articles about peer-to-peer computing, service- and content-based networks, and mobile collaboration overlays.

This publication is composed as a handbook, where the first chapter of each part usually gives a useful overview introducing the covered problem. Therefore, such chapters are a worthy source for a less advanced reader to get acquainted with the issue discussed. Subsequently, chapters with more sophisticated material, necessitating deeper knowledge, are presented.

As a weak side, some chapters are not selected suitably and do not compose a coherent part, so it could be tiresome for a potential reader. Moreover, the depth of insight is not uniformly spread through all aspects. Some of them are presented with details, while others are just mentioned briefly. The figure and graph composition of the book is inconsistent: each author used his/her own style.

To summarize, readers who are not familiar with overlays and wish to gain knowledge in a broad variety of aspects of this topic will benefit from the *Handbook of Peer-to-Peer Networking*. The book may also be suitable as a supplementary reference for an undergraduate course on peer-to-peer networking.

CONFERENCE CALENDAR

2010

OCTOBER

• *ICMWI 2010 - Int'l. Conference on Machine and Web Intelligence, 3-5 Oct.*
Algiers, Algeria.
http://www.icmwi2010.org/

• *CNC 2010 - Int'l. Conference on Advances in Communication, Networks, and Computing, 4-5 Oct.*
Calicut, Kerala, India.
http://cnc.engineersnetwork.org/

♦ **IEEE SmartGridComm 2010 - 1st IEEE Int'l. Conference on Smart Grid Communications, 4-6 Oct.**
Gaithersburg, MD.
http://www.ieee-smartgridcomm.org/

• *ICIN 2010 - 2010 14th Int'l. Conference on Intelligence in Next Generation Networks: Weaving Applications into the Network Fabric, 11-14 Oct.*
Berlin, Germany.
http://www.icin.biz/

• *IEEE ARRAY 2010 – 2010 Int'l. Symposium on Phased Array Systems & Technology, 12–15 Oct.*
Boston, MA.
http://array2010.org

♦ **ATC 2010 - 2010 Int'l. Conference on Advanced Technologies for Communications, 20-22 Oct.**
Ho Chi Minh City, Vietnam.
http://www.rev-conf.org/index.php

• *WD 2010 - IFIP Wireless Days 2010, 20-22 Oct.*
Venice, Italy.
http://www.wireless-days.org/

• *WCSP 2010 - 2nd Int'l. Conference on Wireless Communications and Signal Processing, 21-23 Oct.*
Suzhou, China.
http://www.ic-wcsp.org/

♦ **IEEE CCW 2010 - 2010 IEEE Annual Computer Communications Workshop, 25-27 Oct.**
Lake Arrowhead, CA.
http://www.ieee-ccw.org/

♦ **IEEE EDOC 2010 - 14th IEEE Int'l. EDOC Conference, 25-29 Oct.**
Vitória, Brazil.
http://www.ieee-edoc.org

• *CNSM 2010 - Conference on Network and Service Management, 25-29 Oct.*
Niagara, Canada.
http://www.cnsm2010.org/

♦ **MILCOM 2010 - Military Communications Conference, 31 Oct.-3 Nov.**
San Jose, CA.
http://www.milcom.org/index.asp

• *APCC 2010 - 16th Asia Pacific Conference on Communication, 31 Oct.-3 Nov.*
Auckland City, New Zealand.
http://apcc2010.aut.ac.nz/

NOVEMBER

• *IEEE RIVF 2010 - 2010 Int'l. Conference on Computing and Telecommunication Technologies, 1-4 Nov.*
Hanoi, Vietnam.
http://www.rivf.org/Index/Index.class

• *NETGAMES 2010 - 9th Annual Workshop on Network and Systems Support for Games, 16-17 Nov.*
Taipei, Taiwan.
http://netgames2010.ntpu.edu.tw/

• *IEEE ICCS 2010 - IEEE Int'l. Conference on Communication System, 17-20 Nov.*
Singapore.
http://iccs-2010.org/

DECEMBER

♦ **IEEE CAMAD 2010 - 2010 IEEE 15th Int'l. Workshop on Computer Aided Modeling and Design of Communication Links and Networks, 3-4 Dec.**
Miami, FL.
http://www.ieee.camad.org/

♦ **IEEE GLOBECOM 2010 - 6-10 Dec.**
Miami, FL.
http://www.ieee-globecom.org/2010

• *ITU-Kaleidoscope 2010 - ITU-T Kaleidoscope 2010: Beyond the Internet? Innovations for Future Networks and Services, 13-15 Dec.*
Lonavala, India.
http://itu-kaleidoscope.org/2010

♦ **IEEE VNC 2010 - IEEE Vehicular Networking Conference 2010, 13-15 Dec.**
Jersey City, NJ.
http://www.ieee-vnc.org/2010/

• *ICSPCS 2010 - 4th Int'l. Conference on Signal Processing and Communication Systems, 13-15 Dec.*
Gold Coast, Australia.
http://www.dspcswitsp.com/icspcs_2010/index.html

• *IEEE IMSAA 2010 - 4th Int'l. Conference on Internet Multimedia Systems Architecture and Applications, 15-17 Dec.*
Bangalore, India.
http://www.imsaa.org/

♦ **IEEE ANTS 2010 - Advanced Networks and Telecommunication Systems, 16-18 Dec.**
Mumbai, India.
http://www.ieee-ants.org/

• *HONET 2010 - High-Capacity Optical Networks and Enabling Technologies, 19-21 Dec.*
Cairo, Egypt.
http://honet-ict.org/

2011

JANUARY

• *COMSNETS 2011 - 3rd Int'l. Conference on Communication Systems and Networks, 4-8 Jan.*
Bangalore, India.
http://www.comsnets.org/

♦ **IEEE CCNC 2011 - IEEE Consumer Communications and Networking Conference, 9-12 Jan.**
Las Vegas, NV.
http://www.ieee-ccnc.org/

• *RWW 2011 - 2011 Radio and Wireless Week, 16-20 Jan.*
Phoenix, AZ.
http://www.mttwireless.org/

• *WONS 2011 - 8th Int'l. Conference on Wireless On-Demand Network Systems and Services, 26-28 Jan.*
Bardonecchia, Italy.
http://conferenze.dei.polimi.it/wons2011/

♦ Communications Society portfolio events are indicated with a diamond before the listing;
• Communications Society technically co-sponsored conferences are indicated with a bullet before the listing. Individuals with information about upcoming conferences, calls for papers, meeting announcements, and meeting reports should send this information to: IEEE Communications Society, 3 Park Avenue, 17th Floor, New York, NY 10016; e-mail: b.erlikh@comsoc.org; fax: +1-212-705-8996. Items submitted for publication will be included on a space-available basis.

CONFERENCE CALENDAR

• *NCC 2011 - 17th National Conference on Communications, 28-30 Jan.*
Bangalore, India.
http://www.ncc.org.in/ncc2011/index.html

FEBRUARY

• *NTMS 2011 - 4th Int'l. Conference on New Technologies, Mobility and Security, 7-10 Feb.*
Paris, France.
http://www.ntms-conf.org/innovative-projects.htm

• *ONDM 2011 - 15th Int'l. Conference on Optical Networking Design and Modeling, 8-10 Feb.*
Bologna, Italy.
http://www.ondm2011.unibo.it/

• *ICACT 2011 - 13th Int'l. Conference on Advanced Communication Technology, 13-16 Feb.*
Phoenix Park, Korea.
http://www.icact.org/

♦ **IEEE CogSIMA 2011 - IEEE Conference on Cognitive Methods in Situation Awareness and Decision Support, 22-24 Feb.**
Miami, FL.
http://www.ieee-cogsima.org

• *ISWPC 2011 - Int'l. Symposium on Wireless Pervasive Computing, 23-25 Feb.*
Hong Kong, China.
http://www.iswpc.org/2011/

MARCH

♦ **OFC/NFOEC 2011- Optical Fiber Communication Conference, 6-10 March**
Los Angeles, CA.
http://www.ofcnfoec.org/

♦ **IEEE WCNC 2011 - IEEE Wireless Communications and Networking Conference, 28-31 March**
Cancun, Mexico.
http://www.ieee-wcnc.org/

APRIL

♦ **IEEE ISPLC 2011 - 15th IEEE Int'l. Symposium on Power Line Communications and Its Applications, 3-6 April**
Udine, Italy.
http://www.ieee-isplc.org/

♦ **IEEE INFOCOM 2011 - IEEE Conference on Computer Communications, 10-15 April**
Shanghai, China.
http://www.ieee-infocom.org

• *WTS 2011 - Wireless Telecommunications Symposium 2011, 13-15 April*
New York, NY.
http://www.csupomona.edu/~wtsi/wts/index.htm

NEW PRODUCTS

FPGAs SLASH POWER CONSUMPTION BY 50 PERCENT

Xilinx, Inc.

Xilinx has introduced an FPGA series that slashes total power consumption by 50 percent and offers capacity of up to two million logic cells on a unified architecture that scales across low-cost to high-end lines. Xilinx 7 series FPGAs further extend the range of applications programmable logic can address by breaking new ground in solving customer challenges for lower power and cost without compromising on higher capacity and increased performance. Implemented on 28-nanometer (nm) process technology optimized to deliver low power and high performance, the new lines enable significant levels of productivity as skyrocketing development costs, complexity, and inflexibility of alternate ASIC and ASSP technology make FPGA platforms more relevant to an increasingly diverse community of designers.

The 28nm lines extend Xilinx's Targeted Design Platform strategy introduced with the company's 40nm Virtex-6 and 45nm Spartan-6 lines, now in volume production. The Targeted Design Platform strategy combines FPGAs, ISE Design Suite software tools and IP, development kits, and targeted reference designs to enable customers to leverage their existinf design investments and reduce their overall costs as they meet evolving market needs. In this new generation, Xilinx also takes a critical step in its work to dramatically expand the ecosystem of available IP abd designs that enable customers to focus on differentiation even as they transition to 28nm devices.

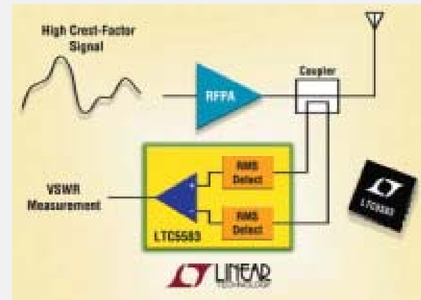
The new FPGAs enable developers to implement programmable solutions to a range of systems that had previously only been achievable in ASSPs or ASICs, including portable ultrasound equipment consuming less than 2 watts and low-cost LTE baseband and femtocell base stations.

All Xilinx 7 series FPGAs share a unified architecture that enables customers to easily scale their designs up or down in capability to reduce the cost and power or increase performance and capability, therebay reducing their investment in developing and deploying products across low-cost and high-performance lines. The architecture is derived from the Virtex-series-based architecture and has been designed to simplify reuse of current Virtex-6 and Spartan-6 FPGA designs.

http://www.xilinx.com

6GHz MATCHED DUAL RMS DETECTORS ENABLE ACCURATE VSWR MEASUREMENTS

Linear Technologies



Linear Technology introduces the LTC5583, a 40MHz to 6GHz dual channel, matched RMS power detector, offering over 55dB isolation at 2.14GHz. In RF power amplifier (PA) applications, the LTC5583 provides a simple solution for accurately measuring forward power, reverse power, and voltage standing wave ratio (VSWR). The device comprises a pair of 60dB dynamic range RMS detectors that are matched to 1.25dB. This provides accurate RF power measurement of high crest-factor signals such as those used in LTE, WiMAX, W-CDMA, TD-SCDMA and CDMA2000 3G or 4G basestations and other high-performance radios employing complex modulation waveforms. Each channel can detect signals accurately from as small as -58dBm to 2dBm, in a log-linear response with a typical linearity of better than +/-0.5dB covering all cellular frequency bands. At higher frequencies, the device is capable of providing 47dB of useful dynamic range up to 6GHz. Unique to the LTC5583, each detector simultaneously tracks the envelope of the modulated input waveform, providing on-chip capability to measure both peak and average signal power.

The LTC5583 has best-in-class channel-to-channel isolation of over 55dB at 2.14GHz, when driven differentially. Unlike other detectors, the LTC5583 can operate single-ended for RF input frequencies up to 2.14GHz, requiring no external balun transformers. This configuration greatly reduces costs without trading off dynamic range and provides isolation better than 40dB. An integrated amplifier measures the difference between the two detector outputs. For applications where one RF input is measuring the forward power and the other the reflected power, the difference output provides real time VSWR results. The two matched detectors are also useful in applications such as monitoring and controlling RF amplifier stage gain. The LTC5583's matching and isolation performance minimize calibration requirements, thus simplifying designs and reducing costs.

http://www.linear.com/5583

30 MHz FUNCTION/ARBITRARY WAVEFORM GENERATORS WITH UNPARALLELED SIGNAL ACCURACY

Agilent Technologies Inc.

Agilent Technologies Inc. has introduced the 33521A 1-channel function/arbitrary waveform generator and the 33522A 2-channel function/arbitrary waveform generator. The 33521A and 33522A are members of the Agilent 33500 Series line of function/arbitrary waveform generators.

The generators provide the lowest total harmonic distortion and jitter in their class. The generators, large graphical display allows for simultaneous parameter setup, signal viewing and editing. The generator also offers more precise functions and true point-by-point arbitrary waveforms that are alias-protected for exceptional accuracy.

With full-bandwidth pulses and real point-by-point arbitrary waveforms, the new 33500 series generators deliver the highest fidelity in their class. R&D and manufacturing engineers will experience a new level of accuracy and flexibility, offering more precise validation of their designs and devices during test.

The Agilent 33500 Series function/arbitrary waveform generators feature:

- 30 MHz sine, square and pulse waveform bandwidth to cover more applications.
- Less than 40 ps jitter and total harmonic distortion of .04 percent providing greater signal fidelity.
- 250 MSa/s, 16-bit sampling rate for higher time-resolution arbitrary waveforms.
- True point-by-point arbitrary waveforms for more accurate representation of user-defined signals;
- USB and LAN (LXI-C) for easy connectivity; and
- A TCXO timebase (standard) and OCXO (optional) for ultra-high stability.

The Agilent 33521A and 33522A function/arbitrary waveform generators are fully compliant with the LXI Class C Specification. The generators include USB 2.0 and 10/100 Base-T Ethernet (LAN) for quick and easy connectivity to a PC or a network.

http://www.agilent.com

Product Spotlights

u2t

The highly compact and integrated coherent receiver CPRV1010A is suited for 100G DP QPSK applications enabling compact module and line card designs. The receiver offers balanced detection and a linear electrical interface and supports a symbol rate of up to 32Gbaud. It complies to the CCRx-MSA specification.

http://www.u2t.de

Silicon Labs

Silicon Labs, complete portfolio of industry-leading XOs, VCXOs, clock generators, jitter attenuating clocks and clock buffers set a new standard for flexibility, performance and lead time. Low-jitter, high-performance, custom samples are available overnight; production quantities are available in 2-4 weeks. Start entering your requirements today!

http://www.silabs.com

Xilinx, Inc.

The Spartan®-6 FPGA Connectivity Kit is a complete, easy-to-use development and demonstration platform using the low-cost Spartan-6 LXT device family. This kit provides the hardware, software, IP and targeted reference designs needed to create high-speed serial systems and other connectivity applications right out of the box.

http://www.xilinx.com

GL Communications

GL's PacketExpert™ is a quad port (electrical & optical) tester that can perform independent Ethernet/IP testing at wirespeed. Bi-directional RFC 2544 testing, wirespeed BERT, frame capture, and other features are supported. It takes confusion out of Ethernet testing at all protocol layers, from raw Ethernet frames to IP/UDP packets. It can be used as a general purpose Ethernet performance analysis tool for 10, 100, and 1 Gbps Ethernet local area networks. Key performance indicators include frame loss, Tx/Rx frame count, bit error count and rate, bandwidth, and more. Detailed reports in PDF format are generated for both BERT and RFC 2544. PacketExpert™ further supports IPV4, IPV6, smart loopback, user-defined loopback, configurable headers for MAC/IP/UDP, error insertion, and sequence number generation

http://www.gl.com

Anritsu Company

Anritsu Company introduces the MS272xC Spectrum Master series that provides the broadest frequency range ever available in a handheld spectrum analyzer. Providing frequency coverage up to 43 GHz in an instrument that weighs less than 8 lbs., the MS272xC series is also designed with an assortment of applications to test the RF physical layer, making it easier than ever for field technicians, monitoring agencies and engineers to monitor over-the-air signals, locate interferers, and detect hidden transmitters.

http://www.us.anritsu.com

Vectron

The OX-405 is a low-noise, small form factor oven-controlled crystal oscillator (OCXO) targeted specifically at military radar applications. Both airborne and ground based radars will benefit from the OX-405's small size and phase noise performance. The OX-405 is the smallest low noise (below -165 dBc/Hz) OCXO at 100 MHz.

http://www.vectron.com/products/ocxo/ox-405.htm

Global Communications Newsletter

October 2010

The New Broadband Era: A 100 Mb/s Universal Service in the USA

By Ana Vazquez Alejos, Muhammad Dawood, and Rafael Asorey Cacheda
New Mexico State University, USA; University of Vigo, Spain

To provide universal Internet access to U.S. citizens, the FCC presented an ambitious plan to Congress in March 2010 [1]. According to this plan, about 100 million American citizens without access to broadband service are to be provided 100 Mb/s connectivity by 2020. The main target of the FCC plan is to make this service affordable to rural and low income urban areas, thereby ensuring universal service in a way similar to the service provided by the European Union to its citizens. An intermediate goal of the FCC plan will be to provide 50 Mb/s connectivity to 100 million households by 2015. Provision of 1 Gb/s connectivity to institutions such as schools, hospitals, and government buildings is presented as a long-term goal.

From an economical point of view, the FCC, by launching this service, hopes to achieve the goal of affordability for U.S. homes by stimulating the competition in this vital sector of the U.S. economy. The development of a specified competition policy is one of the strongest aims of the FCC broadband plan. The plan will seek to encourage private innovation and investment. The policies and actions recommended in this plan fall into three categories: 1. fostering innovation and competition in networks; 2. devices and applications; 3. redirecting assets that government controls or influences in order to spur investment and inclusion, and optimizing the use of broadband to help achieve national priorities.

However, the FCC does not detail how the infrastructure and access to affordable connections will be provided and financed. It is estimated that $12–$25 billion may be needed to expand and modernize the present infrastructure, thereby necessitating a large public investment. This kind of investment may not be available at a time some Internet providers and operators are undertaking improvements of their own networks. To this end, one of the approaches under consideration is a public auction of spectrum to be reclaimed from television broadcasters. Hence, the FCC hopes to reach a mobile broadband spectrum from the available 50 MHz to 500 MHz in 10 years. It is likely, however, that without public funds, it will not be possible to accomplish these goals, although achieving them would doubtlessly place the United States on top of the most developed information-driven societies.

Another no less important point is apprehension about network neutrality being compromised by government regulations. The broadband plan seems to have a price in a critical faction. This somewhat expected action would be in line with the recent Declaration of Granada adopted by the EU to justify the right of a government to keep some kind of control on the Internet.

Among the technical benefits that can be derived from the accomplishment of the FCC plan, 25 times faster web access is noteworthy. Other minor benefits include multiple aspects of present and future living, such as teleteaching, teleworking, and telemedicine. The main collateral benefit is doubtlessly the creation of employment opportunities in a time of economic crisis that has badly affected the information technology (IT) sector as well as so many others. Additionally, the use of broadband will let every citizen track and manage their real-time energy consumption, thereby helping the United States to lead the clean energy economy worldwide.

Since there are large areas in the United States without any kind of Internet access, another important and interesting intent of the plan is to deploy the fastest and most extensive wireless networks of any nation in an attempt to lead the world in mobile innovation. The white spaces (WS) broadband plan that FCC approved in 2008 is again claimed to encourage societies, industry, and investors to make working devices and networks that make use of this broadband WS. We remember that WS is the name used for the VHF spectrum gaps released by the TV broadcasters after the analog switchoff. The 700 MHz portion was auctioned by the FCC in 2008 at an amount of $20 million per 1100 licenses.

Additionally, the plan incorporates two important points that make it largely novel and proactive with respect to similar plans offered by other international information societies, such as South Korea, Singapore, Germany, Finland, and Spain. Another of the many goals of the program is to guarantee a nationwide public safety wireless network for first-responder agents.

The present year may be a suitable moment to support the consideration of the broadband Internet as an indispensable service to an information-driven society. The plan is not free from criticism, however. The intended data rates of 100 Mb/s may not be enough and could hardly be considered broadband Internet in 2020 due to the fast evolution of Internet applications and services. As an example of the fast growing speed demand, we find that the IPTV business is expected to blossom and grow exponentially over the next few years. On the other hand, the FCC plan is described as too aggressive. However, service providing commercial companies such as Google Inc. have announced the deployment of an experimental 1 Gb/s Internet network in some communities to test its feasibility.

Many more questions are still open, and future decisions will define the success of this plan. Perhaps the present critical economic situation will strongly shape its adoption.

(Continued on Newsletter page 4)

Project iREACH: Informatics for Rural Empowerment and Community Health in Cambodia

By Brian Unger, University of Calgary, Canada, Chea Sok Huor, iREACH Project Manager, Cambodia, and Helena Grunfeld, Victoria University, Australia

The International Development Research Centre (IDRC) of Canada has funded an ambitious unique project designed to influence ICT policy in Cambodia. Two pilot sites were developed in poor rural areas, one in Kamchay Mear of Prey Veng Province and one in Damnak Chang Eur of Kep. Each pilot comprises both a cluster of 10 WiFi hotspots, called hubs, and a central office connected to the Internet via satellite and to the 10 hubs by a WiMAX network. Each pilot site is managed by a locally elected management committee with support from project staff.

Project staff at each pilot include a pilot manager, a multimedia development coordinator, several content developers who work directly with community members to create audio and video programming, a research coordinator, and an IT technical support person. The key staff at the central Phnom Penh office are a project manager, a research manager, and an IT technical support manager. The latter two have relevant university degrees and several years' experience. The two IT support people at the pilot sites have little formal IT education.

During the IT specification and implementation stages, computer science and engineering experts were available as consultants. Two fourth year geomatics undergraduate students from the Faculty of Land Management and Administration, Royal University of Agriculture, also performed a transmission "line of sight" study. The results of this study helped with the determination of antenna heights, locations, and other network topology issues.

The project was launched by the Cambodian Ministry of Commerce in May 2006 with initial funding of US$1.3 million from IDRC. With a small further commitment by IDRC, the project is funded through May 2011. Significant further support is needed to achieve sustainability.

iREACH benefits are being studied and documented, and

Young and old people participate to iREACH activities.

Cambodians videoconferencing at pilot sites between remote villages.

lessons are emerging around community capacity building and empowerment; technical challenges in rural environments; developing relevant and appropriate ICT services; creating a community-based enterprise; deploying a range of participatory monitoring and evaluation approaches; and working within a centralized and fluid political context.

iREACH is intended as a model of an ICT-based rural community initiative run by the community to serve its needs. The rural communities are envisioned as using ICTs to improve the their economic and social health.

Project Activities and Expertise

Over the years, iREACH staff and volunteers have developed unique expertise in several areas, including:

- Generating and sharing information through e-modules on health, agriculture, and education using Ihe internet, videoconferencing, and radio programming
- Providing ICT services for the community including email, Internet access, overseas calls, audio and videoconferencing, distance learning, and English language training
- Empowering the community by building skills, capacity, and confidence in ICT service development and use, and enabling community participation in all related activities including gender awareness and equality
- Conducting participatory research and surveys in rural areas, including appropriate training for local people and young volunteers

Initial Results on How Villagers Have Benefitted

A full report on participatory research conducted in 2009 and followed up again in February 2010 is nearly complete. Using a qualitative methodology, information was collected from villagers in 22 focus groups in 2009 and 19 groups in 2010.

This research identified many aspects of iREACH that have contributed positively to the livelihoods of users, and other villagers who have benefited through the knowledge passed on by users. Participants from diverse backgrounds demonstrated a capacity for internalizing the information they obtained and using the resulting knowledge in productive ways. Benefits of the project have also played an important

(Continued on page) 4

Third IEEE Lebanon Communications Workshop at American University of Beirut

Rima Cortbawi, Lebanon

The IEEE Communications Society Lebanon Chapter in collaboration with the IEEE American University of Beirut (AUB) Student Branch organized the Third IEEE Lebanon Communications Workshop 2009 (IEEE LCW '09) on 21 November at AUB's Hostler Auditorium. The workshop attracted around 230 participants including telecom engineers, professors, and students. The program started with a plenary session followed by two technical sessions, and for the first time this year, the proceedings were concluded by a panel discussion on the impact of electromagnetic radiation on health.

IEEE Communications Society Lebanon Chapter Chair Dr. Zaher Dawy (AUB) opened with a welcoming speech, thankful for the "strong participation" and "distinguished speakers from Lebanon and abroad." He pointed out that IEEE LCW '09 is instrumental in achieving IEEE ComSoc Lebanon Chapter objectives: to contribute to telecom advancement on a national level, to add to telecom awareness, and to strengthen ties between the academic and industrial worlds. He concluded by thanking the workshop's corporate supporters Detecon Consulting, National Instruments, Cisco, and Nokia Siemens Networks.

AUB's Dean of the Faculty of Engineering and Architecture (FEA) Ibrahim Hajj commended student interest in particular, adding that "students will drive the technology of the future." Hajj added that "countries are not born rich and powerful; they get that way through innovation and research," both of which require collaboration and team work. "This workshop is a small step in the direction of creating [the needed] atmosphere of cooperation," he concluded.

Chair of the IEEE Lebanon Section Dr. Elias Nassar (Notre Dame University) introduced the IEEE as the "largest professional organization in the world" and showed the division of memberships by numbered regions around the world. Nassar then outlined the technical activities of the IEEE, mentioning benefits to professionals from joining the Institute and adding that "staying technically current" through IEEE pays off in the professional world.

The plenary session included a presentation by the head of Competence Practice Communication Technology and member of the Executive Board at Detecon International Dr. Hans-Peter Petry, who elaborated on the next generation of mobile networks, discussing assessment methodologies as well as benchmark parameters. The plenary session was followed by two technical sessions with featured presentations by distinguished invited speakers from Nokia Siemens Networks (Dr. Jijun Luo, LTE System Product Manager), Intel (Mr. Cuneyt Yucel, Director of Market Development for META Region), Ericsson (Mr. Dory Chakour, Radio Access Network Director for Middle East), Cisco (Mr. Moustafa Kattan, Senior Consulting Systems Engineer), and the Lebanese Telecom Regulatory Authority (TRA).

The workshop concluded with an interactive panel discussion on the impact of electromagnetic radiation on health with specialist panelists representing both the public health and engineering points of view on the subject. The panelists included Dr. Sobhi Abou Chahine (Beirut Arab University), Dr. Rima Habib (AUB), Mr. Ibrahim Duhaini (Rafic Hariri University Hospital), Mr. Nayef Azar (Alfa managed by Orascom), and Mr. Hassan Dhaini (TRA).

For more information, check http://ewh.ieee.org/r8/lebanon/com/ieee_lcw.html

42nd World Telecommunications and Information Society Day 2010 in Pune and Mumbay

By Avinash Joshi, Tech Mahindra, India

WTISD 2010 in Pune

The IEEE ComSoc Bombay Chapter celebrated the 42nd WTISD 2010 with traditional grace, jointly with the Institute of Engineers (India) Pune Local Centre and also with the IETE Pune Centre. A press conference was jointly held on 14 May 2010, at Patrakar Bhavan to give a brief outline of the agenda for 17 May, and the Chairman of the IETE Pune Centre (Mr. Ravi Datar) spoke about this year's WTISD theme and the objectives of celebrating the day. The Secretary Institute of Engineers (India) Pune Local Centre, Mr. D. B. Dhone, and Chairman of the IEEE Comsoc Bombay Chapter (Dr Avinash Joshi) spoke about their respective organizations.

The venue of WTISD 2010 was the Firodia Hall of the Institute of Engineers (India) building in Shivaji Nagar Pune. The chief guest of the function was Shri V. K. Mahendra, principal general manager at BSNL Pune Telecom. The function started with a welcome address by Er. Shri A V Surve, Chairman, Institute of Engineers (India) Pune. Thereafter the Chairman of IETE Pune, Shri RV Datar, read out the message received from the ITU Secretary General on this year's theme, "Better City, Better Life with ICTs." He also spoke about the various activities of IETE Pune Centre. Shri TV Mahadevan, Honorable Secretary, IETE Pune Centre, introduced the eminent chief guest.

Dr. Bharat Chaudhari, Chairman, IEEE Pune Subsection, spoke on behalf of Dr. Avinash Joshi who was traveling and hence unable to participate in the function. Dr. Chaudhari apprised the audience of the IEEE and Communications Society activities, in particular.

The star attraction of the program was the keynote address by chief guest Shri VKMahendra. The audience was highly appreciative of the keynote address and listened with rapt attention. Shri Mahendra in his address first recalled the history of telecommunications as well as IT in the world since their inception and simultaneously brought out the corresponding developments in India recalling the manual electromechanical Strowger and crossbar, and then electronic and digital switching eras. He then came to the ITU theme and explained how modern ICT with its global connectivity through the Internet and digital pathways can serve society in various ways. He also brought out the scope ICT provides for introducing innovative services that can serve society in an inclusive manner. Shri Mahendra concluded his speech by outlining various BSNL Pune Telecom activities in the pipeline.

The occasion was utilized to present the annual ISF awards. Prof. SV Kharkar, Distinguished Fellow and founder member of IETE Pune Centre, declared the results, and the chief guest distributed the prizes.

The function was expertly coordinated by Shri D. G. Patil, committee member of the Institute of Engineers (India), despite short notice.

Er Shri D. B. Dhone Secretary, Institute of Engineers

(Continued on Newsletter page 4)

WTISD 2010/*continued from page 3*

(India) Pune Local Centre, proposed a vote of thanks. The function concluded with high tea. Wide press coverage of the event followed in the next day's newspapers.

WTISD 2010 in Mumbai

The IEEE Bombay Comsoc Chapter jointly organized the program with TCS-Nortel Lab Mumbai, one of the oldest units in TCS jointly working with Nortel for more than 15 years in code-division multiple access (CDMA) technology based development, support, and R&D activities. The topic "Trends in Communication" was aptly chosen by Dr. Madhukar Pitke, Founder-Director of the Center for Development of Telematics (CDOT), Fellow of IEEE Indian Academy of Sciences and IETE, who guided associates in opportunities and challenges in the telecom world and also shared some instances related to the ITU council theme "Better City, Better Life with ICTs." He shared the importance of peer recognition and inspired associates to write papers for IEEE journals where they can get high value and recognition for their work. He also focused on the telecom basics that need to be built by every person working in this domain. The highlight of this session was the interaction between the TCS telecom engineers and Dr.Pitke, who answered all the audience's technical questions in detail.

Mr. Ashok Jagatia, IEEE Bombay Section Secretary, addressed the audience, shared the benefits of WCET certification, and urged TCS associates to take up this certification examination to enhance their competencies. Helso shared details about the Industry Now program with TCS management.

Introduction of the speaker and a vote of thanks were given by Dnyaneshwar Kamble, IEEE Bombay ComSoc Secretary. At the end of this program high tea was served.

Global Communications Newsletter
www.comsoc.org/pubs/gcn

STEFANO BREGNI
Editor

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

IEEE COMMUNICATIONS SOCIETY

KHALED B. LETAIEF, VICE-PRESIDENT CONFERENCES
SERGIO BENEDETTO, VICE-PRESIDENT MEMBER RELATIONS
JOSÉ-DAVID CELY, DIRECTOR OF LA REGION
GABE JAKOBSON, DIRECTOR OF NA REGION
TARIQ DURRANI, DIRECTOR OF EAME REGION
NAOAKI YAMANAKA, DIRECTOR OF AP REGION
ROBERTO SARACCO, DIRECTOR OF SISTER AND RELATED SOCIETIES

REGIONAL CORRESPONDENTS WHO CONTRIBUTED TO THIS ISSUE

BIN HU, USA (HUBENJAMIN@YAHOO.COM)
BORHANUDIN MOHD ALI, MALAYSIA (BORHAN@ENG.UPM.EDU.MY)
EWELL TAN, SINGAPORE (EWELL.TAN@IEEE.ORG)

A publication of the
IEEE Communications Society

PROJECT iREACH/*continued from page 3*

role in focusing attention on social issues, including domestic violence and security.

It emerged that iREACH has brought many benefits to the people and has been responsive to the needs of communities within its catchment areas. Some of the key findings of iREACH are:

1) It has become part of the social infrastructure in the communities and has driven many positive changes by heralding new forms of engagement between different sectors of the communities, including those who are more marginalized. The project's spirit of volunteerism in its activities, from research to the preparation of audio programs and passing on computer skills to community members, has enhanced positive social capital.

2) It has generated interest in learning among children, youth, and adults. Adults expressed their appreciation for the opportunity to learn through informal education. Students have become more inquisitive, with 200 XO laptops donated by Elaine Negroponte in 2009.

3) It has contributed to improvements in farm practices and income through informal education, primarily in the form of village-to-village broadcasts and lectures.

4) It has improved health, through sanitation and other health related information, and this has reduced people's medical expenses. This has also encouraged many to make more effective use of government health facilities.

5) It has facilitated communication, interaction, and cooperation within families and villages, and between villages equipped with hubs.

6) It has made information accessible on topics produced by other agencies, thereby strengthening the impact of those other initiatives.

7) Research suggests that the project has not yet had much influence on government policy, and there was limited reference to the project being used for, or having spawned, entrepreneurial activity.

8) The project has had an empowering influence, particularly on women, who are in a stronger position to effect improvements in their quality of life. The influence of the project in reducing domestic violence emerged as a strong contributing factor to this, both through education and by providing a meeting place. There were also a few references to villagers being more empowered to engage with people in authority (e.g., report issues to the police and communicate with council members). (Contact e-mail: unger@ucalgary.ca)

References

[1] H. Grunfeld *t al.*, "Informatics for Rural Empowerment and Community Health (iREACH Cambodia). How Villagers Have Used and Benefitted from iREACH: Report on Participatory Evaluation," in preparation, 2010.

NEW BROADBAND ERA/*continued from page 1*

Among other issues, it is not clear that this plan will bring broadband to low population density areas, of low interest to service providers. The absence of public funds might endanger deployment in most areas and thus reduce its benefits. President Obama has compared the deployment of this huge plan to the history of the trail, and for sure it will mark a milestone in the short and recent history of the Internet.

References

[1] http://www.broadband.gov/

GUEST EDITORIAL

MILITARY COMMUNICATIONS

Torleiv Maseng *Randall Landry* *Kenneth Young*

Since its introduction in the late 1990s, the concept of network-centric operations to enable information sharing has been a fundamental element of the vision of military organizations throughout the world. As the complexity of operational environments increases, military communications network technologies must provide key characteristics such as self-organization and decentralization, which will speed information flow and increase situational awareness. In addition, given the rapid evolution of commercial communications network technologies, and the applications and services that utilize them, military organizations are increasingly looking at adopting commercial communications technologies where practical. This is especially true in the mobile wireless communications that are at the heart of tactical military operations and have seen such rapid adoption in the commercial space. The five articles that make up this year's Feature Topic on Military Communications provide an excellent overview of these trends in military communications — increasing dependence on advanced technologies and leveraging of commercial technology where possible.

The first article, "A Comparative Review of Commercial vs. Tactical Wireless Networks," presents a comparison of requirements, uses, and constraints between current commercial wireless networks (specifically third-generation [3G] and 4G networks) and tactical wireless networks (specifically mobile ad hoc networks [MANETs] built using software-defined radios). The author proposes a model for tactical wireless networks that is based on the Future Force vision enabled by the Joint Tactical Radio System program pursued over the last decade. As is the case for commercial wireless architectures, the proposed model has well defined separate entities and well defined interfaces. The author argues that adopting such a model would allow tactical wireless networks to achieve the phenomenal successes (lower cost and wide-scale deployment) seen in commercial wireless networks.

The article "Peer-to-Peer Communications for Tactical Environments: Observations, Requirements, and Experiences" describes the application of peer-to-peer (P2P) approaches to realize key military applications such as Blue Force Tracking, inter-team communications, remote unmanned vehicle control, and sensor data mining/fusion in tactical environments. Experiences from several tactical networking experiments are used to derive requirements for P2P middleware suitable for tactical environments. The case study of an Agile Middleware Computing solution that meets these requirements is presented along with experimental results that quantify its effectiveness in the target environment.

One approach to enabling tactical networks to adapt to rapidly changing conditions and user needs is the use of cognitive functions across the protocol stack to better utilize scarce spectrum, and dynamically adapt network functionality and configurations. The article "Cognitive Tactical Network Models" proposes a cognitive tactical network model that captures the specifics of tactical military networks and shows how a network designer might use the model to formulate cognitive network design problems and solve them using a design tool based on the model. The authors describe such a design tool, including the various design choices and configuration settings that could be optimized at each layer of the protocol stack. Using such a tool, a network designer could both efficiently explore the many dimensions of the design space as well as examine the robustness of various designs in the face of unforeseen circumstances.

The article "Robust Web Services in Heterogeneous Military Networks" looks at the potential to using service-oriented architecture (SOA) approach based on Web services to enable the NATO information infrastructure, which is a federation of different information and communications systems. Extending such an approach to tactical networks must deal with the issues of exchange of large XML documents in a bandwidth-challenged environment, the difficulty of using TCP in networks with high delay and/or intermittent connectivity, and interconnection of heterogeneous network types. The authors propose a middleware system called Delay and Disruption Tolerant

(Continued on page 52)

(Continued from page 50)

SOAP Proxy (DSProxy) that forms an overlay network, and uses a variety of innovative techniques to reduce overhead and add delay tolerance. A field trial using DSProxy demonstrated the feasibility of using this approach to support unmodified Web services software across heterogeneous tactical networks.

Another aspect of using Web services in tactical environments is explored in the article "Web Services Discovery across Heterogeneous Military Networks." The authors address the issue of locating Web-service-related resources that meet certain criteria in the dynamic and heterogeneous environments found in NATO military operations. An approach using service discovery gateways is proposed as a means of interconnecting the service discovery mechanisms across the heterogeneous networks without requiring changes to existing clients and services. The article also discusses Web services discovery standards, tailor-made solutions for multihop MANETs, and an interoperability gateway prototype that was implemented and successfully tested in a field trial.

The editors hope that these articles provide our readers with a good cross-section of some of the challenges and developments in military application of wireless communications. The five articles that make up this year's Feature Topic on Military Communications were selected from nine high-quality manuscripts received. We encourage researchers and developers in this area to review the 2011 Call for Papers for the Military Communications Feature Topic and consider submitting an article in one of the relevant areas.

BIOGRAPHIES

TORLEIV MASENG (torleiv.maseng@ffi.no) is director of research at the Norwegian Defense Research Establishment, where he is responsible for communications and information systems. He worked as a scientist at SINTEF in Trondheim for 10 years, involved in design and standardization of GSM. For seven years he was a scientist at the NC3A NATO research center in The Hague. During 1992–1994 he was involved in the startup of the new private mobile operator NetCom GSM in Norway, where he had technical responsibility. Since 1994 he has held a chair in radio communications at the University of Lund, Sweden. In 1996 he took up his employment at the Norwegian Defense Research Establishment (FFI) located at Kjeller, 20 km outside Oslo. Since 2005 he is also Professor II at the University of Oslo. He is the author of more than 150 papers, holds patents, and is a Technical Editor of *IEEE Communications Magazine*. He has received an award for outstanding research and has arranged large international conferences.

RANDALL LANDRY (rlandry@mitre.org) received his M.S. and Ph.D. in electrical engineering from the University of Vermont in 1992 and 1994. He is currently with the MITRE Corporation in Bedford, Massachusetts, where he serves as program director for a portfolio of programs that deliver ground-based communications capabilities to Air Force users in garrison and deployed. He has also served as department head for communications and networking, and conducted research in support of the U.S. Department of Defense. As a member of Corporate R&D at Texas Instruments, Dallas, he was previously involved in research and development of highly integrated switching architectures for gigabit networking and holds several patents in this area. He has also served as director of optical and wireless networking in the telecommunications industry. He has been the principal investigator on a number of research programs ranging from satellite communications to tactical wireless networking. Recent research interests focus on network science, autonomic networking, cross-layer design methodologies, dynamic resource management in wireless networks, and the performance evaluation of multihop wireless networks. He has published numerous technical articles, and served as technical program committee member and session organizer for major IEEE communications conferences.

KENNETH YOUNG [SM] is executive director for Government Project Development in the Applied Research organization at Telcordia Technologies, Piscataway, New Jersey. He received his B.S. in physics from St. Joseph's University, and his M.S. and Ph.D. in physics from the University of Pennsylvania. His research interests are in the design and application of mobile ad hoc networking technology for tactical environments. He is the program manager for the Army Research Laboratory's Communications and Networks Collaborative Technology Alliance, a government-industry-academic consortium that performs basic research in survivable wireless mobile networking, signal processing, and tactical information protection. He heads another team developing advanced mobile technology under the U.S. Army CERDEC's Proactive Integrated Link Selection for Network Robustness program. He chairs the Communications Society's Tactical Communications and Operations Technical Committee and is on the advisory board for the Military Communications (MILCOM) Conference. He is a Telcordia Fellow.

A Comparative Review of Commercial vs. Tactical Wireless Networks

George F. Elmasry, DSCI

ABSTRACT

This article presents a comparison between commercial and tactical wireless networks, pointing to their different requirements, expectations, needs, and constraints for information assurance and so on. This comparative study demonstrates why commercial wireless networks have made more technological leaps than tactical wireless networks. The article introduces a model for tactical wireless networks based on the architectures that have been pursued (especially the Joint Tactical Radio System and the Future Force vision) within the last decade or so. The model draws parallels with commercial wireless networks to allow tactical wireless networks a plausible opportunity to achieve similar technological advancements.

INTRODUCTION

Commercial wireless networks (specifically 3G and 4G) [1] have significantly progressed within the last two decades; with improved service, most locations now have excellent connectivity, and overall costs have been reduced drastically. Today, one can purchase a low-cost smart phone with multimedia services anywhere in the United States, travel internationally, then utilize said device for wireless service at nearly any other location. Even the developing countries which missed out on the wired, microwave and satellite eras now have excellent wireless infrastructure [2]. The monumental success of commercial wireless networks is evident by the steep price drops of high quality services, prevalent connectivity and compact end-user equipment with innovative features.

The latest tactical wireless networks are a modest improvement over the original networks that were in use when Congress adopted the goal of having light agile forces for fast military deployment [3]. This wave of tactical wireless networks began simultaneously with the latest wave of commercial wireless networks. This resulted in the development of a new generation of mobile ad hoc networks (MANETs), which would allow the Warfighter to obtain connectivity and deliver two-way information anywhere in the war theater with the speed and quality required by the tactical needs. Then, the Joint Tactical Radio System (JTRS) and Future Force networks came to light [4], while the ability to bring the global information grid (GIG) to the tactical edge, with full IP-based MANETs, began facing many delays and substantial increases in cost [5]. This prompted high-ranking military/government officials to ask: Why is it possible to have low-cost and high-quality service from commercial vendors but not from tactical wireless networks?

Tactical networks differ from commercial networks in their security constraints, anti-jamming needs, mobility/dynamic topology, scarcity of bandwidth, excessive delay, and other characteristics. The plethora of major differences can justify the lag in achieving the goals set for tactical wireless networks. This article argues that there are noteworthy lessons to be learned from the phenomenal success of the commercial model. The tactical wireless network community can be inspired to consider a different model that borrows from commercial networks and constructs a more contained tactical wireless model. This model would meet tactical requirements, leave ample room for advancement, and simultaneously make it possible to build and deploy the desired tactical wireless infrastructure at a low cost.

The next section provides a quick overview of the commercial wireless model, while the following section presents a notional view of the current tactical wireless model. I then describe the proposed generic tactical wireless node model, with the following section summing up the lessons learned. The article concludes with a summary.

COMMERCIAL WIRELESS MODEL

This section assesses the general architecture of commercial wireless networks [1], highlighting their successful creation of excellent coverage and delivery of affordable yet reliable services. A typical commercial service provider has mobile end users and fixed network towers with base stations (BSs). Multiple BSs may be connected to a BS controller (BSC), and BSCs are connected to a core network through a gateway.

The details regarding the functionality of commercial wireless protocols are beyond the scope of this article. This section points to certain regulations established through standardization committees, such as the Third Generation

0163-6804/10/$25.00 © 2010 IEEE

Partnership Project (3GPP) [1]. An important aspect of this technology is the definition of open standard interface control documents (ICDs). Consider a typical message flow (end user to end user), where packets flow from an end user to a BS, to a BSC, then to the core network; and onward to a BSC, to a BS, and to the other end user. Figure 1 shows the four main entities of this flow, with three main ICDs needed to define the interfaces. ICD I defines the end-user-to-BS air interface, ICD II defines the BS-to-BSC interface, while ICD III defines the BSC-to-core-network interface. Note that there are more detailed ICDs, such as those defining end-user handover between two different BSs. This section advocates an open architecture with well defined entities and interfaces, which would provide the following advantages:

Containment: Organizations that develop the technologies (i.e., vendors) need not be concerned about defining interfaces. Open architecture makes them adhere to predefined interfaces and focus on functionalities within the entity (user equipment, BS, BSC, core network, gateway, etc.)

Competition: To build a network, a service provider can purchase some BSs from vendor A, others from vendor B, some BSCs from vendor C, others from vendor D, a core network gateway from vendor E, and so on. The service provider can then interconnect these devices seamlessly and build a reliable network. This creates competition and reduces cost.

Specialization: With this approach, some vendors can specialize in building user equipment and successfully market it at a low price, others can specialize in BSs, and so on. Specialization leads to innovation and drastic cost cutting.

Innovation: With standardization committees defining the entities and alleviating concerns about interoperability, the vendors can focus their efforts on important issues such as traffic engineering, resource management, quality of service (QoS), and call admission control (CAC).

To a large extent, one can attribute the affordability and increased quality of commercial wireless to the above advantages. One also must recall the role of market size; with millions of end users [2], vendors invested in developing the best competitive entity possible, and service providers invested in planning the commercial wireless infrastructure. These investments were based on the projected return on investments that would be realized through the ever growing wireless user population. Obviously, there is no market size for tactical networks. However, the lessons learned from the commercial wireless model can and should be adopted for the tactical world, as discussed later.

Tactical Wireless Model

Tactical wireless networks are extremely complex. Unlike commercial wireless, there is no fixed infrastructure. One cannot use a fixed location for a BS. Also, there is no equivalent to the commercial infrastructure's notion of a static core network. This section depicts a notional view of tactical wireless networks to assist in a meaningful comparison of the two.

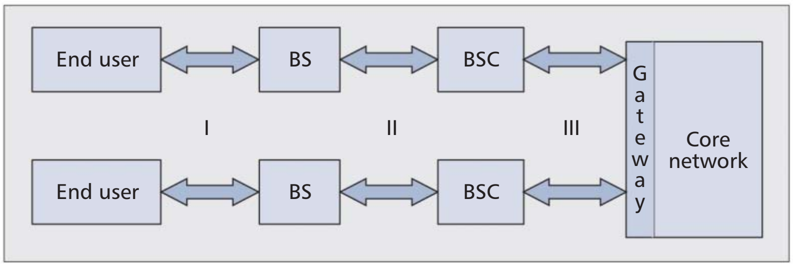


Figure 1. *Commercial wireless definition of entities and well defined ICDs.*

Tactical wireless networks built with the JTRS in mind have tiers of subnets (islands of MANETs). These subnets are built up with waveforms (a waveform is a wireless multiple access radio frequency technology). Figure 2 depicts a notional view of such layered (hierarchical) islands (subnets) of MANETs. There is the soldier radio waveform (SRW) tier [6], which can have two subtiers, one for soldier-to-soldier communications and one for networking sensors. Above that, there is the wideband networking waveform (WNW) tier [7], which also has two subtiers; one forms local subnets for vehicle-to-vehicle communications, and the other is for global connectivity, to generate a single subnet over the entire theater. Note that at each tier, there can be multiple subnets with different frequencies (except for the global subnet), forming islands of MANETs. Some selected nodes can have multichannel capability to access different subnets and work as gateways between the subnets. Above these lower echelons layers, comes the upper echelon layer with another wireless mobile core network (e.g., warfighter information network-tactical, WIN-T [8]), which itself can have a mix of fixed nodes and tactical command mobile nodes. This tactical core network can use waveforms such as high-band networking waveform (HNW) [9] or network-centric waveform (NCW) [10], which offers satellite communications on the move (OTM); in addition, the tactical core network utilizes microwave links for networking stationary or on-the-quick-halt nodes.

Contrary to commercial wireless networks, tactical wireless networks nodes can be all mobile; there is no fixed infrastructure, and the end-user nodes are part of the infrastructure. Gateway nodes (with multichannel radios or multiple waveforms) are not a single point of failure since a subnet can have multiple nodes that can relay traffic between hierarchal layers to create full connectivity anywhere in the theater.

Another important aspect of tactical wireless networks is the node architecture. The National Security Agency (NSA) has mandated the use of the High Assurance Internet Protocol Encryptor (HAIPE) [11]. A tactical node must have the internal notional architecture shown in Fig. 3. Each security enclave (plain-text subnet) must have its own HAIPE encryptor to transfer data through a cipher-text core network. This creates a plain-text (red) IP layer, separated by HAIPE from a cipher-text (black) IP layer. The NSA requires that no information be shared between the red enclaves and the black core with the exception of the type of service (TOS) byte in

Communications | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | qMags

The combination of packet loss, limitation of available bandwidth over some core network paths, and the red/black separation issue leave the red enclaves with many challenges regarding admission control, multicast, and ensuring protocols such as TCP work, etc.

the IP header. This is referred to as red/black separation, where passing congestion information from the black core to the red enclaves (where admission control is needed) is prohibited. Passing topology information is also prohibited, yet it is specially needed for multicast since the JTRS uses multiple access waveforms (a single packet over the air can reach all nodes in a multiple access waveform). For multicast to work efficiently at the red IP layer, one needs to know which destinations in a multicast address are reachable over the same waveform. Another HAIPE constraint is the encryption key, which is downloaded onto a given radio based on its intended deployment. Before the JTRS node can join a different subnet with a different encryption key, it needs to be shut down, new encryption key needs to be loaded, and the node needs to reboot. This constraint eliminated one of the best advantages of MANET radios since it limits roaming. As a result, in actual deployment one can see nodes being orphaned and subnets split for some period of time.

This architecture applies to all tactical nodes. For a large command and control (C2) node, each plain-text enclave can be a local area network (LAN) with a LAN router and hundreds of end users. Commercial off-the-shelf (COTS) routers (Cisco or Juniper) are used for the black IP layer where multiple waveforms from different radio types can be used to form multiple links to different nodes at different tiers. The same node architecture applies to a lower echelon node using a single-channel JTRS SRW with a single security enclave being a single IP port [12]. Note that within JTRS, all the layers are developed in software, making JTRS a software programmable radio (SPR).

With JTRS non-legacy waveforms, there is little commonality between the different waveforms (SRW, WNW, etc.). Each waveform can have different radio implementations (medium access control [MAC] and data link layer [DLL]), black IP layer implementation, red IP layer implementation, and even different versions of HAIPE.

Within tactical networks, mobility and jamming can cause high bit error rates beyond what the radio DLL or MAC can correct, resulting in a large percentage of erroneous packets. With HAIPE dropping any packet with even a single bit error, red enclaves face a large percentage of packet loss. The combination of packet loss, limitation of available bandwidth over some core network paths, and the red/black separation issue leave the red enclaves with many challenges regarding admission control, multicast, ensuring protocols such as TCP work, and so on.

Proposed Generic Tactical Wireless Node Model

This section presents a generic model for open architecture tactical nodes that avoids some of the pitfalls of the JTRS model (although adhering to the same general architecture) and simultaneously borrows concepts from the commercial

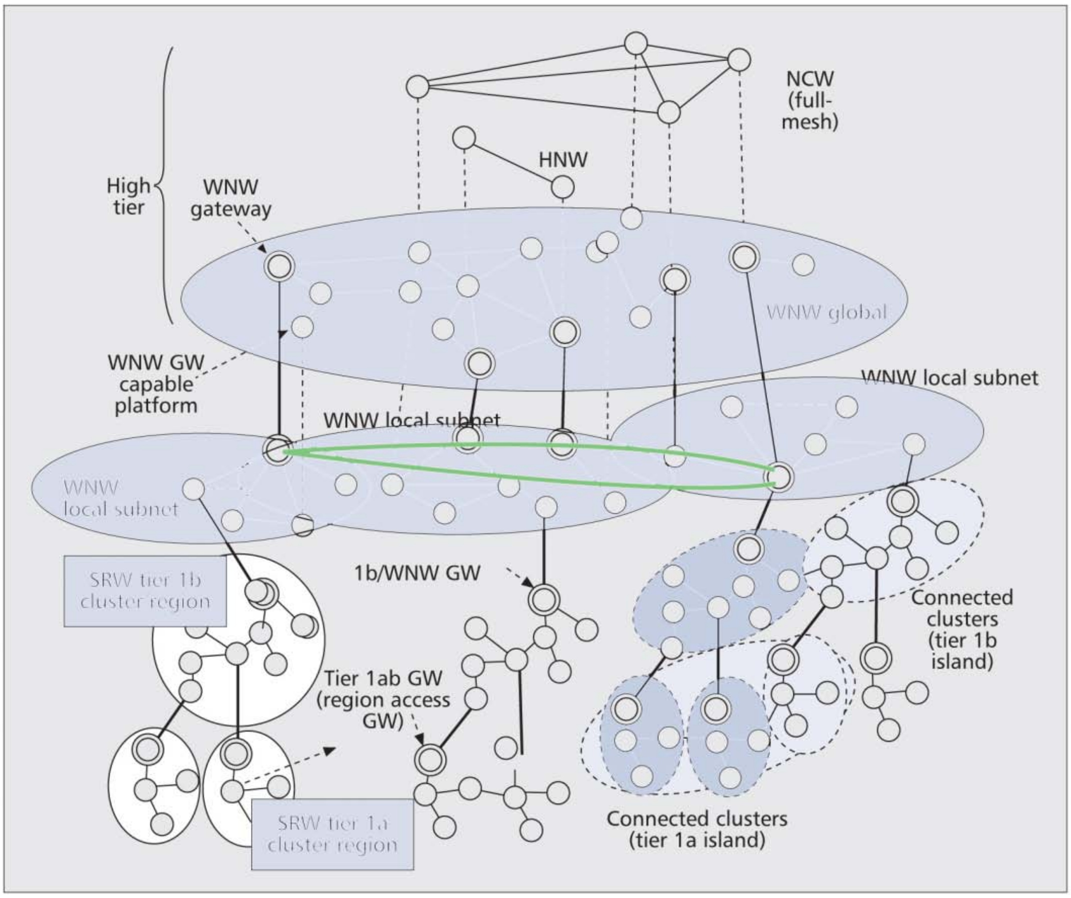


Figure 2. *Notional view of tactical wireless networks with hierarchal subnet islands.*

model. This section starts by pointing out some of the JTRS pitfalls.

No separation of entities: A JTRS waveform (SRW or WNW) has been defined as a single entity. The WNW waveform has many layers and sublayers (e.g., mobile Internet [MI] and mobile data link [MDL]) specific to this waveform only. The current JTRS approach defines a complex waveform as a single program of record that covers all the protocol stack layers from the plain-text side transport layer to the physical layer; this creates no commonality between the peer layers from different waveforms. In other words, JTRS is *not* an open architecture model.

Limited room for innovation: Only one vendor is supported by the government to integrate a waveform, with subcontractors awarded different layers. With the WNW waveform, although a design of a single layer to replace MI and MDL is theoretically possible [13], there is no room for such innovation since the entities are *not* defined in an open architecture manner.

No room for competition: Deployment of a tactical network is left to a single integrator who faces many overwhelming challenges, especially with interoperability, and tends to only trust the in-house solutions. This can lead to cost overruns before bringing the technology to the warfighter.

No room for contributions from specialized organizations: If the radio frequency (RF), DLL, and MAC layers were defined as a separate entity (radio entity) with a well defined ICD to the black IP layer, small firms specializing in this type of technology would have room to produce low-cost waveforms. These specialized firms could produce a wide variety of radios adhering to the ICD details for interface to the black IP layer.

Figure 4 presents an open architecture model for tactical wireless nodes that can be adapted to a C2 node as well as to solider light weight radios (similar to the well defined tactical model in Fig. 3). This model attempts to separate the entities and to use well-defined ICDs. It attempts to create entities for the specialized organizations to focus, create, innovate, and contribute, thus helping to reduce cost and improve performance.

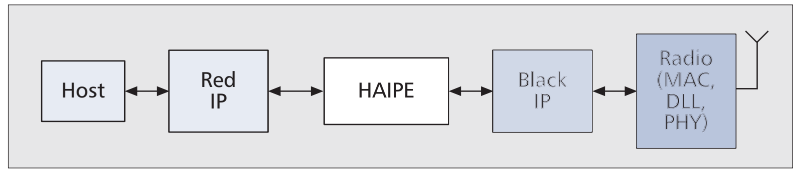


Figure 3. *Notional view of a tactical node with the HAIPE separating the plain-text and cipher-text IP layers.*

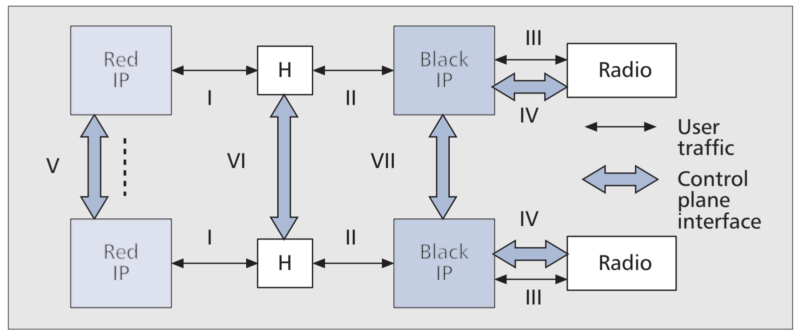


Figure 4. *Tactical wireless model with well defined entities and ICDs.*

Tactical Wireless Model Entities

Red IP Layer — This layer is a fully capable IP layer with added functionality to address the existence of HAIPE, which makes the red IP layer need extra capabilities, such as performing multicast over the encrypted core. Note that HAIPE makes the entire encrypted core look like a black box to the red IP layer, which has to pay attention to the multicast protocol. Other functions, such as CAC over HAIPE [14, 15], flow control, and delivery assurance over HAIPE [16], can be additional red IP layer functionalities that will make this layer different from standard IP. A common software solution for this red IP layer can be defined across all tactical nodes and be deployed as part of an SPR or as a LAN router on standalone hardware for large-scale nodes such as C2 nodes.

It should be noted that the absence of the red IP layer definition as a standalone entity in the JTRS radio had created conflict when the WNW waveform vendor implementation of multicast at the red IP layer was questioned as a violation of NSA's information assurance integrity [17]. This implementation passed some topology information from the black core to the red enclaves, indicating which destinations could be reached through the same waveform. This made the red IP implementation of multicast efficient (there was no need to duplicate packets for each destination over the same waveform). Although multicast efficiency was urgently needed since a large percent of traffic is multicast-based, this breach was considered a violation of NSA's red/black separation. Such issues need to be addressed during standardization and not left to surface during deployment. Also, there is no common agreement on how different red IP layers will communicate across different programs of record as well as across joint forces (i.e., there is no definition of ICD V in Fig. 4). Having the stakeholders come together at a forum, similar to the standardization committees for commercial wireless, and agree on a solution before starting the development is the engineering model that must be adopted if the tactical wireless concepts are revisited. Note that different programs have come up with different solutions for CAC over HAIPE, flow control over HAIPE, delivery assurance over HAIPE, and so on. These solutions are not always compatible [14, 15]. Defining ICD V in Fig. 4 will create commonality, and will make the R&D community focus on optimizing these add-on techniques instead of being bogged down in the interoperability details.

HAIPE — Although HAIPE standards are very well defined, commonality is an important issue. Different JTRS radios are currently planned to be deployed with different versions of HAIPE. Although HAIPE versions are backward compatible, commonality can be very advantageous. Consider, for example, CAC over HAIPE [14, 15]. Reference [14] is a solution based on HAIPE version 3.0 or earlier, while [15] is based on version 3.1 and shows that a more powerful

If a well-defined interface between the black IP layer and the radio is created, one can contain radios to their intended purpose: MAC, DLL, and Physical layers. One can then establish a focus and have ample room for innovative development of different types of radios.

CAC can be developed for this version where the explicit congestion notification bits are allowed to pass between the red and black sides.

Black IP Layer — This is the most complex entity for tactical wireless networks. Starting from COTS IP routers, tactical wireless black IP ran into many complex problems. These problems include how to create load balancing over different radio links; how to adapt to changes in the radio bandwidth (BW) (i.e., the service rate of the black IP layer changes as the radio throughput changes); how to contain open shortest path first (OSPF) protocol traffic in scaled subnets (radio OSPF [ROSPF] is specific to the JTRS WNW radio and is not common across all tactical nodes that are not JTRS nodes); how to accommodate and standardize protocols like Point-to-Point Protocol over Ethernet (PPPoE) used between some satellite terminals and COTS routers; and how to overcome intermediate blockage of the physical layer through the use of techniques such as network coding [17]. Thus, one sees three areas to be considered for the black IP entity:

- Black IP layer functions and how the user traffic flows to the radios (ICD III in Fig. 4 defining the user plane).
- Cross-layer signaling with the radios (ICD IV) where protocols like ROSPF require information from the lower layers to help contain the increase in link state advertisement (LSA) traffic when a subnet is scaled. This ICD can include standardization of protocols like PPPoE to create commonality on how to synchronize the black IP layer's service rate to the radio's fluctuation in BW.
- Peer interface with the peer black IP layers (in remote nodes) to address issues like load balancing (ICD VII).

Radios — If a well defined interface between the black IP layer and the radio is created, one can contain radios to their intended purpose: MAC, DLL, and physical layers. One can then establish a focus and have ample room for innovative development of different types of radios. There is even room to adopt very mature commercial technologies as long as interfaces III and IV are adhered to. Organizations specializing in this line of technology will have the opportunity to introduce cutting edge waveform technologies varying from waveforms with small areas of coverage using omnidirectional antennas (such as 802.11) to large-area waveforms (e.g., WiMax) to even point-to-point links or satellite links, all adhering to ICDs III and IV.

Tactical Wireless Model ICDs

Please see [1, 2] and other 3GPP and IETF references for the latest commercial ICD details. Before introducing the tactical wireless model ICDs, it is worth mentioning that for commercial models, only the minimum interoperability needs are defined in the ICDs; most details are left for the implementer, as long as the interfaces work seamlessly. The ICDs within the tactical wireless model, shown in Fig. 4, can be defined as follows.

ICD I, Red IP to HAIPE — This ICD covers the user traffic flow and can address issues including the maximum size allowed for a red IP packet entering the HAIPE. This packet size can be chosen to avoid fragmentation over the black IP layer (which can reduce throughput efficiency), accommodate the maximum Ethernet frame size, and consider the overhead introduced by HAIPE. Note that there is a clear difference between the HAIPE constraint (red/black separation) and IPSec in commercial networks. This ICD can explicitly define these differences. Please see [14, 15] about implementation of CAC over HAIPE; [16] on how user traffic flow over HAIPE such as TCP can be in proxy over HAIPE; and [18] on how control signaling such as Resource Reservation Protocol (RSVP) can be in proxy over HAIPE. The details of these differences are beyond the scope of this article. It is worth mentioning here that the concept of using measurement-based admission control (MBAC) in [14] was based on commercial wireless research in how admission control can be done in a federation of networks.

ICD II, HAIPE to Black IP — As the HAIPE standards are well defined, in this model one needs to adhere to a singular version of HAIPE to create commonality.

ICD III, Black IP to Radio: User Traffic — This ICD covers user traffic flow and can include standardization of issues like QoS prioritization (i.e., which packet is given precedence in service when there is contention for BW). Reference [19] details how this QoS area should be tied to the red IP layer implementation of QoS in order to create a coherent end-to-end solution for QoS across tactical networks.

ICD IV, Black IP to Radio: Control Plane — This ICD covers all cross-layer signaling issues between the radio and the black IP layer like BW fluctuations and how link states can be reported from the radio to the black IP layer. Queue depth at the black IP layer (reflecting traffic demand) can be reported from the black IP layer to the radio (some radio terminals with dynamic resource management between multiple terminals can use this to dynamically reallocate resources based on the traffic demand), and so on. The key here is to identify the needed capabilities, reach the best solution, and create standardization.

ICD V, Red IP Peer-to-Peer — This defines the peer relationship between the red IP layers in different nodes. This ICD is essential in order to have a common solution for issues like multicast over the encrypted core. This ICD could include standardization of the best possible solution for CAC over HAIPE, multicast over HAIPE, RSVP over HAIPE, and so on.

ICD VI, HAIPE Peer-to-Peer — As mentioned for ICD II, HAIPE standards are well defined, and adhering to one version can help create common solutions across the tactical nodes.

ICD VII, Black IP Peer-to-Peer — This ICD will define issues including load balancing over differ-

IEEE Communications | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | qMags

ent paths (e.g., sending voice over terrestrial links and data over satellite links), and how to implement dynamic area IDs where (in the case of multiple areas at the encrypted core) a mobile node can detach itself from one area and join another.

The above model is essentially the JTRS model, but with well defined separate entities and well defined interfaces. In addition to creating smaller problem spaces, these definitions encourage commonality, containment, and competition while simultaneously allowing ample room for innovation in the specialization areas.

Lesson Learned

With the conception of the Future Force vision in the 1990s, the Department of Defense (DoD) community could have created pilot programs (without funding the actual development) to implement models such as the open architecture model presented in Fig. 4. These pilot programs would have involved the stakeholders and relied on R&D expertise to create ICDs and flush out all the interoperability and efficiency issues prior to funding programs of record to develop technologies in uncharted territories. Many organizations, especially small businesses, would have invested their own resources and focused on creating software-based solutions (for entities including the red IPs, radios, etc.) while adhering to the predefined ICDs. More complex entities, such as the black IP layer, could have become focused programs on their own. Note that the tactical router idea has been floating around for quite some time, and the presented black IP layer with cross-layer signaling to the radio considers many aspects of a tactical MANET router. Although black IP could be the most complex entity in this model, it remains much simpler than the JTRS concept that covers all the stack layers.

The open architecture model presented does not require the integrators to define interfaces or conduct R&D to find out how layers peer with each other. Rather, this action is left for R&D organizations ahead of the development phases. With this model, the job of integrators will be to harness tested technologies and build the systems (much as commercial service providers do). Different organizations can produce plenty of waveforms (MAC, DLL, and physical layers) with a variety of capabilities that adhere to the defined ICDs and can be utilized as plug-and-play with the black IP layer.

Notice that the premises of this model are simple engineering principles such as "do not attempt to solve a problem before you define it" and "keep it simple and straight." Even though the DoD community knew the value of SPRs, knew it could be materialized, and knew it will serve the warfighter well, there was a rush to make it happen instead of taking the time to define entities (to define less complex problem spaces) and invest in defining interfaces. This is the main, although not the only, reason they did not excel to the extent of the commercial model. Having realized the challenges with the current model, the industrial organizations in the DoD arena are now attempting to create a consortium to address standardization issues with tactical networks [20].

Summary

This article presents a comparison between commercial and tactical wireless networks. The article points out why commercial wireless networks have made many technological advances that tactical wireless has yet to catch up with. The article also introduces a model for tactical wireless networks utilizing the lessons learned from the tactical network models — JTRS and Future Force vision — pursued within the last decade or so. The model also attempts to draw from the success of commercial wireless to make it possible for tactical wireless networks to achieve similar results.

References

[1] 3GPP; http://www.3gpp.org/
[2] UMTS Forum; http://www.umts-forum.org/
[3] United States Army; http://www.globalsecurity.org/military/library/report/2003/Future_Force_Black_Book_26_Aug_03_Final.pdf
[4] R. North, N. Browne, and L. Schiavone, "Joint Tactical Radio System — Connecting the GIG to the Tactical Edge," *Proc. IEEE MILCOM '06*, Washington, D.C., Oct. 23–25, 2006.
[5] J. Stine, "A Cautionary Tale on Testing and Evaluating Tactical Wireless Mobile Ad Hoc Networks," *Proc. ITEA J.*, Mar./Apr. 2007, pp. 53–62.
[6] "Soldier-Level Integrated Communications Environment (SLICE) Soldier Radio Waveform (SRW) Functional Description Document (FDD)," v. 1.3, Nov. 2003.
[7] "Wideband Networking Waveform (WNW) System Segment Specification," Boeing, spec. no. AJ01120, Feb. 12, 2003.
[8] http://peoc3t.monmouth.army.mil/win_t/win_t.html
[9] http://www.govcomm.harris.com/solutions/products/000056.asp
[10] J. Wiss and R. Gupta, "The WIN-T MF-TDMA Mesh Network Centric Waveform," *Proc. IEEE MILCOM '07*, Orlando, FL, Oct. 29–31, 2007.
[11] United States National Security Agency, "High Assurance Internet Protocol Encryptor Interoperability Specification," v. 3.1.0, Dec. 31, 2006.
[12] http://jpeojtrs.mil/
[13] S. Dastangoo *et al.*, "Performance Analysis of Distributed Time Division Multiple Access Protocols in Mobile Ad Hoc Environments," *Proc. IEEE MILCOM '09*, Boston, MA, Oct. 19–21, 2009, paper no. U122.
[14] G. F. Elmasry, C. J. McCann, and R. Welsh, "Partitioning QoS Management for Secure Tactical Wireless Ad-Hoc Networks," *IEEE Commun. Mag.*, Nov. 2005, pp. 116–23.
[15] G. F. Elmasry *et al.*, "ECN-Based MBAC Algorithm for Use Over HAIPE," *Proc. IEEE MILCOM '09*, Boston, MA, Oct. 19–21, 2009, paper no. U310.
[16] Y. L. Grushevsky *et al.*, "Adaptive RS Code for Message Delivery over Encrypted Military Wireless Networks," *Proc. IEEE MILCOM '06*, Washington, D.C., Oct. 23–25, 2006.
[17] R. Ahlswede *et al.*, "Network Information Flow," *IEEE Trans. Info. Theory*, vol. 46, no. 4, July 2000, pp. 1204–16.
[18] G. F. Elmasry *et al.*, "Reservation-Based Quality of Service (QoS) in AN Airborne Network," *Proc. IEEE MILCOM '09*, Boston, MA, Oct. 19–21, 2009, paper no. U240.
[19] G. F. Elmasry *et al.*, "Achieving Consistent PHB across the GIG, A QoS Common Denominator," *Proc. IEEE MILCOM '08*, paper no. NC7-1.
[20] http://www.opengroup.org/tech/direcnet-task-force/

Biography

George F. Elmasry [SM] (gelmasry@dsci.com) received his B.Sc. in electrical engineering from Alexandria University, Egypt, and his M.S. and Ph.D. degrees in electrical and computer engineering from New Jersey Institute of Technology. He is currently working as a senior VP at DSCI heading its C3 division. Before DSCI, he was with General Dynamics C4S, in the Warfighter Information Network-Tactical (WIN-T) group. Prior to General Dynamics he worked for Lucent Technologies in the Wireless Networks Wideband CDMA group. His research interests are in the areas of wireless networking, ARQ, joint source and channel coding, and multidimensional interleaving, where he has over 50 publications and patents.

The open architecture model presented does not require the integrators to define interfaces or conduct R&D to find out how layers peer with each other. Rather, this action is left for R&D organizations ahead of the development phases.

Peer-to-Peer Communications for Tactical Environments: Observations, Requirements, and Experiences

Niranjan Suri and Giacomo Benincasa, Florida Institute for Human and Machine Cognition
Mauro Tortonesi and Cesare Stefanelli, University of Ferrara
Jesse Kovach and Robert Winkler, U.S. Army Research Laboratory
Ralph Kohler and James Hanna, U.S. Air Force Research Laboratory
Louis Pochet, U.S. Air Force Reserves
Scott Watson, Space and Naval Warfare Systems Center, Pacific

ABSTRACT

Tactical edge networks present extremely challenging environments for communications given their wireless ad hoc nature and the inherent node mobility. Military applications such as Blue Force Tracking, inter-team communications, remote unmanned vehicle control, and sensor data mining/fusion thus have to deal with unstable links with limited bandwidth and variable latency. The peculiar characteristics of tactical networks call for peer-to-peer approaches to realize complex, adaptive, and fault-tolerant applications to be deployed in the battlefield. This article reports on our observations from several tactical networking experiments in which we have deployed state-of-the-art applications and services that leverage P2P communications. More specifically, we discuss why P2P approaches are critical for tactical network environments and applications. We then analyze the requirements that should be satisfied by P2P middleware for tactical environments. Finally, we discuss a case study, the Agile Computing Middleware, and present experimental results that demonstrate its effectiveness.

INTRODUCTION

Tactical edge networks present an extremely challenging communications environment for application developers and users. These networks are built from ad hoc wireless connections between mobile nodes, providing unstable links with limited bandwidth and variable latency. In urban environments, performance is further degraded by occlusions from buildings and interference from consumer electronics and civilian transmitters. Meanwhile, a greater number of assets equipped with higher-fidelity sensors, along with the proliferation of new end-user applications, are placing ever increasing bandwidth demands on the network. Systems operating in these tactical networks must be capable of providing reliable and timely information exchange within this unreliable and congested communications environment. Creative solutions that combine multiple techniques are needed to successfully address these challenges and realize the goals of network-centric warfare.

Client-server approaches such as service-oriented architectures (SOAs) are commonly adopted as the basis to realize applications and services in military systems running on higher-echelon command and control networks. In an unreliable bandwidth-constrained tactical environment that is subject to network partitioning, client-server architectures can introduce centralized points of failure and performance bottlenecks. Moreover, the unicast point-to-point connections result in excessive bandwidth consumption when data is sent to large numbers of clients. Peer-to-peer (P2P) approaches do not rely on designated server nodes that must be reachable and therefore can continue to (partially) function in a partitioned network. Also, P2P systems can make use of multicast and other advanced data distribution schemes that minimize the transmission of redundant information. Finally, because communications do not need to be routed through central servers, P2P technologies can leverage the fact that many applications place greater importance on communication between nearby nodes as opposed to distant nodes. These characteristics make P2P architectures a better fit for tactical networks than traditional client-server designs.

This article focuses on P2P systems design as applied to middleware and application development for tactical networks. P2P systems are usually classified as structured or unstructured network architectures [1]. In structured P2P sys-

0163-6804/10/$25.00 © 2010 IEEE

tems, nodes cooperate to maintain a distributed database that contains information about the location of resources (i.e., nodes, files, services, etc.) within the network. In unstructured P2P systems, instead, nodes have to advertise resource location or discover resources by broadcasting queries over the network. A resource location database reduces the number of messages required for resource discovery and facilitates the discovery of non-replicated resources. However, in highly dynamic network environments such as tactical networks, the bandwidth and computational overhead required to maintain the distributed database in a consistent state typically outweighs its benefits. On the other hand, unstructured P2P systems have to deal with scalability issues related to the broadcast of queries across the network, but their inherent resilience to node churn makes them better suited to mobile ad hoc and tactical environments.

Based on experiences from experiments and exercises, we present scenarios, such as Blue Force Tracking (BFT), remote unmanned vehicle control, and sensor data mining/fusion, that benefit from P2P approaches. These observations are subsequently developed into specific technical requirements. The observations and requirements that are presented take into account the perspectives of the Air Force, the Army, and the Navy and Marine Corps. While the focus is on P2P aspects, we briefly discuss the intersection of P2P with client/server approaches and SOAs.

Finally, the article presents a case study using the Agile Computing Middleware, a software framework that enables the realization of applications and services for tactical edge networks. We discuss the Group Manager discovery and the DisService information dissemination components, which support P2P applications. We provide experimental results that demonstrate the effectiveness of the P2P approach. The Agile Computing Middleware, initially developed to address the requirements of Army battlefield environments, has been extended to support Air Force requirements, and is currently being extended to satisfy Navy and Marine Corps requirements.

Peer-to-Peer Communication Scenarios in Tactical Environments

During the last eight years, we have collectively participated in a number of tests, experiments, demonstrations, and exercises involving tactical environments. These include the Horizontal Fusion Quantum Leap experiments in 2003 and 2004 (QL-1 and QL-2), the C4ISR[1] On the Move (C4ISR OTM) experiments in 2006, 2007, and 2008, the Joint Forces Command (JFCOM) Empire Challenge in 2009 (EC-2009), JFCOM BoldQuest (BQ) exercises in 2007 and 2009, the Office of Naval Research (ONR) ISR-C2[2] experiment in 2009, the Defense Advanced Research Projects Agency (DARPA)/SSC Pacific SIE-DTN[3] field test in 2010, and the Joint Expeditionary Forces Experiment (JEFX) in 2008, 2009, and 2010. Figure 1 shows the typical nature of the networks we have observed in these exercises — clusters of edge nodes that are interconnected via reach-back links. In particular, we note that the size of the individual clusters range from 10 to 20 nodes. P2P approaches are most applicable within these individual clusters, with other communication protocols being more appropriate for the reach-back links [2]. In this section we distill and summarize the scenarios and discuss the benefits of P2P approaches.

Figure 1. *Tactical networking environment with peer-to-peer clusters.*

Blue Force Tracking

Perhaps the most fundamental need for tactical users is Blue Force Tracking (BFT) — applications that provide situational awareness information regarding the presence and location of friendly forces. BFT is critical to avoid friendly fire accidents. The dynamic, geographically sensitive nature of most tactical situations makes BFT applications well suited to P2P architectures. Even in an unreliable communications environment that prevents connections to a central BFT server, nodes in close proximity are likely to still be able to communicate with each other and share critical location information. P2P-based BFT systems enable this mode of operation. Furthermore, P2P approaches can be more bandwidth-efficient — a critical resource in tactical networks.

BFT also provides a good example of the proximity/precision correlation. When users are operating in close proximity, it is more important that their relative positioning be exchanged reliably, quickly, and frequently between each other. As they are located farther apart, the information may be aggregated and exchanged less frequently. This requirement occurs in other application scenarios as well, and advanced P2P architectures can provide this capability.

Resource Discovery and Awareness

Another fundamental need of tactical users and applications is discovery and awareness of available resources or assets. Resources can include deployed sensors, manned vehicles, autonomous air and ground vehicles, and even other users themselves. Resources may pertain to the platform itself or to services provided by a platform.

[1] *C4ISR is an acronym for Command, Control, Communications, and Computers for Intelligence, Surveillance and Reconnaissance, and usually refers to the set of systems and networks that enable net-centric operations.*

[2] *ISR-C2 is an acronym for Intelligence, Surveillance, and Reconnaissance for Command and Control — a concept similar to C4ISR.*

[3] *SIE-DTN is the Service Interoperability Environment with Disruption Tolerant Networking — a field test conducted by the Navy's SPAWAR Systems Center Pacific and DARPA.*

Unmanned ground and air vehicles are becoming more common in military operations. As the number and type of unmanned assets increase, the importance of distributed P2P technologies for discovering and controlling these assets increases as well.

For example, an unmanned air vehicle (UAV) may provide a *video feed* service. Resource discovery is an essential prerequisite for many other operations, such as tasking of autonomous vehicles, gathering data from sensors, and communicating with other users. For example, at QL-2, C4ISR OTM, and ISR-C2, dismounted soldiers were discovering available robotic assets, deployed perimeter sensors, as well as other soldiers. In the JEFX and BoldQuest exercises, a user on the ground was discovering airborne assets along with data feeds being generated by the airborne assets.

An equally important aspect that complements resource discovery is continued awareness of the resource. For example, consider a user that has tasked perimeter trip-wire sensors to notify the user of any incursions, but never receives any notifications from the sensors. This may simply be because the sensors have not detected anything. It could also be because the sensors are malfunctioning or out of communications range. There is a significant difference between these two scenarios, and the user must be able to differentiate between them. Systems must provide status information regarding assets and the communications links between them so that users can be confident that the system is functioning properly.

Client-server discovery systems generally rely on one or more centralized registries, which become critical points of failure in the unreliable tactical network environment. With a centralized system, two nodes that are able to communicate with each other, but not with the central registry, will not be able to discover each other. There is also the initial problem of locating the registry, usually solved by statically configuring each client with the address of the registry server. If the registry must be moved, or if a client would be best served by using another registry server, the client's configuration must be changed. This is not always possible. For these reasons, resource discovery and awareness services in tactical edge networks are best provided through P2P approaches.

Geographically Directed/Constrained Operations

In military applications, many requests for information or assets are constrained by geography. For example, a trip-wire perimeter sensor that is triggered may cause a user (or an automated process) to look for a visual sensor that can provide imagery in the vicinity of the sensor's location. Similarly, a user may wish to find the closest available robot or UAV to a building to be examined or monitored. The geographical constraints may apply to the node itself or, in the case of a sensor node, to the sensor's area of coverage. A related feature is a geographic subscription to information. For example, the RouteScout application for JEFX 2010 allows the mission path to be used as a geographically constrained search for information along the path. For the Marine Corps, we have developed a mechanism to model the mission characteristics, including the set of possible paths, which is subsequently used by peers to push relevant information to the users' nodes.

Sensor Data Exchange

Unattended ground sensor systems currently fielded by the military generally rely on satellite communications for data transmission. This approach has a number of drawbacks. Satellite communications require a transceiver with a clear view of the sky, and low-power satellite modems suitable for embedded sensor applications are very slow (on the order of a few kilobits per second). By using P2P models combined with mobile ad hoc network technologies, sensors can use high-speed short-range radios to exchange and exfiltrate data to other sensors and nearby units. With a high-speed P2P link, a sensor can send high-resolution imagery and motion video that would be impractical to transmit over a low-speed satellite link. Additionally, P2P technologies can be used to build store-and-forward sensor networks, where a sensor that is not in communications range of a data consumer will locally store detection data and forward it to consumers as they come in range. These P2P approaches can work side by side with existing satellite reach-back systems, enabling advanced applications without interfering with current infrastructure and doctrine.

Unmanned Vehicle Operations

Unmanned ground and air vehicles are becoming increasingly common in military operations. As the number and type of unmanned assets increase, the importance of distributed P2P technologies for discovering and controlling these assets increases as well. For example, if a combat unit has a number of unmanned assets and a number of controllers, a P2P discovery mechanism is required so that any controller can discover and connect to any asset, as it may not be known in advance which controllers will be used with a particular asset. P2P systems are also useful for integrating unmanned assets into the BFT systems described earlier. Additionally, a major focus of current unmanned systems research is developing collaborative behaviors where multiple assets cooperate to accomplish a task. To execute these behaviors, assets must be able to discover other nearby assets and then exchange state information as the task is being performed. The proximity/precision correlation applies here as well. P2P systems are useful for accomplishing these goals in a decentralized manner.

Sensor Data Mining, Data Fusion, and Distributed Cross-Cuing

Multiple factors are driving the battlefield to be increasingly covered with distributed sensors that are capable of generating large quantities of data. Data mining, data fusion, and cross-cuing are all mechanisms to reduce data overload by aggregating, correlating, transcoding, and filtering raw data into useful information. When such sensor operations are in close geographical proximity, P2P approaches are more effective because they can exploit multicasting over short-range wireless links that require less power and provide higher bandwidth.

P2P systems can also be used to provide automated, dynamic, self-configuring cross-cuing capabilities between unattended sensors and other assets. For example, a small sensor that contains a motion detector or vibration sensor but no camera can be used to trigger a different camera-equipped sensor to take an image or record a video, cuing the camera as necessary. Additionally, if a UAV or other similar asset happens to be nearby, the imager on the UAV could also be cued and tasked.

SECURE CHAT, MEDIA EXCHANGE, AND SOCIAL NETWORKING

Military operators have embraced aspects of social networking to improve effectiveness of mission execution. Tools include instant messaging, marking up and sharing multimedia objects such as pictures and video, collaborative map annotation, and other forms of communication. Soldiers use these tools to maintain situation awareness, coordinate, and exchange information in an effective and clandestine manner. They share useful tidbits of intelligence information, contact subject matter experts to ask for opinions, and dynamically replan their strategy as necessary. Many of these operations involve communicating with peers who are on the same mission or in their physical proximity. Therefore, P2P systems are the best approach to supporting these scenarios.

TECHNICAL REQUIREMENTS FOR PEER-TO-PEER SYSTEMS

In this section we develop technical requirements for P2P systems that can address the types of scenarios described in the previous section. We have addressed many of these requirements in developing the Agile Computing Middleware, described later in this article.

AUTOMATIC CONFIGURATION

The dynamic nature and high operational tempo of most tactical environments does not allow time for system configuration. Therefore, wherever possible, systems should automatically configure themselves within the network and provide the capabilities necessary for the users. P2P systems naturally address this requirement by removing the need to configure clients with information about servers that need to be accessed. Systems should also minimize requiring users to configure network settings and specify IP addresses of other information producers and consumers. Gateways and other structural aspects of the P2P network should be automatically discovered as nodes join and leave networks.

BANDWIDTH-EFFICIENT PEER DISCOVERY

Discovery of other peers is the most fundamental requirement for P2P systems, as it is a prerequisite for most other operations. Peer discovery is closely related to service advertisement and search. Typically, peer discovery applies at the platform or node level, whereas advertisement and search apply at the service level. For example, at C4ISR OTM, each ground robot advertised a Robot Agent service that was discovered by operator control units (OCUs) used to tele-operate the robots. Similarly, high-mobility multipurpose wheeled vehicles (HMMVWs) advertised numerous services such as Blue Force Tracking, data fusion, and language translation services that were looked up by other peer users.

Peer discovery must be bandwidth-efficient and adapt to different types and qualities of network links. Ground-based military operations on urban terrain (MOUT) scenarios such as QL and C4ISR OTM use wideband radios that provide reasonable bandwidth on the order of 1 Mb/s. However, the airborne networking environment is more constrained. For example, during BQ 2007 and JEFX 2007, long-range UHF links were utilized between the aircraft and ground nodes, which provided a bandwidth of 10 kb/s or less. In BQ 2009 and JEFX 2009 the move to new prototype radios (e.g., DARPA Quint Networking Technology) provided much higher bandwidth. However, these broadband radios and waveforms provide significantly shorter range, and certain operations still require low-bandwidth UHF links. Therefore, it is critical that peer discovery be adaptive in terms of bandwidth utilization.

There are several commercial and open source efforts that address peer discovery. Examples include JXTA [3], Zeroconf [4] implementations such as Apple's Bonjour and Avahi's mDNS/DNS-SD, XMPP [5], Session Initiation Protocol (SIP) [6], and uPnP [7]. Our experience with these components shows that they do not perform well on tactical networks. These components have been designed for infrastructure networks that do not have the bandwidth and reliability challenges of tactical edge networks. We compare the performance of JXTA with the Agile Computing discovery mechanism in the section on experimental results.

PEER-LOSS DISCOVERY

Peer loss discovery is important for providing situational awareness to tactical users. In client-server environments, a user often detects the loss of the other entity simply by detecting the loss of the connection. However, in P2P environments where the communication tends to be point-to-multipoint (e.g., using multicast), peer loss discovery is more difficult. Peer loss discovery directly contributes to increased awareness of surrounding resources.

DYNAMIC INFORMATION CHANNELING ACCORDING TO TOPOLOGY AND RESOURCE AVAILABILITY

To support data mining and data fusion, as well as geographically directed resource discovery, it is necessary to provide mechanisms to convey information through a subset of the network dynamically determined by topological or resource availability constraints and manipulate data as it is transferred between peers. This is a challenging task that requires monitoring of the network topology as well as dynamically discovering and allocating computational resources along the communication path. In addition, to minimize the overhead of on-demand realloca-

Military operators have embraced aspects of social networking to improve effectiveness of mission execution. Soldiers use these tools to maintain situation awareness, coordinate, and exchange information in an effective and clandestine manner.

The dynamic nature of tactical networks requires disruption tolerant approaches to information dissemination. Sometimes, information must be delivered to nodes that periodically disconnect from the rest of the network, requiring reliability mechanisms such as caching and periodic retransmission of important data.

tion protocols, proactive communication path rerouting and computational resource reassignment should be implemented based on predictions about future user requirements.

Adaptive Dissemination

Support for information dissemination is essential in tactical networks. To minimize communication overhead, different protocols and algorithms can be used depending on the set of receivers and the requirements of the application. In our experience tactical applications present multiple patterns of data dissemination. BFT is transmitted from each node to every other node in a many-to-many pattern. Sensor fusion requires many nodes (sensors) to transfer data to one node (fusion node) and then onto some consumers in a many-to-one-to-few pattern. For maximum efficiency, the data dissemination system must adapt its strategy based on the dissemination pattern. For instance, information that must be delivered to most of the nodes in the network is best transmitted as an all-nodes broadcast. On the other hand, information that must be delivered only to a small set of nodes, especially when in close proximity, should adopt different mechanisms. The same considerations apply to the case of multiple information sources.

Adaptation should exploit prioritization, differential update rates, reliability, and sequencing requirements as well. As discussed earlier, there is often a direct relationship between physical proximity and the requirements for precision in the information exchange between nodes. An adaptive dissemination system can exploit this to reduce bandwidth utilization. Similarly, not all consumers require the cost associated with reliable and/or sequenced delivery of information. The dissemination system should be flexible in supporting these properties when required and saving bandwidth when they are not necessary.

Disruption Tolerant and Robust Information Dissemination

The dynamic nature of tactical networks requires disruption-tolerant approaches to information dissemination. Sometimes, information must be delivered to nodes that periodically disconnect from the rest of the network, requiring reliability mechanisms such as caching and periodic retransmission of important (non-obsolete) data. The SIE-DTN experimentation has shown that disruption tolerance significantly improves the ability for nodes to receive data under intermittent connectivity conditions.

Systems should also provide mechanisms for robust information delivery. Techniques such as forward error correction, scattering (transmitting different portions of an object to receiving nodes and letting them recompose the original object), and layered coding (the decomposition of an object into smaller and lower-fidelity usable parts that can be rejoined by the receiver) should be adopted to improve the availability of large information objects.

Finally, systems should exploit common patterns in node mobility and information to improve the timeliness and availability of information dissemination. This requires learning mechanisms that can process node mobility and data/service usage information, identify common behaviors, and produce reliable forecasts that can be used as the basis for decision making in information caching and routing.

Tracking Information Pedigree

Tracking pedigree of information entails securely maintaining meta-information regarding the source of the information, the time when the information was originated or generated, and the path the information has traversed before reaching a consumer. Information pedigree is often essential to decision making. While this requirement applies to many types of information systems, we highlight it here because P2P systems make it more challenging to track information pedigree. For example, with a client-server system, if the server is a trusted entity, the server can often assign or validate some aspects of the pedigree of the information provided by a client, and pass that on to other clients. In a P2P system, if some of the peer nodes are not trusted, additional mechanisms may be necessary to either ensure that the information satisfies necessary security requirements or adjust the pedigree appropriately.

Policy-Based Control over Information Exchange

Another requirement relates to control over the nature and destination of information as it is managed and disseminated in P2P systems. The frequent occurrence of joint and coalition operations, and the associated restrictions on information sharing imply that P2P systems should address this requirement. Client-server approaches are more amenable to such control over information sharing, as the servers can act as gatekeepers of the information and enforce any policies on sharing. This problem must be addressed within P2P systems if they are to be pervasive and span multiple administrative domains at the tactical edge. An alternative is a hybrid approach that uses P2P within a single administrative domain and trusted server gateways to span domains.

Integration with Service-Oriented Architectures

SOAs have been widely adopted by military systems running on higher-echelon command and control (C2) networks. SOA-based approaches allow for extensive service reuse and the integration of heterogeneous services, with significant savings in development costs and time for building distributed applications. SOAs also allow the dynamic (re)composition of services at run-time, thus enabling the ad-hoc realization of complex and adaptive distributed applications.

Because of these advantages, SOAs are often proposed in tactical networks despite the challenges these architectures face in that environment [8]. Traditional SOA implementations are based on centralized service directories and strong assumptions about relatively static network topologies. Therefore, they are susceptible

Communications | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

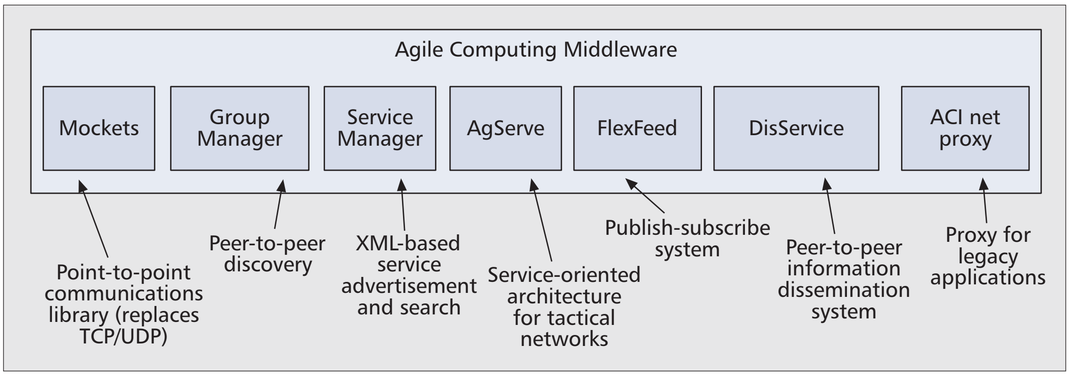


Figure 2. *Components of the agile computing middleware.*

to poor scalability (also caused by their RPC-based paradigm that hinders the adoption of caching) and are limited by relatively high computational and bandwidth requirements and lack of support for service migration.

These limitations call for tactical network-specific SOA implementations as well as ad hoc integration solutions between P2P and SOA middlewares [9, 10]. As a result, P2P systems for tactical applications should include an interoperability layer that enables resource discovery and service accessibility across SOA and P2P architectures.

AGILE COMPUTING MIDDLEWARE

The Agile Computing Middleware [11] has been developed over the course of the last nine years to address many of the requirements enumerated in the previous section. Figure 2 shows the key components of the middleware. In particular, this section presents two components — Group Manager and DisService — that realize P2P discovery and information dissemination. The ACINetProxy, which helps integration with legacy systems, is also briefly described. The next section presents experimental results that demonstrate the effectiveness of the P2P approach in Group Manager and DisService.

GROUP MANAGER

The Group Manager component supports peer discovery and has been optimized to be bandwidth-efficient for tactical edge networks. It supports a flexible combination of proactive advertisement and reactive search. When advertisement is activated, the frequency of advertisements depends on the node movement and the churn rate of the network. Hence, a fast-moving entity such as an aircraft would advertise more frequently, whereas a stationary ground sensor advertises slowly. Nodes can also control the *strength* of their advertisement, which is defined in terms of a metric (number of hops, distance, or link quality). The frequency and strength of advertisements provide a trade-off between discovery speed, discovery range, and bandwidth utilization. Providing control over this behavior allows nodes to select the best possible combination to suit their needs. Ground nodes that need to maintain radio silence could choose instead not to advertise at all, thereby preventing their discovery.

Discovery occurs within the context of a group, a concept that allows partitioning of network nodes into different sets. A group is identified by a name with a simple hierarchical notation, such as `mil.army.arl.hf2004.ugvs`. Groups are a convenient abstraction to aggregate resources based on mission or ownership (e.g., country or branch of the armed forces). Nodes enable discovery among themselves by simply joining the same group. For example, all the nodes that are assigned to a specific mission could join a common group, thereby allowing them to be discovered by each other. Nodes are allowed to join multiple overlapping groups, and to advertise different capabilities within each group. Therefore, a node may advertise service X to other nodes participating in the same mission and service Y to other nodes that are on a different mission but belong to the same country.

Group Manager also supports private groups with restricted access and encrypted communications. Private group security leverages a preshared passphrase, often proposed in tactical military applications, where direct access to the trusted authority (TOC) is not available for encryption key generation and key exchange. Private groups make it very simple for dynamically formed ad hoc teams to verbally exchange a passphrase that may then be used to restrict access to the group. We have used this capability to easily build secure P2P chat applications that dynamically manage chat members based on their mission, role, or any other attribute. Note that encryption in private group communications does not replace standard cryptographic measures provided by most tactical radios; instead, it represents an additional security layer on top of them.

Group Manager handles discovery and maintains group membership in a completely decentralized manner. This design choice implies that there is no attempt to maintain a consistent group view of all the nodes that are members of a group, as it would require an excessive amount of bandwidth. Distributed hash table (DHT) approaches were also discarded for their high bandwidth requirements and given the high churn rate of tactical networks.

DISSERVICE

DisService is the P2P information dissemination component of the Agile Computing Middleware. In DisService information is pushed and received

SOAs have been widely adopted by military systems running on higher-echelon command and control networks. SOA-based approaches allow for extensive service reuse and the integration of heterogeneous services, with significant savings in development costs and time for building distributed applications.

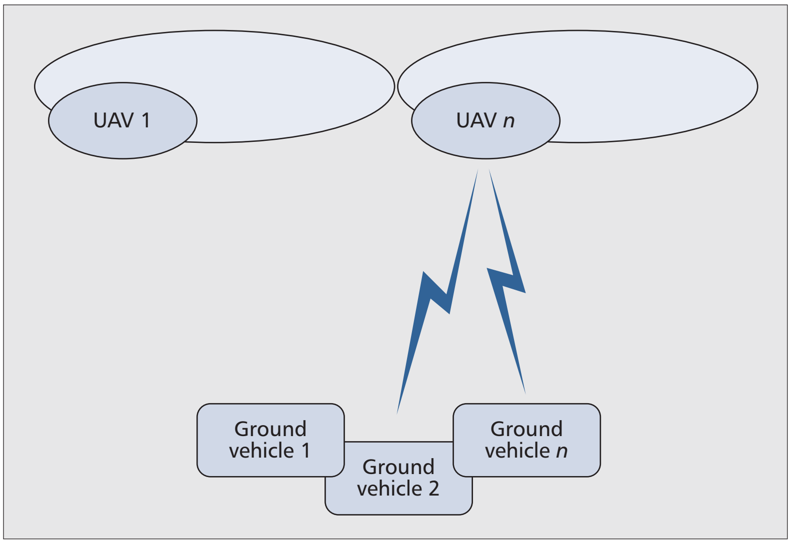


Figure 3. *Peer discovery experiment scenario.*

in the context of a group, using the same abstraction proposed by Group Manager. DisService is disruption tolerant and realizes several novel architectural approaches to improve dissemination and availability of data.

One of the fundamental assumptions in DisService is that in a P2P wireless environment, the cost to broadcast or multicast a message to a neighbor is the same as with a unicast. Therefore, DisService always prefers broadcast or multicast communications, which enable all other direct neighbors to also receive and archive the data.[4] This feature, called *opportunistic listening*, significantly increases the availability of data, especially in a tactical network where nodes and links are unstable. The next section presents experimental results that show the benefits of this feature.

Another DisService feature is flexible support for sequential and reliable message delivery. When a node subscribes to a group, it can choose whether messages should be delivered sequentially and/or reliably. Typically, sequencing increases latency and reliability increases bandwidth utilization. Therefore, DisService allows applications to make trade-offs that improve their efficiency. For example, a node could prefer unreliable but sequenced delivery for BFT information, while requiring reliable unsequenced delivery for sensor reports.

The third unique feature of DisService is that the burden of reliability, if desired, is placed on the receiver as opposed to the sender. In both the TCP and DTN models, the sender has the responsibility to retransmit the data until the receiver, or an intermediary in the case of DTNs, has acknowledged receipt. DisService instead adopts a different model where the receiver, if so desired, requests missing messages (or parts of messages) from other nodes until they have been received. The motivations behind the adoption of this model are the point-to-multipoint nature of DisService and the observation that different peers that are recipients of the same data may have different requirements for reliability. Furthermore, this approach exploits the opportunistic listening capability of DisService, where a node looking for the missing data may be able to obtain it from any other peer as well. The conventional model, where the sender wishes to ensure that a set of target nodes receive data, is also available.

DisService also supports multiple dissemination algorithms, including reliable flooding, probabilistic (epidemic) protocols, and heuristic protocols, which can be selected based on the type of information being disseminated. The specific dissemination algorithm chosen depends on the number of subscribers to a group, the nature of their subscription (reliable or not), and the relative priority. Reliable flooding is an expensive algorithm, but is appropriate for high-priority messages that must be delivered reliably. Probabilistic protocols use attributes such as the number of neighbors of nodes to determine the dissemination paths. Heuristic protocols further enhance probabilistic approaches by exploiting domain-specific knowledge to guide the probabilistic dissemination. One example of a heuristic is favoring the use of a UAV as a good intermediate node for information propagation, given the increased reach and visibility of a UAV. Another example is a node acting as a relay to a neighbor that has only one other neighbor (which implies that the first node is the only way for the second node to receive any information).

Finally, DisService provides a different approach to handle dissemination of large messages, such as multimedia objects (pictures, video, etc.). In these cases only the metadata describing the large message is disseminated to the subscribers. Each subscriber, upon evaluation of the metadata, may subsequently choose to retrieve the complete message. DisService supports two mechanisms to handle the actual dissemination of large messages. In the simple case the message is transferred as a whole from one of the source nodes to the destination node, which can now act as an additional source node. In addition, DisService supports a second approach which fragments the large message into chunks that are scattered across the network and replicated on a bandwidth-available basis. When a node wants to retrieve a message, it queries its peers to find the chunks, retrieves them, and reassembles the original message. Furthermore, where possible, the chunks are created using a data-type-specific layered encoding algorithm that allows each chunk to be independently usable. For example, large images are broken up into multiple complete but lower-resolution images that are individually usable but, when combined, recreate the original high-resolution image.

Integration

The best approach to leveraging the capabilities of the Agile Computing Middleware is to directly take advantage of the application programming interfaces (APIs) of the various components. However, we recognize that this is not always possible, especially with legacy applications. The middleware includes the ACINet-Proxy component, which provides a transparent mechanism to capture all the outgoing traffic generated by legacy TCP- and UDP-based appli-

[4] *Note that the cryptographic and security features included in tactical radios will ensure that nodes that should not be receiving the data will not do so. If further restricting is necessary, the notion of private groups, as described for the Group Manager, can be applied to DisService as well.*

IEEE Communications | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | qMags

cations and redirect it using middleware components. This capability is particularly useful for redirecting TCP connections to use Mockets [2] and UDP multicast to DisService. ACINetProxy is configurable via policies to support deployment on a wide range of scenarios and permits fine-grained performance tuning (e.g., prioritization and shaping of traffic).

EXPERIMENTAL RESULTS

This section presents results from two experiments that evaluate the performance of the Group Manager and DisService components.

PEER DISCOVERY RESULTS

In this first experiment we compare the performance of Group Manager with the JXTA middleware for P2P computing. In particular, we compare the discovery aspects of Group Manager and JXTA, and measure the relative bandwidth utilization. JXTA was chosen because its discovery mechanism is utilized by the System of Systems Common Operating Environment (SOSCOE), part of the U.S. Army Future Combat Systems and other programs. The scenario, shown in Fig. 3, consists of two to four UAVs and four to 18 ground vehicles (GVs). The connectivity between the GVs and the UAVs varies based on the position of the UAVs. In this particular experiment the GVs were attempting to discover a service advertised by the UAVs. The connectivity between the nodes was emulated using an enhanced version of the Naval Research Laboratory's Mobile Ad Hoc Network Emulator [12].

The results are shown in Table 1. While the peer discovery was completed successfully with both components, the Group Manager used substantially less bandwidth than JXTA. As the results show, JXTA used between 3.58 to 7.77 times more bandwidth than the Group Manager in default mode. The Group Manager also supports an enhanced mode that uses periodic pings to check whether nodes are still reachable, thereby addressing the peer loss discovery requirement. Even with this enhancement (which JXTA does not support), the Group Manager is more efficient, with JXTA using between 1.58 and 2.29 times more bandwidth. Also note that in this particular experiment, JXTA was operating in edge mode, where it does not use a DHT. In rendezvous mode with a DHT, JXTA used significantly more bandwidth.

DISSERVICE OPPORTUNISTIC LISTENING RESULTS

This second experiment demonstrates opportunistic listening, one of the unique features of DisService. Opportunistic listening exploits peer nodes as surrogate listeners to capture information that a subset of the peer nodes are trying to receive. The scenario is shown in Fig. 4, and consists of a set of 10 ground nodes of which two nodes are trying to receive data being multicast by a UAV. The UAV makes one pass through the area as it transmits messages to the ground nodes. The bandwidth of the airborne network between the UAV and the ground nodes is assumed to be 230 kb/s, and the bandwidth of the P2P network between the ground nodes is assumed to be 1 Mb/s. In the baseline case, we report on the percentage of messages received via UDP multicast by the two interested ground nodes. In the DisService case, the ground nodes exchange missing data between themselves to maximize the number of messages correctly received by the two ground nodes.

The results, with varying message sizes and network reliability, are shown in Table 2 and illustrate the benefit of opportunistic listening. For example, with a UAV link reliability of 70–80 percent, only 20 percent of messages of size 7 kbytes are successfully delivered due to IP fragments being lost. With DisService, the target nodes are able to obtain these missing fragments from other peers and reassemble the messages intact, thereby improving performance significantly. Even with very poor UAV link reliability of 30–50 percent, DisService is able to receive a

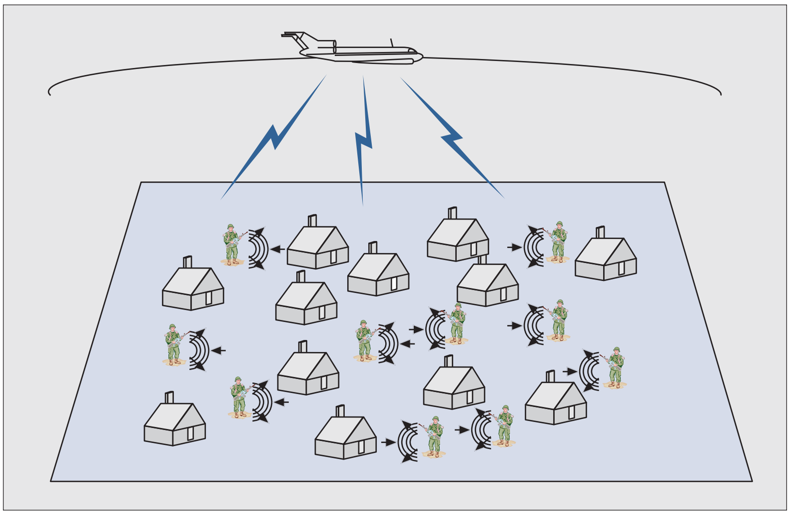


Figure 4. *Opportunistic listening experiment scenario.*

Number of nodes		Group Manager		JXTA
		(Default)	(With pings)	
UAVs	GVs	Bandwidth utilization (bytes/s)		
2	1	83	372	645
2	2	164	672	1106
2	3	247	866	1536
2	4	329	1217	2012
4	6	828	1856	2963
4	10	1350	3061	4849
4	14	1146	2937	6713
4	18	1451	5471	8658

Table 1. *Group Manager performance (bandwidth) for peer discovery.*

UAV link reliability	Ground net reliability	Msg size (bytes)	Msgs sent	Msgs received		Success rate		DisService P2P overhead (bytes/Msg)
				Multicast	DisService	Multicast	DisService	
70% to 80%	80%	1024	1390	820.6	1357.2	59.06%	97.68%	101
		7168	203	31.9	202.5	15.75%	100.00%	44,518
		15,360	96	4.5	96.0	4.69%	100.00%	69,938
		35,840	41	0.0	41.0	0.00%	100.00%	174,163
30% to 50%	80%	1024	1377	590.6	1367.3	42.89%	99.30%	14,380
		7168	200	2.0	193.0	1.00%	96.50%	62,139
		15,360	91	0.0	89.3	0.00%	98.13%	135,077
		35,840	41	0.0	34.0	0.00%	82.93%	368,985

Table 2. *DisService opportunistic listening performance.*

large percentage of the messages. We also report on the P2P traffic exchanged between the ground nodes as they reconstruct the messages, which could be regarded as overhead. As the results show, the overhead per message is quite high, especially as the size of the messages increases and the reliability decreases. However, this communication takes place over the P2P ground network, which has higher capacity and is less congested. Furthermore, we expect to optimize this bandwidth utilization in the near future.

CONCLUSIONS

Our experiences have led us to study how to realize robust and efficient applications and services for tactical edge networks. We have observed that the nature and characteristics of tactical networks, combined with the requirements of network-centric warfare present a compelling case for P2P systems. We then distilled these experiences into a set of technical requirements that must be addressed to develop effective P2P systems for tactical environments. Following these guidelines, we have developed the Group Manager and DisService components of the Agile Computing Middleware, which realize P2P discovery and information dissemination. We have tested those components in the context of several experiments and exercises that allowed us to verify the soundness as well as the effectiveness of the P2P approach. We continue to enhance the Agile Computing Middleware and conduct further experiments. Future work includes optimizing DisService, comparing with reliable multicast approaches, as well as comparing opportunistic listening with forward error correction. We hope that this article will motivate and guide the development of future P2P systems for tactical networks.

ACKNOWLEDGMENTS

This research was sponsored in part by the U.S. Army Research Laboratory under Cooperative Agreement W911NF-04-2-0013, by the U.S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0009, by the Air Force Research Laboratory under Cooperative Agreement FA8750-06-2-0064, and by the Office of Naval Research under grant N000140910012. Figure 1 is courtesy of William Howell, Florida Institute for Human and Machine Cognition.

REFERENCES

[1] X. Shen *et al.*, *Handbook of Peer-to-Peer Networking*, Springer, 2009.
[2] N. Suri *et al.*, "Communications Middleware for Tactical Environments: Observations, Experiences, and Lessons Learned," *IEEE Commun. Mag.*, vol. 47, no. 10, Oct. 2009, pp. 56–63.
[3] B. Traversat *et al.*, "Project JXTA 2.0 Super-Peer Virtual Network," Sun Microsystems, Inc. tech. rep., 2003; http://research.sun.com/spotlight/misc/jxta.pdf
[4] IETF Zeroconf Working Group; http://www.zeroconf.org
[5] XMPP Standards Foundation; http://xmpp.org/
[6] J. Rosenberg *et al.*, "SIP: Session Initiation Protocol," RFC 3261, June 2002; http://www.rfc-editor.org/rfc/rfc3261.txt
[7] The UPnP Forum; http://www.upnp.org
[8] K. Lund *et al.*, "Using Web Services to Realize Service Oriented Architecture in Military Communication Networks," *IEEE Commun. Mag.*, vol. 45, no. 10, Oct. 2007, pp. 47–53.
[9] D. Galatopoullos, D. Kalofonos, and E. Manolakos, "A P2P SOA Enabling Group Collaboration through Service Composition," *Proc. ICPS '08*, Sorrento, Italy, July 6–10, 2008.
[10] N. Suri, "Dynamic Service-Oriented Architectures for Tactical Edge Networks," *Proc. 4th Wksp. Emerging Web Services Tech. '09*, Nov. 2009, pp. 3–10.
[11] N. Suri *et al.*, "An Adaptive and Efficient Peer-to-Peer Service-Oriented Architecture for MANET Environments with Agile Computing," *Proc. 2nd IEEE ACNM '08*, Salvador de Bahia, Brazil, Apr. 7–11, 2008, pp. 364–71.
[12] Mobile Ad Hoc Network Emulator (MANE); http://cs.itd.nrl.navy.mil/work/mane/index.php

BIOGRAPHIES

NIRANJAN SURI [M] (nsuri@ihmc.us) is a research scientist at the Florida Institute for Human and Machine Cognition (IHMC). He received his Ph.D. in computer science from Lancaster University, England, and his M.Sc. and B.Sc. in computer science from the University of West Florida, Pensacola. His current research activity is focused on the notion of agile computing, which supports the opportunistic discovery and exploitation of resources in highly dynamic networked environments. His other research interests include coordination algorithms, distributed systems, net-

working, communication protocols, virtual machines, and software agents. He has authored or co-authored over 60 papers, has been on the technical program committees of several international conferences, and has been a reviewer for the National Science Foundation as well as several international journals.

GIACOMO BENINCASA (gbenincasa@ihmc.us) is a research associate at the Florida Institute of Human and Machine Cognition, working on information dissemination in tactical networks and on biologically-inspired security infrastructures. He holds a B.Sc. and an M.Sc. in computer engineering from the University of Modena and Reggio Emilia, Italy.

MAURO TORTONESI (mauro.tortonesi@unife.it) received his Laurea degree in electronic engineering and his Ph.D. in computer science engineering from the University of Ferrara, Italy. From 2004 to 2005 he was a research associate at the Florida Institute of Human and Machine Cognition. Currently, he is an assistant professor in the Engineering Department of the University of Ferrara, Italy. His research interests include distributed and mobile computing, QoS management, business-driven IT management, and industrial automation.

CESARE STEFANELLI [M] (cesare.stefanelli@unife.it) received his Laurea degree in electronic engineering and his Ph.D. in computer science engineering from the University of Bologna, Italy. He is currently a professor of computer science engineering in the Engineering Department of the University of Ferrara. His research interests include distributed and mobile computing, adaptive and distributed multimedia systems, network and systems management, and network security. He is a member of AICA.

JESSE KOVACH (jkovach@arl.army.mil) is a computer engineer with the Battlefield Information Processing Branch of the U.S. Army Research Laboratory, specializing in information dissemination in tactical networks, and sensor and robotic systems integration. He designs, develops, and tests prototype systems that incorporate advanced concepts developed within the laboratory as well as elsewhere in the defense R&D community. His work has been incorporated into a number of fielded systems and has been used in multiple high-visibility demonstrations, including C4ISR On the Move and Empire Challenge. He holds a Bachelor's degree in computer engineering from the University of Maryland.

ROBERT WINKLER (winkler@arl.army.mil) has been a computer engineer with the U.S. Army Research Laboratories since 1987. He graduated magna cum laude in computer science from the University of Maryland at College Park in 1988 and received his Master of Science degree in computer science from Johns Hopkins University in 1995. His current applied research interests include databases, artificial intelligence, and data mining.

RALPH KOHLER, JR. [F] (Ralph.Kohler@rl.af.mil) is a principal engineer at the U.S. Air Force Research Laboratory (AFRL), where his primary research interests are the nexus of information management, wireless and tactical networks, and the capabilities that combining the two make possible. He received his Master's degree from Syracuse University and a member of the AIAA.

JAMES HANNA (james.hanna@rl.af.mil) is a senior research engineer with the Enterprise Information Management Branch in the Information Directorate at AFRL. He has been a computer engineer for AFRL for the past 21 years and has over 23 years experience developing software applications. For the past two years he has led the development of Phoenix, an SOA-based information management infrastructure. He is also the lead in-house engineer researching issues related to distributed survivable information management and advanced architectures to address the complex challenges of deploying information management in tactical environments. He has a B.Sc. in computer science from the State University of New York Institute of Technology, Utica.

LOUIS POCHET (loupochet@gmail.com) is the senior integration engineer in the Combat Support Office of the U.S Air Force Reserve and of a small technology startup, GMECI. Before entering the Reserves, he developed applications, protocols, and waveforms for the U.S. Department of Defense as an active duty research engineer, deploying to theaters of operation multiple times. He has focused on network discovery and configuration over existing and experimental tactical IP networks, including QNT and ANW2, for the last five years.

SCOTT WATSON (scott.c.watson@navy.mil) is a senior systems architect with the Composeable Services Branch in the Command and Control Department of the Space and Naval Warfare Systems Center — Pacific. He has over 10 years' experience designing software systems and applications. He is currently the principal investigator for the Mobile Modular Command and Control for Company level operations (M2C3) project. He has a B.S. in computer science from San Diego State University, California. He recently received the Navy Civilian Meritorious Service award for work in the area of disruption-tolerant networking.

Communications | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | qMags

MILITARY COMMUNICATIONS

Cognitive Tactical Network Models

Ossama Younis, Latha Kant, Anthony McAuley, Kyriakos Manousakis, David Shallcross, Kaustubh Sinkar, Kirk Chang, and Kenneth Young, Telcordia Technologies, Inc.

Charles Graff and Mitesh Patel, U.S. Army CERDEC

This work was sponsored by U.S. Army CERDEC under contract DAAD-10-01-C-0062.

ABSTRACT

Unlike commercial MANET applications, tactical networks are typically hierarchical and involve heterogeneous types of radio communications. Future tactical networks also require cognitive functions across the protocol stack to exploit scarce spectrum and dynamically adapt functions and configuration settings. In this work we highlight the need for novel design tools for cognitive tactical networks. We define a system design model that will provide the foundation for generic network design problem formulations via the use of cognitive techniques covering both dynamic frequency adaptations and machine-learning-related aspects of cognition. We use the system model to identify several potential cognitive design knobs and describe how the different design knobs can potentially be adjusted at different timescales of operation. These knobs are used in formulating a cognitive network design problem. Finally, we discuss how a network designer can potentially benefit from the proposed model result, a cognitive network design toolset we have recently developed.

INTRODUCTION

The Federal Communications Commission (FCC) reports that most of the licensed spectrum is heavily underutilized [1]. At the same time, most available spectrum is licensed, and little is left for future wireless network applications. This is why new cognitive mechanisms have been proposed recently to efficiently exploit available spectrum and improve network performance [2]. Cognitive techniques are also proposed to provide radios with adaptation and learning capabilities to be able to autonomously react to dynamic conditions of the network.

Military applications also suffer from spectrum scarcity and thus can greatly benefit from cognitive functions. One challenge to cognitive network deployment in military applications is the lack of a cognitive network model that captures the peculiar features of military applications. Such features include the typical hierarchical nature of military groups (rank and urgency), island-like distribution of platforms, and the need for special connectors (unmanned aerial vehicles [UAVs] or satellite). Another challenge is the lack of a cognitive network design tool to answer *what if* questions and provide insights on expected network behavior using high-level analytic models. Such a tool should be able to contrast learning vs. adaptation techniques, estimate incurred overhead, and provide design recommendations on parameters, such as needed channels. Addressing these challenges paves the way to application of cognitive networks on SDR radios (e.g., Gnu [3] or WARP [4]) or potential cognitive radios (e.g., Boeing's JTRS GMR [5] and DARPA WNaN radio [6]).

In this work we propose a novel cognitive tactical network model that captures the features described above. We use this model to identify several potential *cognitive design knobs* and describe how the different design knobs can potentially be adjusted at different timescales of operation. This lays the groundwork for problem formulation of cognitive tactical network design. We briefly discuss how we have used the model and formulation to develop our new cognitive network design tool. We conclude this work in the final section.

SYSTEM MODEL FOR COGNITIVE TACTICAL NETWORKS

Here, we define the entities that are used in the military model of cognitive networks. Then we motivate the envisioned steps for designing a cognitive tactical network.

COGNITIVE SYSTEM COMPONENT DEFINITIONS

Platform: A physical entity that hosts a number of radio routers. Examples of a platform include a vehicle and a soldier.

Radio router: An entity that hosts a number of potential interfaces, with every interface corresponding to a single waveform. A good example of a radio router is a ground mobile radio (GMR) [5]. GMR radio is able to host waveforms like wideband network waveform (WNW) and soldier radio waveform (SRW), and has routing functionality to route IP packets across the waveforms over radio interfaces. There are two types of radio routers: *non-gateway* and *gateway*. N_1J denotes a non-gateway that has only a single radio interface. Note that J denotes the platform ID. G_iJ denotes a gateway that has i ($i > 1$) interfaces.

Interface: Participates in a single waveform at any given time and belongs to a radio router.

Waveform: Uses physical (PHY) and medium access control (MAC) schemes through which a group of interfaces can send traffic. A distinct set of frequency channels are allocated to each waveform. Our model frees a waveform from being associated with specific channels. We use w to denote the number of channels allocated to a waveform.

Frequency group: A frequency group is a set of frequency channels chosen from the frequencies allocated to a waveform, and are to be assigned to MAC subnets (defined later). The frequency groups can overlap (i.e., a single frequency channel can potentially belong to multiple frequency groups), but higher priority will only be given to a particular group at any given time. We use g to denote the number of frequency channels in a frequency group. When a frequency group is assigned to a *subnet* (to be defined below), the said subnet has *priority* over the frequencies in the frequency group assigned to it.

MAC subnet: A frequency group can be assigned to a group of radio interfaces and serves as the common transmission medium. A group of radio interfaces that are assigned the same frequency group and are in range of each other is referred to as a *MAC subnet*. (This differs from the concept of *IP subnet*.)

Frequency-set: The set of frequency channels a MAC subnet uses at a given time instance. Frequency-set differs from frequency group in that the frequency group is the *nominal* primary frequency set for which the MAC subnet has priority. However, based on real-time network conditions, frequency-set can consist of a different set of more or fewer channels. If we use h to denote the size of a frequency-set, a MAC subnet may, in general have:

- $h > g$: The subnet is exploiting available secondary frequencies in addition to its own g primary frequencies.
- $h < g$: The subnet is not using all of its g primary frequencies.

Cognitive System Design Model

In this section we present the cognitive system design model through illustrative examples. Here, the focus is on how a network would function given the concept of dynamic spectrum assignment at both the primary spectrum allocation (PSA) of available frequency bands among platforms and dynamic spectrum access (DSA), with dynamic frequency borrowing/lending.

Step 1: Define Traffic Flows — An estimate of traffic flows among platforms the network must support. Each flow has one source and either one destination (unicast) or multiple destinations (multicast).

Step 2: Platform Laydown — We envision a mission plan that is given to the commander and network planner. The mission plan, which is the input to the design/planning process, includes the following items:

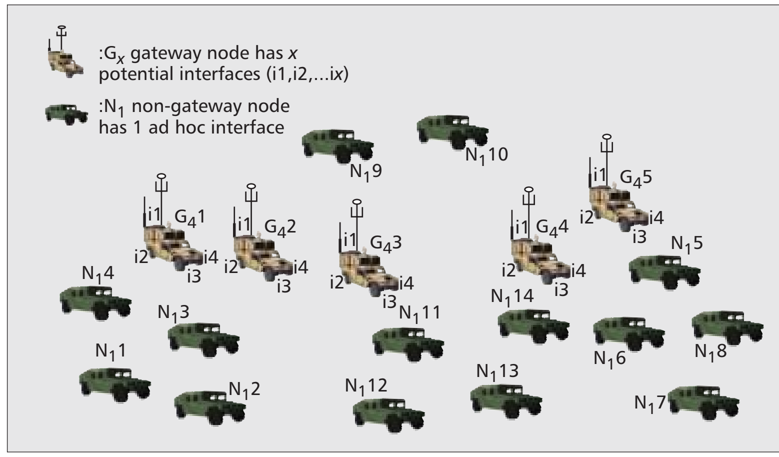


Figure 1. *Platform laydown.*

- Platform specifications, such as the number of radio interfaces and set of possible waveforms.
- Transmission range, the range corresponding to the maximum transmission power for a flat earth line-of-sight propagation channel.
- Data rate, which is variable for an adaptive waveform. It adapts its transmission format vs. signal-to-noise ratio (SNR).
- Locations, which are estimates of the waypoints (as a function of time). Platform location and asset profiles are shown in Fig. 1. It is noted that a platform is represented by X_{ij}, with $X = G$ representing a gateway platform and $X = N$ representing a non-gateway platform. Subscript i represents the number of radio interfaces, and j represents the unique platform ID of a platform. If $i > 1$, the platform is a gateway platform (i.e., with more than one radio interfaces).
- **Path loss**: Information that affects path loss, such as terrain and foliage. Path loss is specified for a range of frequencies.

Each platform type is allocated potential associations between the radio interfaces and waveforms. For example, we assume that all gateway platforms have multiple potential interfaces, and all non-gateway platforms have a single interface.

Step 3: MAC Subnet Design — MAC subnet design has two steps: selection of interface waveform and creation of MAC subnets. The two operations may be performed separately or together.

In Fig. 2, for example, two waveforms (waveform 1 and waveform 2) are assigned to gateway platforms (each of which has four potential interfaces) and are given either two or three actual interfaces. All non-gateway platforms can have either waveform 1 or waveform 2, but we configure all to be waveform 1 except for $N_1 9$ and $N_1 10$.

Interfaces of the same waveform are further divided into *clusters* or MAC subnets. The creation of the MAC subnets should be designed to increase the scalability of frequency assignment (which should be decided based on groups rather than links) and

MAC. The allocation of MAC subnets can be based on diverse goals and constraints, including flows among platforms, membership in command, and security.

Step 4: Topology Design — With the platform location and radio interfaces defined, the topology control algorithm can assign the power level of each radio interface, which decides the links among platforms. The goals include producing a connected network that satisfies a set of constraints (Fig. 3), including:

- Robustness: For example, *k*-connectedness.
- Available radio interface: Only the same interface type can be connected.
- Energy/power constraints: Network-wide or per individual platform or per interface.
- Delay: For example, network diameter constraints.
- Improve loss performance and capacity performance by considering traffic loading.

Step 5: Waveform Frequency Set — Frequency channels associated with a waveform can be defined within the limits of the cognitive radio hardware and software. In many cases, it is expected that the waveforms will be non-overlapping. However, overlap may be permitted.

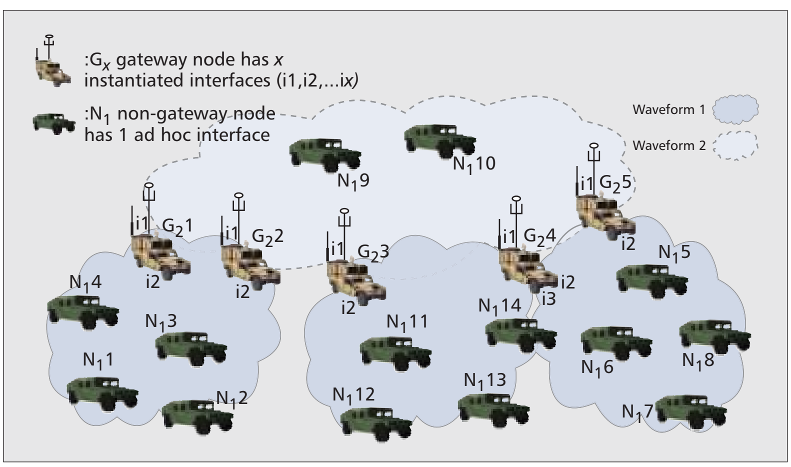


Figure 2. *Platform group and frequency group assignment.*

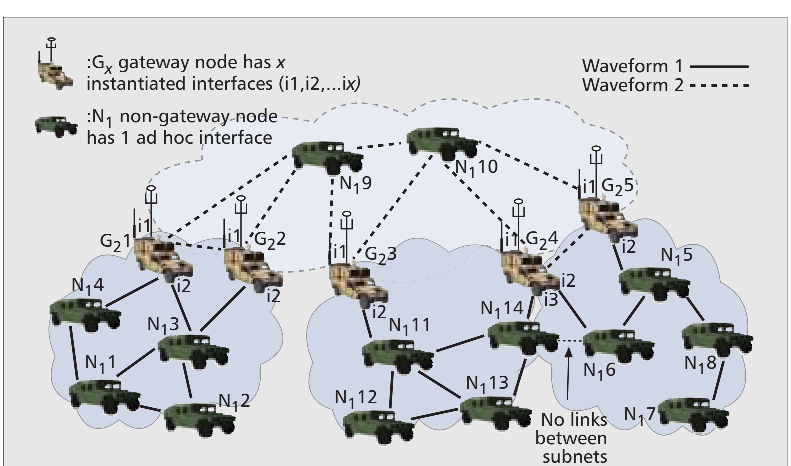


Figure 3. *Topology design.*

Step 6: Frequency Grouping — The frequency channels for which a frequency group has priority will be predefined, designed once during planning, or dynamically assigned. We use the following example to illustrate how frequency groups are formed and how they are assigned to a subnet. Assume that five frequency channels F_1–F_5 are assigned to a waveform. Of these five channels, F_1, F_2, and F_3 form frequency group G_α and F_4, F_5, and F_3 form frequency group G_β. We note that as shown in this example, frequency groups can overlap. However, for the overlapped channel, one group needs to be assigned as the *priority group* at any instance. For example, G_β may have the priority subnet for F_3 based on policy (e.g., alleviating congestion due to heavy load in subnets that are using G_β).

To offer more flexibility, it is possible to have frequency channels shared across waveforms, as shown in Fig. 4. In particular, two waveforms (waveforms 1 and 2) can have an overlapped frequency $\{F_6\}$. We plan to exploit the benefits of this additional level of flexibility in developing the design tool. Note that a *hopset* and its color in the figure refer to a frequency group.

Step 7: Primary Channel Allocation — The primary frequencies must be assigned for each subnet. We use the allocation of groups to subnets to define the primary frequencies a subnet can use. In general, the frequency group assignment algorithm should aim at reducing interference across subnets. Figure 5 gives an example of assigning the same group to subnets A and C, which are far enough apart not to interfere. The figure also shows the initial frequency assignments, assuming that F_3 is assigned to group G_β. Our model also assumes that one channel per waveform is used for sensing and information exchange among members of the same waveform. There is no need for information exchange among subnets of different waveforms.

Step 8: Dynamic Spectrum Access — Given the primary frequency group assigned to a subnet, frequency-sets are formed dynamically based on the traffic loading of a subnet and the interference.

To realize the concept of *available spectrum* or real-time channel allocation, channels can be dynamically *borrowed* from other frequency groups and form a frequency-set. In particular, suppose F_3 is in group G_β and hence assigned to subnet β. Assume subnets α and β are far away from each other; subnet α can utilize F_3, F_4, and F_5, as long as it does not cause interference to subnet β. In this case subnet α's frequency-set would consist of $\{F_1, F_2, \ldots, F_5\}$, and subnet β's frequency-set would consist of $\{F_3, \ldots, F_5\}$. In addition, even if a subnet using group α and a subnet using group β are close to each other, a subnet only uses the frequency channels it needs; the unused channels can be *lent* to any close-by subnets of the same waveform type. In our example, suppose subnet β is lightly loaded and only needs F_5; then F4 and F_3 can be lent to close-by subnet α. In this case subnet α's grequency-set would consist of $\{F_1, \ldots, F_4\}$, and subnet β's gre-

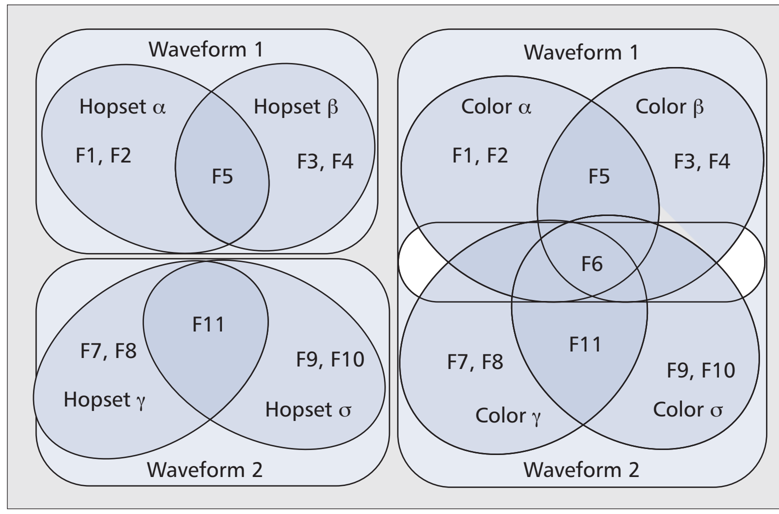


Figure 4. *Shared frequency channels between subnets (left) and waveforms (right).*

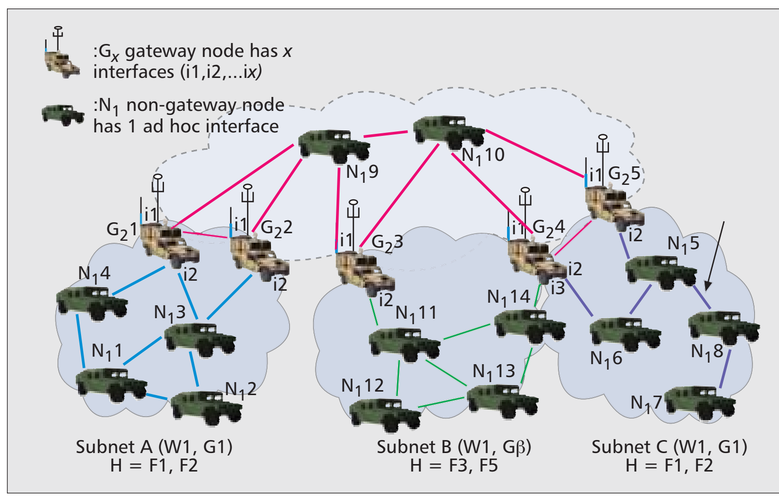


Figure 5. *Channel assignment to subnets.*

quency-set would consist of $\{F_5\}$. However, as soon as loading in subnet β increases, subnet α needs to release those lent channels back to subnet β.

As illustrated in Fig. 6, subnets A, B, and C are of the same waveform type and share the five frequency channels $\{F_1, \ldots, F_5\}$. Since subnets B and C are physically quite close to each other, no overlap is allowed in their frequency-sets. However, since subnet C does not require F_2 it can be *borrowed* by subnet B. For subnets A and B, however, suppose they are physically far away from one another, and as subnet A has very high loading, subnet A is allowed to borrow or use subnet B's frequency band F_3, F_4, F_5. Such exploitation of spectrum can result in much higher capacity in heavily loaded networks.

Again, the fact that subnet A is allowed to borrow subnet B's frequency band F_3, F_4, F_5 is due to the fact that subnet A using F_3, F_4, F_5 will not cause interference to subnet B, and loading on subnet A necessitates this borrowing. We believe that a distributed algorithm for frequency-set assignment can exploit the interference information needed for MAC slot assignment. Note that *all* of subnet A's nodes need to agree on releasing channels to subnet B before such release is granted (unanimous consent).

Through the flexibility realized by these two layers of channel assignment, primary channel allocation and DSA, we believe that network performance can be significantly improved.

Step 9: Computing the Routing Paths — The next step involves computing end-to-end routing paths based on the assignment routing metric(s). In Fig. 7 routing paths are calculated using a simple link cost, the inverse of bandwidth; but, in general, any link metric (or combination of link metrics) can be used.

Step 10: MAC Schedule — The MAC algorithm determines a link transmission schedule. Here, we focus on the time-division multiple access (TDMA)-MAC and provide an example in Fig. 8. In the figure only concurrent active links are shown. Note that concurrent links do not interfere.

Cognitive Network Design: Problem Formulation

This section is devoted to describing how a network designer could potentially use the generic system model described earlier to both formulate cognitive network design problems and solve them using a cognitive network design toolset.

Design Objectives

Based on the system model, we envision diverse problem formulations and design objectives for network planning. For example:

- Produce a minimum set of radios that satisfy a set of delay/loss constraints. In this case the design process starts with each platform having a specific maximum number of radio interfaces. Then provide other input parameters to identify, for every gateway, the minimum set of active radio interfaces needed to satisfy a set of QoS constraints.
- Produce an optimal set of subnets given the mobility pattern and available radio interfaces and platform grouping constraints.
- Produce a minimal set of primary frequency assignments.
- Quantitatively specify the performance gain realized by a primary frequency assignment and a dynamic spectrum frequency-set assignment compared to a static frequency-set assignment scheme.

Design Timescales

In order to produce solutions for the problems described above, each of the design steps should be executed at different timescales. Assuming a mission would typically last for several hours, we have the following time granularity levels, which in turn influence the frequency of potential design steps (i.e., how often a design step is executed).

Order of Several Hours (Once per Mission) — As an example, consider the MAC subnet and waveform frequency sets. Interface grouping (i.e., MAC subnet) and waveform frequency sets should last the entire mission. These assumptions are made according to how frequency planning and subnet design are done for future force networks like warfighter information network-tactical (WIN-T) and future combat systems (FCS). In addition, not explicitly assumed here is that the association between an interface and waveform is also fixed throughout the entire mission. Moreover, we also assume that traffic loading (i.e., IER) and platform movement pattern are given as input in advance.

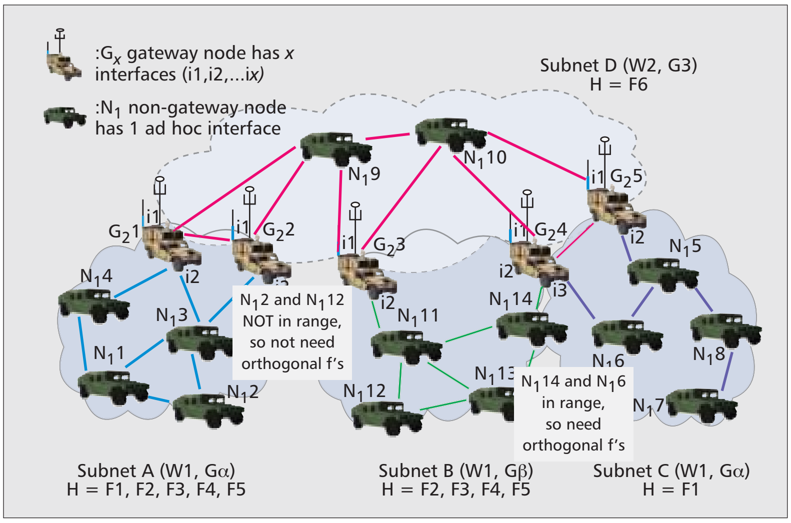


Figure 6. *Frequency-set assignment.*

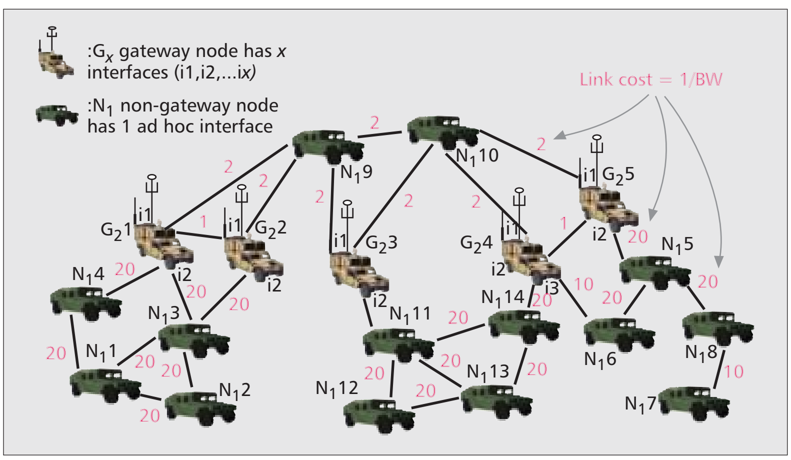


Figure 7. *Routing path computation.*

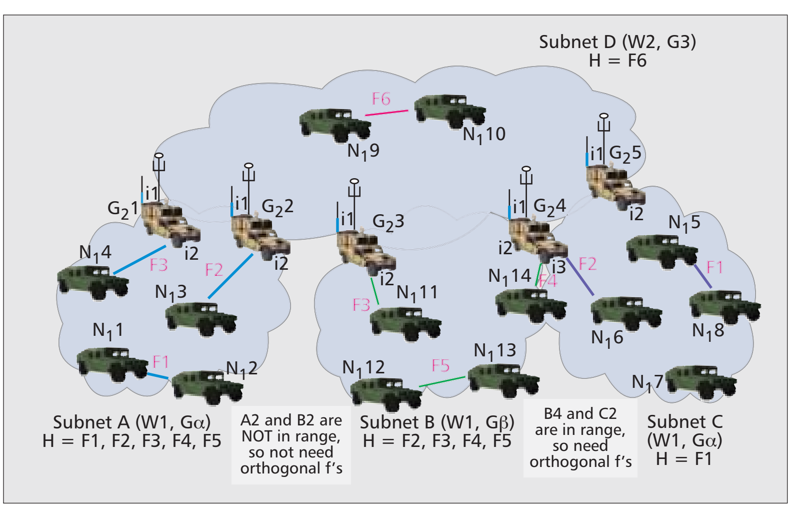


Figure 8. *MAC scheduling.*

Order of an Hour (Several Times in a Mission) — As an example, consider frequency grouping and primary frequency assignment. We assume frequency grouping (i.e., grouping the frequency channels assigned to a waveform into groups) and primary frequency assignment (i.e., assigning the frequency groups to the MAC subnets) can be done several times within a mission. These decisions are based on factors like past traffic pattern and the anticipated network behavior, and the results are distributed through the entire network as policies.

Order of Minutes (Once Every Minute or so) — Consider dynamic secondary frequency access. Although secondary channels must be released immediately if primary users are detected, we do not envision that spectrum opportunities will be exploited every few seconds. We envision that in a well-designed network, secondary spectrum access would have ample damping or hysteresis mechanisms so that secondary frequencies remain stable on the order of minutes.

Order of Seconds — As an example, MAC decisions could be made as frequently as once per TDMA frame, based on instantaneous loading (similarly for carrier sense multiple access [CSMA], in which it is also instantaneous). For most TDMA protocols used in waveforms like WNW or NCW, frame duration is on the order of a second or so. In addition, high mobility requires analysis within seconds.

Levels of Design

Cognitive network design spanning the aforementioned timescales can be performed using the concept of a snapshot during which network status, including traffic loading and node locations, typically remains the same. Thus, a cognitive algorithm cannot be invoked more frequently than a snapshot. The frequency with which the design steps are taken will be determined by the frequency of the snapshots. The frequency of the snapshots in turn determines the speed of adaptation and design runtime. Based on this observation, we propose three main *levels* of design:

Mission-level design: We envision that waveform frequency set design and frequency grouping can potentially be performed once per mission.

Phase-level design: We envision that MAC subnet design, topology maintenance, and primary frequency assignment can potentially be done several times per mission (i.e., in phases).

Snapshot-level design: We envision that secondary frequency access, routing, and MAC can potentially be performed once per snapshot.

Figure 9 shows our three phases of design. Note also that the algorithms at a particular level can also be jointly optimized by formulating them into a joint optimization problem. For example, MAC subnet design and topology can be jointly or separately optimized.

The assignment of functions to phases can be changed based on factors such as the level of uncertainty and dynamics. For example, the execution frequency of a cognitive topology control algorithm can be done in any of the phases (with different pros and cons):

Snapshot level: This should, in principle, yield the best overall connectivity for the entire mission. However, we believe it may result in too much overhead and instability in the network.

Phase level: As proposed in Fig. 9, this should yield sufficient adaptability to cope with dynamic changes while yielding stable designs. The execution frequency of the cognitive topology control algorithm in this case can be set identical to the MAC subnet assignment algorithm (Fig. 9), so one can explore the feasibility of joint optimization in a single optimization formulation.

Mission level: The transmit power of the radio could be fixed over the entire mission, identical to the waveform frequency assignment. For many dynamic missions this may limit adaptability, but one can explore the feasibility of jointly optimizing the waveform frequency assignment and transmission power level.

Sequential Design Flow

The design steps can be executed in a sequential manner. As an example, a simple noniterative design execution option is depicted in Fig. 10.

In this design execution the mission-level design is performed using the data provided for the entire mission (traffic, mobility pattern, etc.), and yields designs for subnet and waveform frequency set. Here, the design steps at the mission level can be performed sequentially (e.g., subnet design first followed by waveform frequency set design) or jointly (e.g., a joint optimization can be performed for both subnet and waveform frequency). Using the mission-level design as input, phase-level design is performed for the duration of the particular phase of the mission (which, again, can be done either sequentially or jointly). Similarly, the snapshot-level design is performed during this phase.

Iterative Design Flow

We also note that design steps can be executed in an iterative manner, and, depending on how the iterations are arranged, there could be several design execution options. After the very last snapshot-level design is completed, the mission-level design can be performed again based on the results of the phase-level and snapshot-level design, aiming to further improve the design and performance. This may result in a new and better mission-level design, which in turn can impact the phase- and snapshot-level design. Obviously, there are numerous ways to arrange the iterations.

Here, an automated design manager can assist a network designer to invoke the *best* cognitive functional module for a given set of network design goals/objectives. However, there are several trade-offs with each of these modules in terms of performance provided and overheads incurred, and a design tool is needed to encompass all these alternative design options and assess the trade-offs [7].

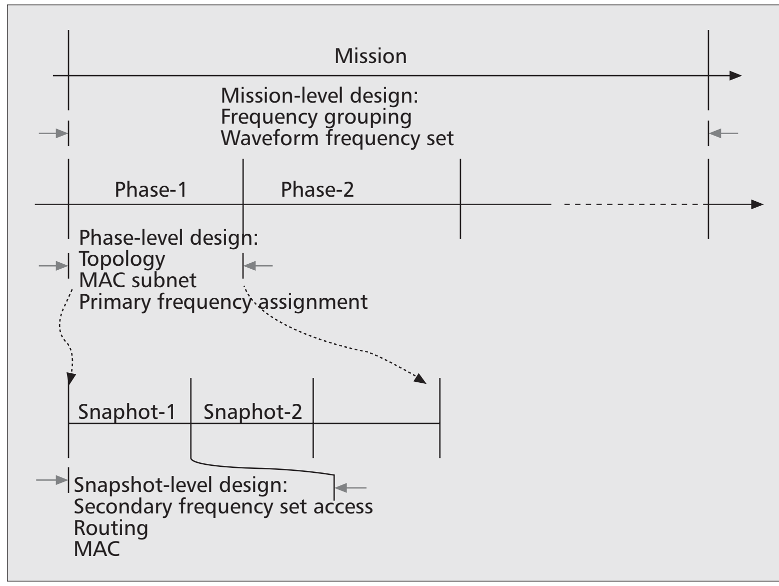


Figure 9. *Mission-level, phase-level, and snapshot-level design.*

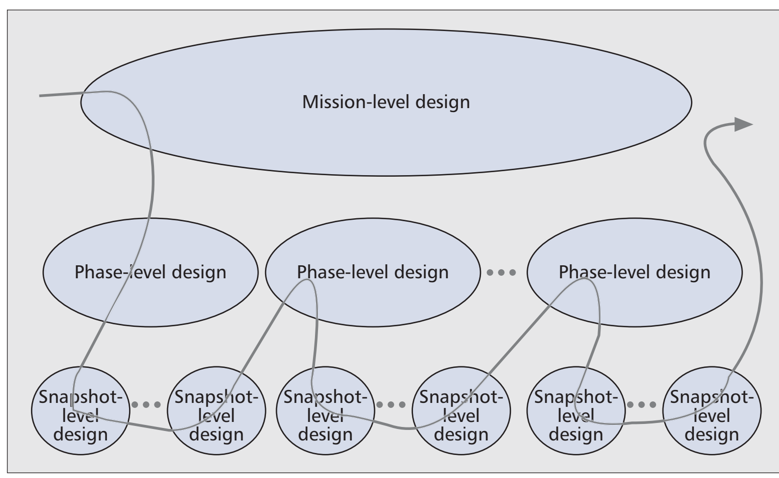


Figure 10. *Non-iterative design execution.*

Cognitive Tactical Network Design Tool

Significant extensions have been enacted to the Network Engineering Design Analytic Toolset (NEDAT) [8] to fit the model and problem formulation of cognitive network design in military applications. Details of the design of the tool can be found in [7, 9]. We summarize here the vision for the design of different layers and functions in the protocol stack, and describe the cognitive nature of the modules.

Subnet formation: Includes algorithms to select (waveform, radio interface) and cluster the network accordingly to fit the hierarchical military model. Every subnet then belongs to only one waveform.

Topology construction: Selects gateway routers in each subnet and creates logical links

Communications | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | qMags

We are currently exploring how to use learning techniques to predict the results of snapshot design, for example, whether or not constraints will be satisfied, what performance metric will be generated, and which parameter values should be used to achieve optimal design.

to connect platforms in subnets (intrasubnet) and across subnets (intersubnet). It should also introduce game-theoretic solutions to allow for distributed control.

Primary spectrum allocation: Divides channels that belong to a waveform among the subnets that belong to the same waveform according to the expected traffic load. The algorithm is traffic-aware.

Routing: Implements cognitive routing techniques that are based on reinforcement learning and cross-layer optimization. We have compared routing with reinforcement learning to that of adaptive routing. Our study has shown that learning can provide stable routing, while adaptive routing can be unstable in many cases.

Dynamic spectrum access: Exploits short-timescale connectivity information to redistribute underutilized channels among overloaded subnets, while keeping waveform-channel associations intact. We have proposed techniques to allow nodes to lend frequency channels, in a distributed fashion, to nearby subnets in need of such channels. We are currently exploring how to use learning techniques to enable stable/optimized channel reallocation.

MAC: Includes new formulation to jointly allocate slots and select channels for each active link that carries load. For distributed protocols (e.g., USAP [10]), the protocol should able to allocate channels while assigning slots. It should also provide functionality to minimize either the number of channels used or slots allocated.

Network design: We are currently exploring how to use learning techniques to predict the results of snapshot design; for example, whether or not constraints will be satisfied, what performance metric will be generated, and which parameter values should be used to achieve optimal design.

Note that in our model, topology design (including subnet formation), frequency grouping, and primary channel allocation are centralized. On the other hand, topology design (link formation), DSA, routing paths, and MAC schedules are all distributed. This matches an operational model in which a centralized entity makes decisions on initial subnets and frequency allocations to subnets, while all other processes can be performed at the radio node level.

Conclusion

In this article we have introduced a novel model of cognitive tactical networks that encompasses the peculiar features of troop deployment and channel scarcity. Based on this model, we have formulated the cognitive network design problem, and showed its different design knobs and execution-level alternatives. We have used this model and formulation in the development of a new cognitive network design tool that allows the designer to rapidly answer what-if questions and can also automatically tune cognitive network parameters [7]. Such a tool can be very useful, not only in assessing design feasibility, but also in exploring the vast space of parameter options and recommending which layers to optimize.

References

[1] FCC, "Spectrum Policy Task Force Report," ET Docket no. 02-135, 2002.

[2] I. Akyildiz, W.-Y. Lee, and K. Chowdhury, "CRAHNs: Cognitive Radio Ad Hoc Networks," *Ad Hoc Net.*, vol. 7, no. 5, July 2009, pp. 810–36.

[3] B. Le, T. W. Rondeau, and C. W. Bostian, "Cognitive Radio Realities," *Wireless Commun. Mobile Comp.*, vol. 7, no. 9, Nov. 2007, pp. 1037–48.

[4] Rice University, "Wireless Open-Access Research Platform"; http://warp.rice.edu/

[5] Boeing, "Joint Tactical Radio System, Ground Mobile Radios (JTRS GMR)"; http://www.boeing.com/

[6] DARPA, "Wireless Networks After Next (WNaN)"; http://www.darpa.mil/sto/solicitations/WNaN/

[7] O. Younis *et al.*, "Cognitive MANET Design for Mission-Critical Networks," *IEEE Commun. Mag.*, vol. 47, no. 10, Oct. 2009, pp. 64–71.

[8] L. Kant *et al.*, "NEDAT- A Toolset to Design and Analyze Future Force Networks," *Proc. IEEE MILCOM*, San Diego, CA, Nov. 2008.

[9] L. Kant *et al.*, "C-NEDAT: A Cognitive Network Engineering Design Analytic Toolset for MANETs," *Proc. IEEE MILCOM*, 2010.

[10] D. Young, "USAP: A Unifying Dynamic Multichannel TDMA Slot Assignment Protocol," *Proc. IEEE MILCOM*, 1996.

Biographies

OSSAMA YOUNIS [M] () is a senior research scientist in Applied Research at Telcordia Technologies, Inc. He was an assistant research scientist in computer engineering at the University of Arizona from July 2005 to June 2007. He received his Ph.D. degree in computer science from Purdue University in 2005. He has also received B.S. and M.S. degrees in computer science from Alexandria University, Egypt. He has served on the Technical Program Committees for several international conferences. His research interests include wireless sensor network protocols and applications, network security, and cognitive radio networks. He is a member of the ACM.

LATHA KANT is chief scientist and director of the Mobile Networking Research Department at Telcordia. She is Telcordia's Lead on the Network Design Project, a joint project between CERDEC and Telcordia, and is leading a team of research scientists involved with the design and development of NEDAT described in this article. She is also leading a research project on "Network Science" as part of the Army Research Laboratory Collaborative Technology Alliance, and is Telcordia's lead on the Future Combat Systems QoS task. Her research interests include design and performance analyses of wireless networking and network management systems.

ANTHONY MCAULEY received his Ph.D. from Hull University, England, in 1985, and has been with Telcordia since 1987. He is currently a chief scientist in the Mobile Network group. His research interests include wireless ad hoc networking, with an emphasis on analytic design and robust protocols for auto-configuration, routing, mobility, security, and QoS. He has also built several Linux mobile internetworking systems, designed efficient error control codes, and built VLSI chips.

KYRIAKOS MANOUSAKIS is a senior research scientist with the Mobility Management Research Department at Telcordia Technologies. His research interests include mobile ad hoc networking, protocol design, and network optimization. He has been involved with the application of stochastic approximation techniques and, in particular, simulated annealing, to the design and analysis of mobile ad hoc communications networks.

DAVID SHALLCROSS is a senior research scientist with the Information Analysis and Services Department, Applied Research, Telcordia. His research interests include research in and application of mathematical programming and optimization in the context of telecommunications. This has included developing prototype design tools for a variety of different sorts of networks, developing improved algorithms for workforce scheduling, research into methods of network topology inference from limited information, and research into methods and consequences of routing in telecommunications networks.

Communications | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | qMags

KAUSTUBH SINKAR is a research scientist with the Mobility Management Research Department at Telcordia Technologies. His research interests are in the areas of computer networking with focus on cooperative networks, multi-interface devices, mesh networks, and network security. He has an M.S. from Brooklyn Polytechnic University.

KIRK CHANG [SM] is a director in the Applied Research Area of Telcordia Technologies. He received his Ph.D. degree from the Electrical Engineering Department of the University of California, Los Angeles. His research interest is in wireless communications, wireless ad hoc networking, IP routing, and QoS, and he holds 12 U.S. patents in these areas.

KENNETH YOUNG [SM] is executive director for government project development in the Applied Research organization at Telcordia, where he leads projects on mobile ad hoc networking technology. He is the program manager for the Army Research Laboratory's Communications and Networks Collaborative Technology Alliance, a consortium that performs basic research in survivable wireless mobile networking, signal processing, and tactical information protection. He heads another team developing advanced mobile technology under CERDEC's Proactive Integrated Link Selection for Network Robustness program. He is a Telcordia Fellow.

CHARLES GRAFF is senior project engineer at CERDEC, U.S. Army, RDECOM Fort Monmouth, New Jersey. As CERDEC Lead for the Network Design project, he leads a team of several engineers at Fort Monmouth who are working jointly with the Telcordia team in developing portions of the NEDAT program. As an expert on wireless and mobile ad hoc networking, he also serves as an advisor to several key projects at CERDEC.

MITESH PATEL is an electronics engineer at the U.S. Army CERDEC Aberdeen Proving Ground, Maryland. He received his B.E. in electrical engineering from the City College of New York in 1995 and his M.S. in electrical engineering from Stevens Institute of Technology in 1999. His research interests include tactical mobile ad hoc networks, quality of service, and secured wireless networks. He has worked on prior programs including the TMSIP, MOSAIC, and PILSNER ATO program providing test and evaluation support. He is currently supporting the Cognitive Networking Technologies program at CERDEC.

Communications | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | qMags

Robust Web Services in Heterogeneous Military Networks

Ketil Lund, Espen Skjervold, Frank T. Johnsen, Trude Hafsøe, and Anders Eggen, Norwegian Defence Research Establishment (FFI)

ABSTRACT

NATO Network Enabled Capability is first and foremost about achieving better interaction between the different actors involved in military operations. This implies more efficient exchange of information. Consequently, the NATO information infrastructure will consist of a federation of systems, including a plethora of different information and communication systems, as well as a mix of new and legacy systems. NATO recommends a service-oriented architecture approach based on Web services to enable such a federation. In this article we explain how the communication protocols normally used in Web services are unsuited for disadvantaged and heterogeneous networks. We then present our prototype proxy, which enables the use of standard unmodified Web services across all network types, including tactical networks with low data rates and frequent disruptions. It is designed to work with existing security mechanisms, and also offers further optimizations in the form of optional plug-ins.

INTRODUCTION

The North Atlantic Treaty Organization (NATO) and the member nations are migrating toward NATO Network Enabled Capability (NNEC). The NATO policy is to adopt civil standards as far as possible in order to be able to procure commercial off-the-shelf (COTS) products. International standards are important for the members to agree on common solutions and for the industry to compete under equal conditions. Two important recommendations made in the NNEC Feasibility Study (FS) [1] are that the information infrastructure should be implemented as a service-oriented architecture (SOA), and that IP should be used as a common network protocol in all network types. An SOA is most commonly realized through Web services using XML formatted documents. This technology is defined in a number of standards, and while this standardization work is still ongoing, Web services remain the most adopted SOA implementation technology. By providing a standardized way of describing interfaces and data formats, Web services enable loose coupling of information consumers and producers, which means that each nation is free to implement clients and services according to its own requirements and preferences. If the military requirements cannot be met by civil standards, specialized military solutions must be developed and issued as NATO standardized agreements (STANAGs).

It is not realistic to expect all NATO nations to adopt the same systems, and there will always be legacy systems that must be integrated and need to coexist with contemporary systems. Therefore, the NATO nations must instead agree on standardized descriptions of interfaces and data formats, and leave it to each nation to implement the interfaces according to the specifications. By implementing such clients and services using Web services, and according to the agreed standards and formats, interoperability between nations is ensured. In addition, the use of Web services means that using COTS software in many cases is a viable solution, contributing to reduced cost and development time.

However, Web services use XML, and XML documents tend to be large, causing significant overhead. This represents a problem when trying to extend Web services into tactical networks. In addition, the transport protocol normally used in Web services implementations, TCP, is not suited for networks characterized by high delay and frequent disruptions.[1] This is of particular importance when considering users in the field who may only communicate with others over *disadvantaged grids* (i.e., tactical communication systems with low data rates, high delays, and frequent disruptions).

WEB SERVICES IN HETEROGENEOUS NETWORKS

In Web services all communication is based on sending XML-based SOAP messages.[2] A SOAP message is an *envelope* consisting of a header and a body. The header contains information related to the handling of the message, such as addressing and security information, while the body contains the application data. In regular Web services, SOAP messages are transmitted using the HTTP protocol, which in turn uses the TCP protocol for reliable transfer of the messages. This protocol set is not suited for use in

[1] *There are optimizations for TCP, such as TCP Reno and TCP Vegas, which might alleviate the delay problem. However, these do not handle disruptions. For a comparison of TCP varieties see [2].*

[2] *One exception to this is REST Web services, which do not rely on SOAP. However, SOAP is a better solution with respect to maintaining security, which is why we focus on SOAP Web services in this article.*

disadvantaged grids, and the main question is how to enable the use of Web services in disadvantaged networks and across heterogeneous networks. Through our research on Web services in disadvantaged grids, we have found that this question can be broken down into three requirements that must be met:

- Reduce the network traffic generated by Web services
- Remove the dependence on end-to-end connections
- Hide network heterogeneity

Addressing Web Services Overhead

The first problem is related to the amount of network traffic generated by Web services. It is necessary to reduce both the size of the individual messages and the number of messages being transmitted. XML is a rather verbose language, and tends to produce much larger messages than binary formats do. Using techniques such as compression will reduce the size and thus the bandwidth requirements of each individual message, but will not reduce the number of messages sent between nodes. In our work we have looked at several ways of limiting the number of messages:

- Employing caching near the clients, which allows for reuse of older messages
- Using the publish/subscribe paradigm, where clients subscribe to information instead of requesting it, allowing the same message to be sent to multiple clients
- Employing content filtering to ensure that only relevant data is transmitted

End-To-End Connections

The second issue is that regular Web services depend on a direct end-to-end connection between the client and the service. TCP is connection-oriented, and designed for wired networks, which means that the control mechanisms are designed for handling congestion, and much less for handling errors. In tactical networks with high error rates and high latencies, the congestion control of TCP will therefore cause suboptimal utilization of the network due to frequent connection timeouts. When multiple networks are interconnected (Fig. 1), TCP's need to establish an end-to-end connection increases this problem; each traversed network adds delay, increasing the risk of connection timeout. Similarly, HTTP is synchronous, which means that when a SOAP request is sent, the HTTP connection is kept open until the SOAP response is returned in the HTTP acknowledgment message. If the connection times out, the SOAP response cannot be routed back to the service consumer.

The obvious solution to this problem is to replace HTTP and TCP with other more suitable protocols. However, this requires modifying the application software. Alternatively, an extra communication layer can be introduced. Within this layer, hidden from the applications, more suitable protocols can be used (e.g., tactical protocols such as STANAG 4406 Annex C and E), which are able to withstand long and variable round-trip times, and have little communication overhead.

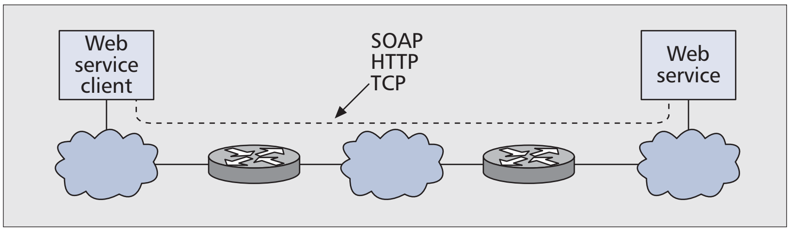


Figure 1. *HTTP and TCP establish end-to-end connections.*

By implementing this extra communication layer in a *proxy* solution, standards compliance can be retained. A proxy is a node in the network between a client and a server through which the network traffic passes. A proxy can be used for several purposes, such as caching, firewalling, and content adaptation. For example, HTTP proxies have been popular on the Internet for years, since they lower response times when surfing the web. Web services proxies follow the same principle as HTTP proxies, in that they function as a middle man between the provider and the consumer of the service. However, they do not just understand the HTTP protocol, they must be able to recognize and process SOAP as well. We have developed an initial SOAP proxy prototype [3] that we extend further in this article.

Introducing an extra communication layer means increased flexibility when it comes to selecting which transport mechanism(s) to use. Additionally, using this approach means that the end-to-end connection dependence is removed in favor of a per-hop-behavior. In this case the application software can often be left unmodified. However, there is a possibility for information corruption along the route, which may not be detected without an end-to-end connection. Also, since packets are acknowledged on a per-hop basis, you do not get end-to-end reliability. These two issues can be mitigated if the client uses application-level solutions for error control and reliability. For example, using XML signatures will ensure that any modifications to a SOAP message are detected by the receiver, despite the lack of end-to-end error control. Furthermore, using, for example, the WS-ReliableMessaging specification provides application-level acknowledgments, so an end-to-end connection on the transport layer is not needed to acknowledge delivery. Thus, client software must use the appropriate Web services specifications to add the desired level of resilience to their SOAP messages.

Network Heterogeneity

The third problem arises when heterogeneous networks are interconnected. In disadvantaged grids it is not uncommon to experience data rates of less than 1000 b/s [4]. In particular, when several users are using the network simultaneously, the effective data rate can become very low because resources are shared. Connecting such networks to faster networks introduces a risk that the gateway between the networks has to drop packets due to its buffers filling up faster than the packets can be transmitted out onto the lower-capacity network. This problem can be

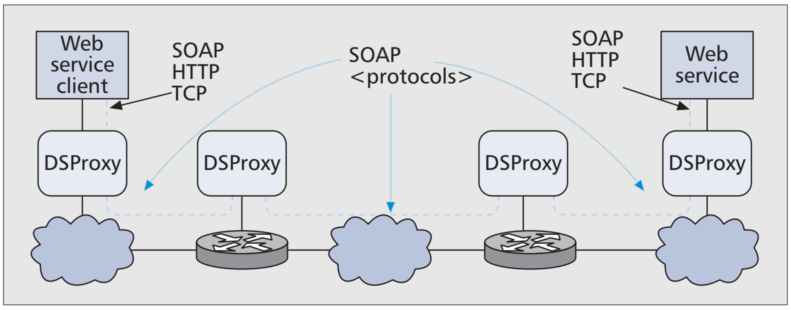


Figure 2. *A DSProxy overlay network.*

countered by introducing store-and-forward capabilities into the network. In addition, a store-and-forward capability can help alleviate the problems that arise from frequent communication disruptions, which can prevent a message from being delivered immediately. Having store-and-forward support can ensure that the message is not dropped and subsequently has to be retransmitted.

When traversing heterogeneous networks, different communication protocols may be required. This means that a message traversing several networks may have to use multiple different protocols on its way from sender to recipient. Therefore, it is necessary to add store-and-forward functionality on the application layer, not on the network layer.

DELAY- AND DISRUPTION-TOLERANT SOAP PROXY

In our previous work [5] we have focused on how to reduce network traffic through the use of techniques such as compression and content filtering. We now extend this work by introducing response caching and publish/subscribe to further reduce network traffic. We have implemented all these mechanisms, combined with the techniques discussed above, in a proxy prototype.

Our prototype middleware system, called the Delay- and Disruption-Tolerant SOAP Proxy (DSProxy), addresses many of the challenges associated with utilizing Web services in disadvantaged and heterogeneous networks. This proxy is an implementation of a wide array of principles and mechanisms that tackle different aspects of these challenges. The proxy software is designed to be modular, and its functionality can be divided into *core functionality* and optional *plug-ins*. The core functionality includes the basic optimizations that are required to make standard Web services work in tactical networks, while the optional plug-ins provide further optimizations such as caching and publish/subscribe support. A key difference between the two types of functionality is that the core functionality does not rely on inspecting the SOAP messages that pass through the proxy. The plug-ins, on the other hand, might require inspecting data or rely on making small modifications to clients and Web services.

CORE FUNCTIONALITY

The DSProxy is a proxy for Web services designed to handle all types of information and traffic flows that are suited to be implemented as Web services. Some types of data, such as voice flows or other types of information with strict real-time demands, are likely to require other forms of optimizations beyond what can be supported by standard Web services.

The DSProxy system comprises multiple proxy instances deployed in a network, and forms what is known as an overlay[3] network (Fig. 2). The nodes that constitute the overlay network communicate with each other and exchange information about their state and the environment in which they operate in order to maintain the overlay. The proxy is a lightweight solution that can be deployed locally on any node in the network. The largest benefit will be achieved if the proxies are deployed both on nodes that bridge networks and in any node that communicates over a disadvantaged grid.

The DSProxy overlay can be maintained in one of two ways: manual configuration by an administrator or dynamic configuration by the proxies themselves. In the first case there is a human in the loop responsible for creating a usable configuration. The second case warrants further discussion, because the proxies make autonomous decisions regarding overlay reconfiguring and routing. In the dynamic overlay proxies advertise their presence by broadcasting messages, allowing them to be discovered by neighboring proxies. Topology information is distributed among the overlay members, which use the information to create the shortest routes through the overlay. If network conditions change, the dynamic overlay reacts by reconfiguration to address these changes.

For these reasons, using an overlay adds resilience to a dynamic network. However, in cases where a multihop network is stable, going through the proxy overlay has more latency compared to invoking a Web service directly. This is because there is a small processing penalty per overlay node hop: In each proxy, uncompressed SOAP messages add 3 ms processing time, whereas compressed SOAP messages add 11 ms processing time (this is because they need to be decompressed in each hop so that the SOAP header can be inspected). Additionally, some latency can be added due to the fact that the messages follow the shortest overlay path, which might be slightly longer than using the direct route. However, this effect will be mitigated by the overlay reconfiguration algorithm, which ensures that the messages are sent via as few proxies as possible.

In order to enable end-to-end connections across multiple and heterogeneous networks, the DSProxy system offers store-and-forward capabilities at each instance in the network. Instead of relying on end-to-end connections between information consumers and producers, as shown in Fig. 1, the nodes relay the information between the parties by initiating single-hop connections between proxies, storing the information, and retransmitting or rerouting when encountering failures. By storing information at

[3] *An overlay network is a logical organization of nodes that form a virtual network on top of physical networks.*

IEEE Communications | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | qMags

intermediate nodes, added robustness is introduced into the network. The messages are stored locally, either in memory or on a storage medium on the node. Once the message is successfully relayed to the next proxy, the message is removed from the storage.

By eliminating the need for end-to-end connections, one is no longer required to use TCP for information exchange; one is free to choose more suitable transport protocols instead. As illustrated in Fig. 2, SOAP is used as a pervasive message format, but the underlying protocols can be replaced on a per-hop basis depending on the network characteristics and the quality of specific links between the DSProxy nodes. This feature, together with the store-and-forward capabilities, enables efficient traversal of multiple heterogeneous networks, while hiding the underlying heterogeneity from the end users. For any particular Web services invocation, the request may traverse a series of different networks, dealing with different link qualities and utilizing different transport protocols. By utilizing compression, SOAP and XML-based information exchange is made possible even in networks with severely restricted data rates. Our prototype uses GZIP compression, but due to the modular design of the proxy, support for additional algorithms can be included later.

A key concern is standards compliance, and our prototype interoperates with COTS Web services software without modification. The system accepts regular incoming Web services invocation requests over HTTP and TCP, and is able to relay such requests to proxies, and finally to Web services endpoints, as shown in Fig. 2.

The TCP connection from the client is terminated at the first DSProxy, meaning that the TCP request from the client is acknowledged immediately. This allows the TCP connection between the client and the first proxy to be kept open for the entire duration of the invocation process, and it will not time out, regardless of how long it takes for the service to respond. Once the response message is available, it is returned to the client using the same connection.

By placing the first proxy as close to the client as possible, preferably on the same physical machine/device, challenges associated with TCP and disadvantaged grids are avoided. This deployment scheme ensures store-and-forward capabilities throughout the network. Likewise, for the last communications hop, between the last proxy in the chain and the Web service, standard HTTP and TCP is used, allowing interoperability with existing Web services without requiring any modifications. By simply modifying the URLs used by clients to invoke Web services, the communication is routed through the overlay network. Any instructions to the proxies are specified as extra parameters added to the URL. Apart from the modified URL, both Web services clients and services are unaware of the presence of the DSProxy system.

Security

Security is a key concern in all military networks, and the increased information flow between systems that is central to the NNEC vision demands flexible security solutions. Current security policies tend to require IPsec or link layer cryptography, while Web services security mechanisms can provide end-to-end application layer security.

The DSProxy core functionality is designed to work with existing security mechanisms, ensured by the core functionality of the proxy being content agnostic. The proxies will not parse, inspect, or interpret the body parts of SOAP messages, which allows for end-to-end security through encryption. Because the information needed to route the messages is located in the SOAP header, there is no need to decrypt the messages being transported. If a Web service client relies on encrypting the entire SOAP message, a new SOAP envelope can be wrapped around the original message, placing routing information and DSProxy parameters within the unencrypted SOAP header. This is done by the first proxy in the chain, which gets this information from the modified invocation URL provided by the client. Because the SOAP header is unencrypted, the system remains compatible with the various Web services security related standards (see [6] for a survey of these). Also, because the DSProxy system does not require any special behavior by protocols that are part of the IP stack, it can also be used with lower-layer security mechanisms, such as IPsec and link layer cryptography.

While the DSProxy core functionality works together with end-to-end security solutions, the optional advanced functionality described below, such as response data caching and the publish/subscribe mechanism, require all intended recipients to be able to decrypt the response message targeted toward one specific client. Because COTS Web services are unaware of these modes of operation, they will encrypt a given response message using the key designated for the client responsible for the actual invocation. In order to make this work with multiple clients, one should rely on group keys distributed to all intended recipients.

Optional Plug-ins

In many types of tactical scenarios it is a common situation that multiple information consumers require the same information provided by the same information producers. This can be applications such as blue force tracking, weather forecasts, or sensor data. In order to optimize bandwidth requirements for such scenarios, the DSProxy system supports response data caching, which offers a kind of economy of scale when serving multiple clients requesting the same information. By caching response data messages at multiple nodes in the network, clients may, depending on the degree of time criticality of the information, suffice with cached response data. This means that instead of having the chain of proxies relay the Web services invocation message all the way to the actual Web service and invoking it, a proxy along the way may determine that it has a valid cached response for the request and return it to the client. Because caching behavior is specified by each client by appending a parameter to the Web service endpoint URL, the client is in control of if and when cached response data will suffice, and how

By placing the first proxy as close to the client as possible, preferably on the same physical machine/device, challenges associated with TCP and disadvantaged grids are avoided. This deployment scheme ensures store-and-forward capabilities throughout the network.

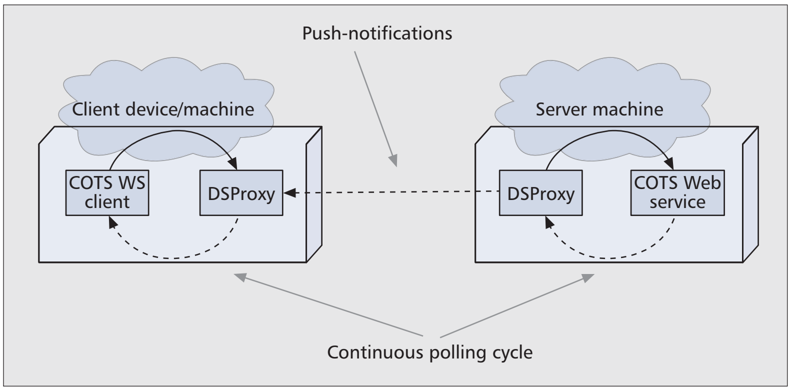


Figure 3. *Standard request/response-based Web services client and server engaging in publish/subscribe communication using the DSProxy overlay network.*

old these messages may be. While a client may request response data caching on a per request basis, its availability is not guaranteed, because the caching behavior is configured for each DSProxy. This configuration option is critical for proxies running on limited devices, which may lack the system resources needed for extensive caching.

For contexts in which multiple consumers subscribe to services expected to produce response messages that change very little over time, great size and bandwidth reductions can be achieved by only sending the message differences between proxies. Once a DSProxy has received a response message from another instance, it is able to reconstruct the next response message using only the difference between the messages. However, this requires the proxies to keep state information, since they need a previous message to deduce the new message from using the differential information. An example of such a scenario is when multiple consumers subscribe to a blue force tracking service. While the messages produced by the Web service in response to invocations comprise HTTP headers, SOAP headers, and SOAP bodies, all that may have actually changed between two subsequent responses may be the units' location coordinates. Experiments have demonstrated that such messages containing hundreds of bytes of information can be reconstructed without loss using differences consisting of merely tens of bytes.

While standard Web services are based on the request/response paradigm, this is not always the best choice for disadvantaged grids. For time-critical information, clients have to poll Web services frequently, which may lead to wasted polls and added network traffic. In such situations it makes sense to have the information producers push the information to the consumers as soon as it becomes available. The DSProxy system enables such publish/subscribe mechanisms to be used together with standard Web services software, requiring only minimal modifications to be made to the clients [7]. In order to subscribe to a Web service, a parameter is added to the URL before routing the request through the overlay network. This parameter instructs all proxies along the invocation path to treat the request as a subscription, and the last proxy in the chain initiates a polling cycle to the actual Web service (Fig. 3). This proxy sends an ordinary request to the Web service at regular intervals, and compares the response to the previous response in order to see whether the data has changed.

Placing the last proxy on the same physical machine as the Web service means that polling only burdens the internal machine resources, and no traffic is transmitted across the network. Once this proxy discovers that new information is available (i.e., the response from the web service has changed), it notifies its subscribing proxies by initiating an outgoing connection. Once the information arrives at the first proxy in the chain, it is made available for the client to retrieve. As standard Web services clients lack the ability to accept incoming connections, the client is required to poll the first proxy regularly to obtain updated information. This typically involves wrapping the Web services invocation calls in a loop structure. This loop must also handle when the poll returns without new data. Again, placing the first DSProxy on the client machine avoids added network traffic. As with response data caching, economies of scale are achieved when several clients subscribe to the same information, since one proxy can subscribe on behalf of many clients as well as on behalf of other proxies.

Publish/subscribe specifications like WS-Notification and WS-Eventing are emerging, but they are not yet in widespread use, compared to regular Web services. WS-Eventing is supported by the DSProxy, and support for WS-Notification is under development. However, the DSProxy publish/subscribe mechanism works with non-publish/subscribe Web services, allowing already existing software to utilize publish/subscribe.

FIELD TRIALS

Our proxy was tested in a series of live field trials at Combined Endeavor in September 2009, with promising results. During these experiments, the DSProxy system was deployed in a small combined setup consisting of four networks that were interconnected using IP on the network level and the proxies on the application level. Figure 4 shows the network setup used during these experiments, where we interconnected the Norwegian and NATO C3 Agency (NC3A) networks. Interoperability between the nations was provided using the tactical communications standard (TACOMS) for federated networking, and the common Region C (RG C) Combined Endeavor backbone. The DSProxy was deployed in the gateways between the networks, and on each individual node within the Norwegian mobile ad hoc network (MANET). This deployment meant that not only did the proxy ensure successful information exchange between the networks, but it also enabled Web services communication within the MANET, where most of the tactical communication took place. This meant that the tactical users did not have to concern themselves with whether or not they were connected to the rest of the network when sending information.

Related Work

In [8] a solution employing proxy servers for disruption-tolerant networking (DTN) is presented. This work uses proxies to translate from applications' native use of end-to-end transport protocols to network-level DTN messages. This solution allows standard TCP and UDP software to take advantage of DTN without modification. However, the solution relies on network-level DTN support.

Faucher *et al*. [9] introduce an experimental Web services proxy aiming to overcome the challenges associated with disadvantaged grids. Their proxy handles SOAP sent over HTTP/TCP. The authors argue that XML compression will enable the proxy to perform better in tactical networks, but they have not implemented this.

Metzger *et al*. [10] have created a proxy adding delay tolerance to the XML-based Jabber protocol. Jabber is mainly used for chat, but a SOAP extension that allows Jabber to carry Web services traffic exists. Their prototype system does not support dynamic joining of group members, limiting its use to preconfigured static users.

Summary

In this article we have discussed the importance of Web services for realizing NNEC. Web services, being based on standards, are central to the NNEC vision because the technology provides interoperability between applications, enabling the NATO nations to interconnect their command and control systems and share information in a functional and cost-efficient manner.

Furthermore, we have highlighted issues related to adopting Web services technology in military networks and identified ways of mitigating these issues. In particular, we have introduced a network of proxies that add delay tolerance to SOAP, in addition to employing other overhead-reducing measures. We have implemented a prototype solution, and shown that it makes it feasible to employ Web services in tactical networks in a field trial.

At Combined Endeavor 2009, we showed that our prototype enables the use of unmodified Web services software across heterogeneous tactical networks, and that it also can augment the functionality of such Web services by introducing support for subscription-based services.

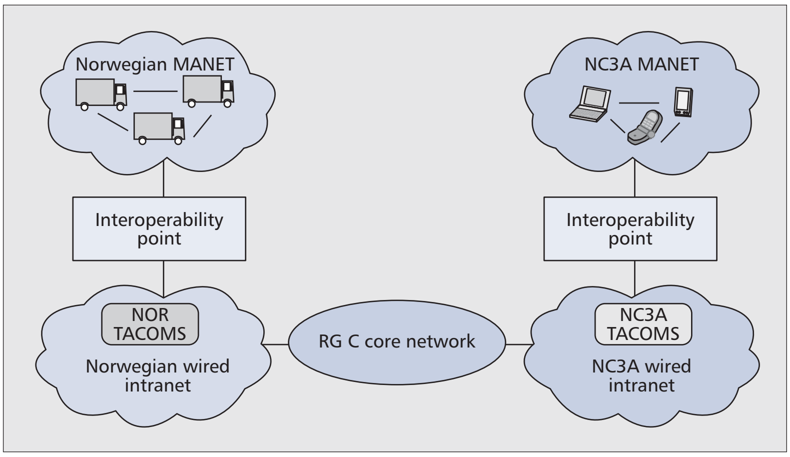


Figure 4. *Combined Endeavor experiment setup.*

References

[1] P. Bartolomasi *et al*., "NATO Network Enabled Capability Feasibility Study," v. 2.0, Oct. 2005.

[2] S. Papanastasiou, *Investigating TCP Performance in Mobile Ad Hoc Networks*, VDM Verlag, May 2008.

[3] E. Skjervold *et al*., "Delay and Disruption Tolerant Web Services for Heterogeneous Networks," *IEEE MILCOM*, Boston, MA, Oct. 2009.

[4] J.-C. St-Jacques, "Challenges for a Distributed Collaborative Environment Functioning over Mobile Wireless Networks," *Wksp. Role of Middleware in Sys. Functioning over Mobile Commun. Net*., NATO IST-030/RTG-012, 2003.

[5] K. Lund *et al*., "Using Web Services to Realize Service Oriented Architecture in Military Communication Networks," *IEEE Commun. Mag*., vol. 46, no. 10, Oct. 2007, pp. 47–53.

[6] N. A. Nordbotten, "XML and Web Services Security," FFI-report 2008/00413; http://rapporter.ffi.no/rapporter/2008/00413.pdf

[7] E. Skjervold *et al*., "Enabling Publish/Subscribe with COTS Web Services across Heterogeneous Networks," *4th Int'l. Wksp. Architectures, Concepts, and Tech. for Service Oriented Comp*., Athens, Greece, July 2010.

[8] K. Scott, "Disruption Tolerant Networking Proxies for On-the-Move Tactical Networks," *IEEE MILCOM '05*, vol. 5, Oct. 2005, pp 3226–31.

[9] R. Faucher *et al*., "Guidance on Proxy Servers for the Tactical Edge," MITRE Corp. tech. rep. no. MTR 060175, Sept. 2006.

[10] R. Metzger and M.C. Chuah, "Opportunistic Information Distribution in Challenged Networks," *Proc. 3rd ACM CHANTS '08*, San Francisco, CA, 2008, pp. 97–104.

Biographies

KETIL LUND (ketil.lund@ffi.no) is a scientist at the Norwegian Defence Research Establishment (FFI), where he has been working since 2006. His research interests include service-oriented architectures, Web services, quality of service, and middleware. At FFI he is currently working within the area of secure pervasive SOA. He received his Ph.D. in informatics from the University of Oslo (UiO).

ESPEN SKJERVOLD (espen.skjervold@ffi.no) is a research scientist at FFI, engaged in theoretical research and practical development in areas such as distributed systems and service-oriented architecture. He is also a part of FFI's team of professionals aiming to support and refine FFI's common systems design and development practices. He holds a Master's degree in distributed systems and computing (2007) from Brunel University, London, United Kingdom.

FRANK T. JOHNSEN (frank-trethan.johnsen@ffi.no) received his Cand.Scient. degree from UiO in 2002. Following this he worked with networks and distributed systems at UiO until he started work as a scientist at FFI in 2006. At FFI he is currently working within the area of secure pervasive SOA. His research interests include Web services, quality of service, and middleware.

TRUDE HAFSØE (trude.hafsoe@ffi.no) is a scientist at FFI where she has been working since 2006. Before coming to FFI she worked with content distribution systems at UiO. She received her Cand.Scient. degree from UiO. Her research interests are Web services, quality of service, and network protocols.

ANDERS EGGEN (anders.eggen@ffi.no) received his M.Sc. degree in computer science from UiO in 1991. Since 1992 he has been employed by FFI, where he is currently a principal scientist. He has worked on many aspects of communication and application protocol evaluation and design, as well as security and SOA solutions for military systems. He has contributed significantly in technical standardizations activities and has been an editor of both IETF and NATO standards.

Communications | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

Web Services Discovery across Heterogeneous Military Networks

Frank T. Johnsen, Trude Hafsøe, and Anders Eggen, Norwegian Defence Research Establishment (FFI)
Carsten Griwodz and Pål Halvorsen, Simula Research Laboratory

ABSTRACT

NATO has identified Web services standards as a key enabler for interoperability between the different military systems used by various NATO nations. Compared to many civilian systems, military networks vary greatly in terms of computing resources, network bandwidth, mobility and stability, and distributed applications use several different networks concurrently or interact across them. In such dynamic and heterogeneous environments, runtime service discovery is a necessity. According to the W3C, "discovery is the act of locating a machine-processable description of a Web service-related resource that may have been previously unknown and that meets certain functional criteria." In this article we present our approach to service discovery, where we combine Web services standards and proprietary solutions using our prototype interoperability gateway. This approach has been experimentally evaluated in a military experiment featuring both mobile ad hoc networks and fixed infrastructure networks, and the results show that transparent discovery between proprietary solutions and Web services discovery standards can be achieved.

INTRODUCTION

The North Atlantic Treaty Organization (NATO) Network Enabled Capability (NNEC) Feasibility Study [1] focuses on the needs of future interoperable military operations. An information infrastructure supporting such operations will have to allow for communication across system and national boundaries while at the same time taking legacy systems into account. A key concern is that open standards and available products should be used when possible, to both ensure interoperability and reduce costs.

Web services technology is becoming increasingly popular for implementing loosely coupled service-oriented civilian systems. Interoperability is the main concern; thus, Web services are based on standards. NATO has identified Web services standards as the enabler for interoperability between the different military systems of the various NATO nations. Interoperability is required on several levels: network hardware compatibility (NATO has workgroups that are addressing this), network protocol compatibility (IP has been chosen by NATO), middleware[1] compatibility, and so on. In this article we focus on Web services technology, that is, the middleware aspect of interoperability.

In the migration toward NNEC, NATO has adopted the concept of service-oriented architecture (SOA) in order to promote interoperability and efficient information exchange in the NATO organization and between the nations. The NATO Core Enterprise Services Working Group (CESWG) has defined a set of NNEC core services, which may be regarded as the foundation for more advanced services in NNEC. Service discovery is one of these core enterprise services to be used for discovering services in both static and dynamic networks.

The information infrastructure in NNEC will be built as a federation of systems provided by the different member countries. In order to achieve interoperability between these systems, the interfaces need to be defined by open standards and profiles. The NATO policy is to adopt civilian standards as far as possible in order to be able to procure commercial off-the-shelf (COTS) products. However, if the military requirements cannot be met by civilian standards, specialized military solutions need to be developed and issued as NATO standardized agreements (STANAGs).

In NATO military operations, information often needs to traverse several different types of networks with different characteristics with respect to data rate, error rate, and frequency of disconnections. Thus, different service discovery solutions need to be chosen for use in the different network types. For fixed networks, it is likely that service registries based on civilian standards like Universal Description, Discovery and Integration (UDDI) or electronic business using Extensible Markup Language (ebXML) may be chosen. For more dynamic networks, other solutions are considered by NATO CESWG. However, it is important that the service discovery solutions chosen can interoperate in order for the service consumers to be able to search for services across different operational levels.

[1] Middleware *is an abstraction layer that can conceal the heterogeneity of applications, operating systems, communication systems, and hardware in a distributed system by providing a common interface to applications. For an introduction to middleware in the context of disadvantaged grids, see [2]. Web services technology is a middleware for building loosely coupled distributed communication systems.*

 0163-6804/10/$25.00

Realizing a Service-Oriented Architecture with Web Services

The core concept in any SOA is that of a *service*, which is a capability offered through a standardized interface. When using standard Web services, the service descriptions are formatted using the Web Services Description Language (WSDL), and the communication is done using SOAP. Figure 1 shows the three central entities in a Web services deployment, and how these communicate with each other. The service provider has a capability that it wants to make available to others, and it does this by offering this capability as a service. In addition to the service itself, the service provider also has to supply a description of the service. Although the technologies used to describe and realize the interface are standardized, the interface and descriptions themselves are often created ad hoc and are not standardized.[2] The service provider can make its services known to potential users by *publishing* the service using a service discovery mechanism. The responsibility of the service provider is limited to providing the service discovery mechanism with the service description, while the distribution of this description is handled by the service discovery mechanism. A service consumer wanting to use a service will first *find* the services that are available to it using the service discovery mechanism. The consumer retrieves the service description, and uses the information therein to invoke the service directly from the service provider.

When Web services operate over military networks, it is important to consider that Web services messages are relatively large compared to typical bandwidth capabilities and other network resource capacities of radio-based military networks. In our earlier work, we have addressed some of the challenges related to service invocation in military networks, more specifically which optimizations are needed for Web services to function in low capacity networks [3], and how to achieve delay tolerance for cross-network invocations [4]. In this article, we give an introduction to Web services discovery in the context of military networks. The current standards are discussed, and we give recommendations for their use across heterogeneous networks.

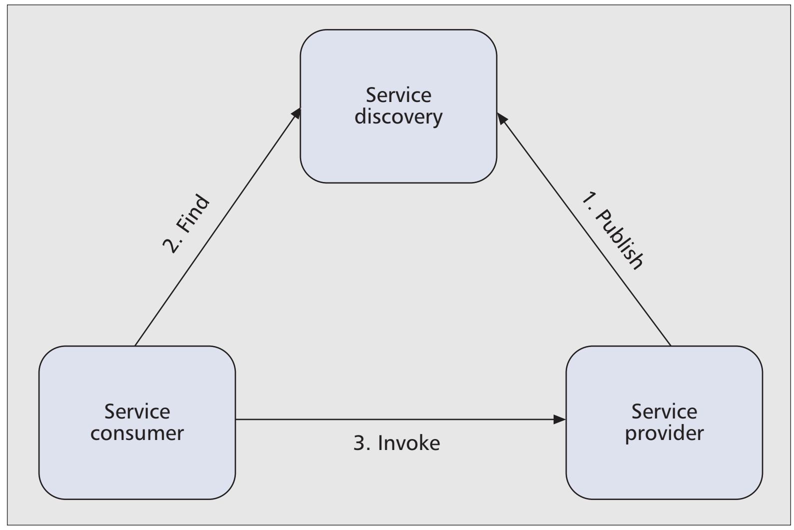


Figure 1. *Web services communication flow.*

Service Discovery

The purpose of doing service discovery can range from finding devices such as printers on a LAN, to finding trading services such as e-commerce providers on the web. In military networks discoverable services would include capabilities such as friendly force tracking and various sensors. To discover services, one can choose from different mechanisms, ranging from a network address or universal resource identifier (URI) provided by out-of-band means, via a decentralized peer-to-peer (P2P) based registry, to a centralized registry, as we discuss in our report [5]. In this article we use the term *service advertisement* to denote the service description itself. An important part of the service advertisement is the service contract, which includes the service interface. By specifying service interfaces abstractly and using a programming-language-independent XML-based interface protocol (WSDL), which provides flexibility in the way services are implemented without requiring changes in the interfaces, Web services achieve loose coupling between interfaces and implementations. Loose coupling makes it possible to reuse existing services as building blocks in new systems. Also, it enables replacing existing services with new, more efficient implementations without having to change other parts of the architecture. Using a service discovery mechanism at runtime provides late binding between consumers and services, which provides us with a dynamic architecture.

Design Time vs. Runtime Discovery

Service discovery and selection can be performed at design time or runtime. Design time service discovery occurs, for example, when a client software developer browses the web for services to include in client software, or an administrator searches for a service address to include in a deployment descriptor for an application. In other words, steps 1 and 2 in Fig. 1 are performed once at design time, and only step 3 is performed at runtime. Most modern integrated development environments include tools that automatically generate Web services clients written in Java or in a .NET programming language. As input, the tools use the service's WSDL file. The service endpoint address is either manually configured in the client software or automatically bound at service invocation time. In the case of design time discovery, the service endpoint is set once in either the configuration file or WSDL, and it does not change during subsequent invocations unless an administrator changes the address at some later point (e.g., when deploying in a different network). This is the least complex case, because there is a human in the loop to read

[2] *Using Web services in NATO requires standardizing the interfaces as well; for example, NFFI is a standard (STANAG 5527) for friendly force tracking.*

When designing a service discovery architecture for NNEC, it is important to note that constraints on network availability and topology, available services, intended users and required robustness vary with each deployment, and thus add to the complexity of such an architecture.

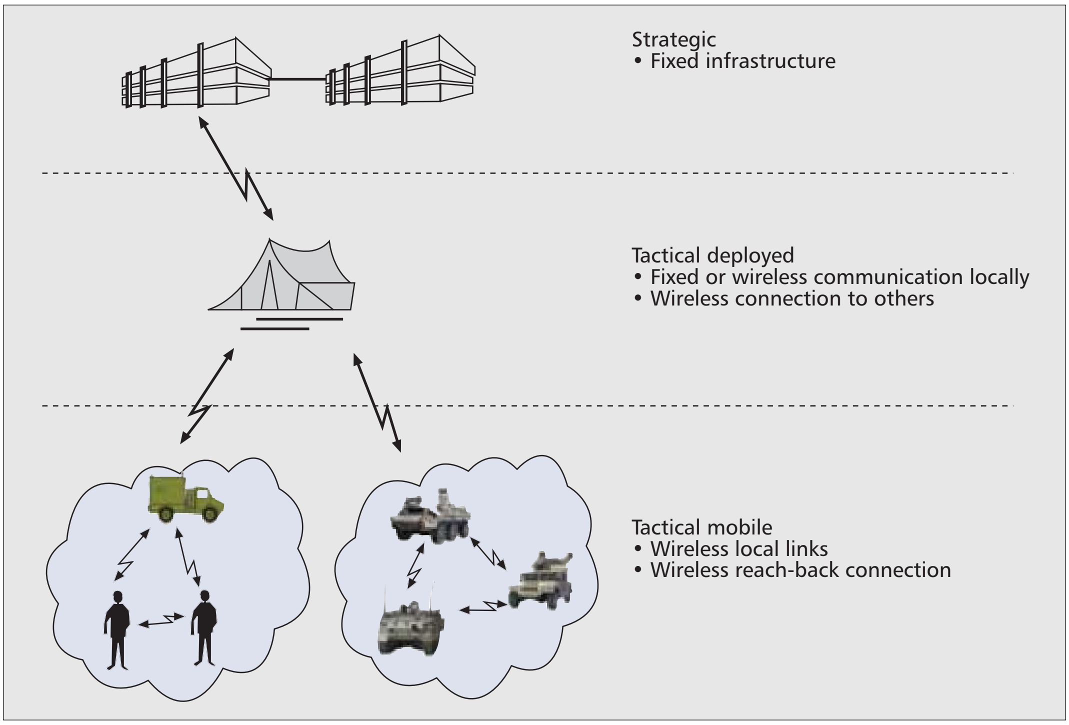


Figure 2. *Operational levels.*

and interpret the descriptions of the services' semantics.

In a military mobile ad hoc network (MANET) node mobility (i.e., nodes can move out of radio range), equipment failure, and enemy actions such as jamming can all cause nodes to disappear from the network in a nondeterministic fashion. In dynamic environments design time discovery is insufficient because services discovered at design time may not be available anymore when they are needed. We need to be able to find the current state of the network and discover services that are available at any given time. Thus, runtime service discovery is our main concern in this article. Here, steps 1 through 3 in Fig. 1 can all be performed at runtime. Runtime service discovery can be divided into two cases. The most common is the case where a client is designed to consume services of a certain interface and can only consume services that provide this interface. Since we are concerned with discovering Web services in this article, we limit ourselves to the case of finding services that are specifically supported by our clients. This means that while all the static metadata in the WSDL are valid the entire time, the address location of the service may change. This means that in dynamic networks, we should be able to discover the current state of the network and find the current addresses of all available services. A more complicated case is when new services, previously unknown to the client, are encountered. Such behavior demands that the computer is able to reason about discovered service descriptions in order to both select a service that is both proper and one it is able to invoke. This goes beyond what can be achieved with WSDL. It requires the use of machine processable semantics. In this article we focus on standard Web services.

Requirements in Military Networks

When designing a service discovery architecture for NNEC, it is important to note that constraints on network availability and topology, available services, intended users, and required robustness vary with each deployment, and thus add to the complexity of such an architecture. There are different operational levels, and each level has different communication needs and technologies. In this article, since we are focusing on middleware and communications, we divide operational levels by the types of networks, and not by different levels in the military chain of command. Using this categorization, we have three *operational levels* (Fig. 2), about which we can state the following [5]:

- The *strategic* level has a large fixed infrastructure and hosts a large number of services. In these networks there will be hundreds to thousands of nodes. Service discovery in this domain requires a solution that can scale to a large number of users and contain information about a large number of services.
- The *tactical deployed* level can use a mostly fixed infrastructure. Such networks constitute the backbones of the deployed networks (e.g., a local headquarters). However, these deployed networks need to communicate with other networks both at the same operational level and higher or lower in the hierarchy. For such communications they employ radio or satellite links, which are prone to disruptions. Thus, at this level, we have fairly large fixed networks, but some dynamicity is expected due to their interconnection with unreliable links. The networks are large, with

hundreds to thousands of nodes and services. At this level, we need solutions that can scale to a large number of users and handle some dynamicity.
- The *tactical mobile* level, the lowest level in the hierarchy comprising the so-called *disadvantaged grids*,[3] differs significantly from the two other levels higher up in the hierarchy. Here, the networks are small, each network with perhaps four to 20 nodes, and a small mission-specific set of services (sensors, positioning information, etc.). The networks use wireless links, which are prone to disruptions. At this level we need a service discovery mechanism that can handle a highly dynamic environment. It should also be resource efficient since resources are scarce.

As we can see, different networks may call for different solutions. In this article we address all three levels and the interoperability between them. In the fixed networks scalability is the most important aspect. In disadvantaged grids, on the other hand, we need solutions that can handle high mobility and at the same time minimize resource use. In these networks, conserving resources and supporting mobility is much more important than scalability, since disadvantaged grids are usually very small networks compared to those higher up in the hierarchy.

Service Discovery in Dynamic Environments

Much work has been done in terms of service discovery for civilian fixed infrastructure networks. A number of different mechanisms have been proposed and implemented for LANs (e.g., SLP, UPnP, and Rendezvous; see [6, 7] for surveys) and wide area networks (WANs) (e.g., the service-oriented P2P architecture [8]). For Web services there are three solutions: WS-Discovery for LAN, and ebXML and UDDI for WAN. We discuss the standards in the next section. Fundamental differences in terms of computing resources, network bandwidth, and issues such as mobility and stability in the network environment make these generic service discovery techniques unsuitable for tactical networks.

Gagnes [9] specifies high-level requirements for a discovery infrastructure for dynamic environments that we summarize briefly. The service discovery capability has to be *available* at all times, as losing this ability can make clients and services unable to reach each other. Thus, the service discovery mechanism cannot rely on centralized components that will constitute a single point of failure, such as a single centralized registry. In addition, the service discovery mechanism must *function independently* and not rely on other external capabilities (e.g., Domain Name Service, DNS). If additional information (e.g., XML schemas) is required for clients to be able to evaluate or use services, these should be provided without dependence on external mechanisms, for instance, by providing these through the service discovery mechanism itself. Furthermore, the discovery mechanism must handle the *liveness* issue: the services a client discovers through the use of the service discovery mechanism must accurately reflect the current state of connectivity. In other words, the services a client finds must be the same services that are available for use by the client at the time of discovery, and should not include services that are no longer available.

The issues described by Gagnes [9] are generic requirements that should be addressed by any service discovery framework in a dynamic environment. The low bit rate and frequent disruptions that characterize tactical networks put some additional requirements on the service discovery mechanisms, and proprietary protocols using cross-layer design are frequent in such networks. Cross-layer design refers to protocol design done by actively exploiting the dependence between protocol layers to obtain performance gains [10]. Cross-layer solutions have low overhead and can be of use in military tactical networks where bandwidth is the limiting factor. However, for interoperability reasons, the compatibility with legacy protocols and equipment must be taken into consideration. If one is capable of utilizing an application-level solution, that is preferable, since binding a service discovery protocol tightly to the underlying network protocols can limit its discovery capability [7].

Pervasive Service Discovery in Heterogeneous Networks

Interoperability is important not only between the different networks in a coalition force, but also between all the operational levels used within a nation. This means that we must be able to discover Web services across different networks in a coherent manner (i.e., we need support for pervasive service discovery). Zhu *et al.* [7] state that:

> *"Existing protocols have various design goals and solutions. Each has its advantages and disadvantages in different situations, so it seems unlikely that a single protocol could dominate in pervasive computing environments. With current protocols, this means clients and services can't discover each other if they don't use a common protocol. We should therefore establish a common platform to enable interoperability among service discovery protocols."*

This statement, originally made about a pervasive computing scenario, also holds true for military operations. The heterogeneity of military networks makes it difficult, if not impossible, to find a single service discovery mechanism that meets the requirements imposed by all network types. At the same time, interoperability is a key concern. This means that finding a suitable form of service discovery for military networks can be divided in two tasks:
- First, one needs to find a protocol (or set of protocols) that meets the requirements of each individual network.
- Second, one needs to find a means of interconnecting the service discovery mechanisms across the heterogeneous networks.

We can classify the different ways of achiev-

Interoperability is important not only between the different networks in a coalition force, but also between all the operational levels used within a nation. This means that we must be able to discover Web services across different networks in a coherent manner.

[3] *Disadvantaged grids are characterized by low bandwidth, variable throughput, unreliable connectivity, and energy constraints imposed by the wireless communications grid that links the nodes [2].*

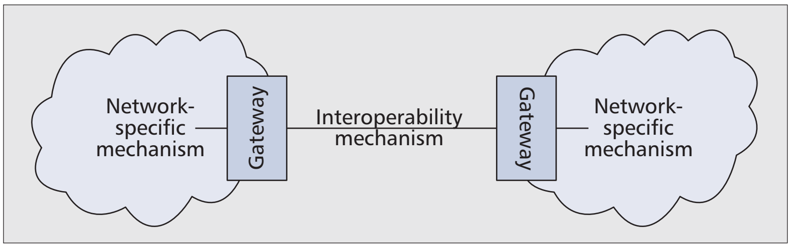


Figure 3. *The gateway approach to pervasive service discovery.*

ing pervasive service discovery in three categories [5]: adaptive discovery, layered discovery, and using service discovery gateways.

Adaptive service discovery means that one single service discovery protocol is used across all network domains. The protocol has to adapt its behavior to adhere to the capabilities of the underlying network. This protocol must be used by all applications in the network.

Layered service discovery adds a common abstraction layer to several existing discovery protocols. This means that each network domain can use a dedicated protocol, and every client must interact with the abstraction layer. The protocol that is layered on top controls and connects the different service discovery protocols.

Using *service discovery gateways*, each network domain can employ the most suitable protocol. In this case, interoperability is ensured by using service discovery gateways between the domains that can translate between the different service discovery mechanisms. Interoperability is ensured by the creation and interpretation of service descriptions in clients, servers, and gateways. The expressive power of the different service descriptions determines whether interoperability is fully or only partially possible. In the case of partially compatible service descriptions, the gateway needs to host the missing metadata in order to achieve full interoperability. For example, this is needed when interconnecting a cross-layer discovery solution with one of the Web services standards [5].

Considering these three techniques, we determine that the gateway approach (Fig. 3) is best suited for military networks because it is the only one of the three approaches that does not require changing the already existing clients and services. Gateways have the benefit of adding low complexity to the system (i.e., needing deployment only in the connection point between the networks). Low complexity is good in the long run, since it means low development and maintenance costs. An important goal for NATO is to keep costs and complexity low. Since NATO military networks use heterogeneous technology, they are already in need of one or more so-called *interoperability points* [1] to act as bridges between the networks. By deploying service discovery gateways at these interoperability points, we can achieve service discovery interoperability in a transparent way and let each network continue to use its service discovery mechanism of choice, while still supporting pervasive service discovery.

Employing the Web Services Discovery Standards

While gateways allow each NATO partner to use one or several Web services discovery standards in their networks and adapt them to partners' service discovery mechanisms through gateways, it is still common practice for many countries to adopt non-military technology where applicable. We argue that the existing Web services discovery standards can do a satisfactory job in several networks and will therefore probably be used wherever possible. In the following we discuss the existing standards' applicability for the three operational levels.

There are three standards by the Organization for the Advancement of Structured Information Standards (OASIS) related to Web services discovery: UDDI, ebXML, and WS-Discovery. Table 1 summarizes the most important differences between the three standards. UDDI and ebXML are both registries, and provide similar functionality that is suitable for both design time and runtime discovery in stable networks. WS-Discovery is a decentralized mechanism, which does not have the single point of failure that a registry does. However, being decentralized, WS-Discovery relies on flooding the network with queries, and is thus not a resource-efficient mechanism. It uses SOAP over UDP with time to live (TTL) = 1, meaning that it will discover services within one hop in a network. WS-Discovery has less expressive power than the registries and is only usable for runtime discovery. All three standards are suitable for use in strategic networks and can also be used in tactical deployed networks, as we have shown in an operational experiment. The standards are not, however, particularly usable in disadvantaged grids: the registries have liveness and availability issues, whereas WS-Discovery is bandwidth intensive and cannot function independent of external capabilities [5].

Service Discovery Beyond the Standards

As we have seen, the Web services discovery standards form a good starting point for service discovery in military networks. They can be used in both fixed infrastructure networks on the strategic and tactical deployed levels, and to some extent also in single-hop wireless networks. What the standards cannot provide is a discovery capability suitable for multihop MANETs or networks with very limited capabilities. These specific problems require tailor-made solutions, and FFI has previously developed several such solutions (e.g., Search+, Mercury), each addressing a separate set of requirements [5]. With our recent focus on Web services, we have developed a mechanism we call Service Advertisements in MANETs (SAM). SAM is an application-level solution for Web services discovery. It is designed for use in small MANETs with highly mobile nodes. SAM integrates periodic service advertisements, caching, location

[4] *Combining service discovery with node location information is based on the observation that many applications not only need to know which services are provided, but also where the mobile nodes are.*

Communications Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

	UDDI/ebXML	WS-Discovery
Type	Registry	Fully decentralized
Suitable for	WAN and LAN	LAN and single-hop MANETs
Discovery capability	Design time and runtime discovery	Runtime discovery only
Availability	Central components; can become unavailable during network partitioning	Decentralized operation ensures availability of the discovery service.
Independence	Only ebXML is independent in that it is both a registry and a repository. UDDI is only a registry.	Dependent on external resources (e.g., fetching XML schemas)
Liveness	No liveness support	Can actively query the network (reactive protocol)

Table 1. *An overview of the Web services discovery standards.*

	SAM
Type	Fully decentralized
Suitable for	LAN, single-hop, and multihop MANETs
Discovery capability	Runtime discovery only
Availability	Decentralized operation ensures availability of the discovery service.
Independence	Independent of external resources. Local repository contains necessary information. Possible to piggyback metadata on service advertisements.
Liveness	Periodic service advertisements (proactive protocol); timeout ensures that stale data is removed.

Table 2. *An overview of the experimental SAM discovery mechanism.*

information piggybacking,[4] and compression in order to be resource efficient. For an overview of SAM, see Table 2.

The SAM mechanism is designed to meet the requirements of one specific network type and usage pattern: tactical mobile networks. Combining SAM with Web services discovery standards in an interoperability gateway forms a basic toolset in which one can choose to deploy the mechanism that best meets the capabilities of a given network.

A gateway, shown in Fig. 4, periodically queries all services in the WS-Discovery and proprietary domains. The services that are available (if any) must then be looked up in the gateway's local service cache. This local cache is used to distinguish between services that have been discovered, converted, and published before, and new services that have recently appeared in each domain. If a service is already present in the cache, it has been converted and published before, and nothing needs to be done. On the other hand, if the service is not in the cache, it is translated[5] from one service description to the other, published in the network, and added to the local cache. Also, for each query iteration the gateway compares the local cache containing all previously found services with the list of services found now. If any service has disappeared from its domain since last time (i.e., the service is present in the cache but not in the current set of discovered services), the gateway deletes the service from the other domain as well by using its native service deletion mechanism. After being removed from the network, the service must also be removed from the local cache. This behavior allows the gateway to mirror the active services from one domain to the other, and remove any outdated information. Provided that the service discovery mechanism in each domain has an up-to-date view of the services in the network, this view is being propagated through our interoperability gateway. We have implemented such a gateway prototype solving transparent interoperability between WS-Discovery and a cross-layer solution, and also between WS-Discovery and SAM. The latter was successfully tested at Combined Endeavor in 2009, a joint NATO and Partnership for Peace field trial. For further details, see our report [5].

CONCLUSION

Achieving systems interoperability by using successful civilian standards, such as Web services, is a critical goal for NATO. Some military networks' characteristics are sufficiently different from the network characteristics assumed by the designers of civilian standards that the military networks cannot deploy implementations of such standards unless they are significantly modified. Different nations must

[5] *The description schemas of the various discovery mechanisms are quite different. Our current prototype implementation of the gateway mediates between those schemas using a manually implemented 1–1 mapping. For future work, we are planning on using semantic technologies for more flexible and expandable mediation.*

be able to use their own systems, and network-specific mechanisms and interfaces are necessary. An important step toward a common solution for interoperability among NATO nations is the interoperability of service discovery solutions. To enable runtime service discovery in dynamic and heterogeneous environments, we have presented SAM, which integrates functionality such as periodic service advertisement, caching, piggybacking, and compression. To mediate between proprietary and standard solutions, and across network and national boundaries, we have presented a gateway approach. Our experiments in a heterogeneous military environment show that the proposed gateway approach for interoperability is feasible and that transparent service discovery is achieved.

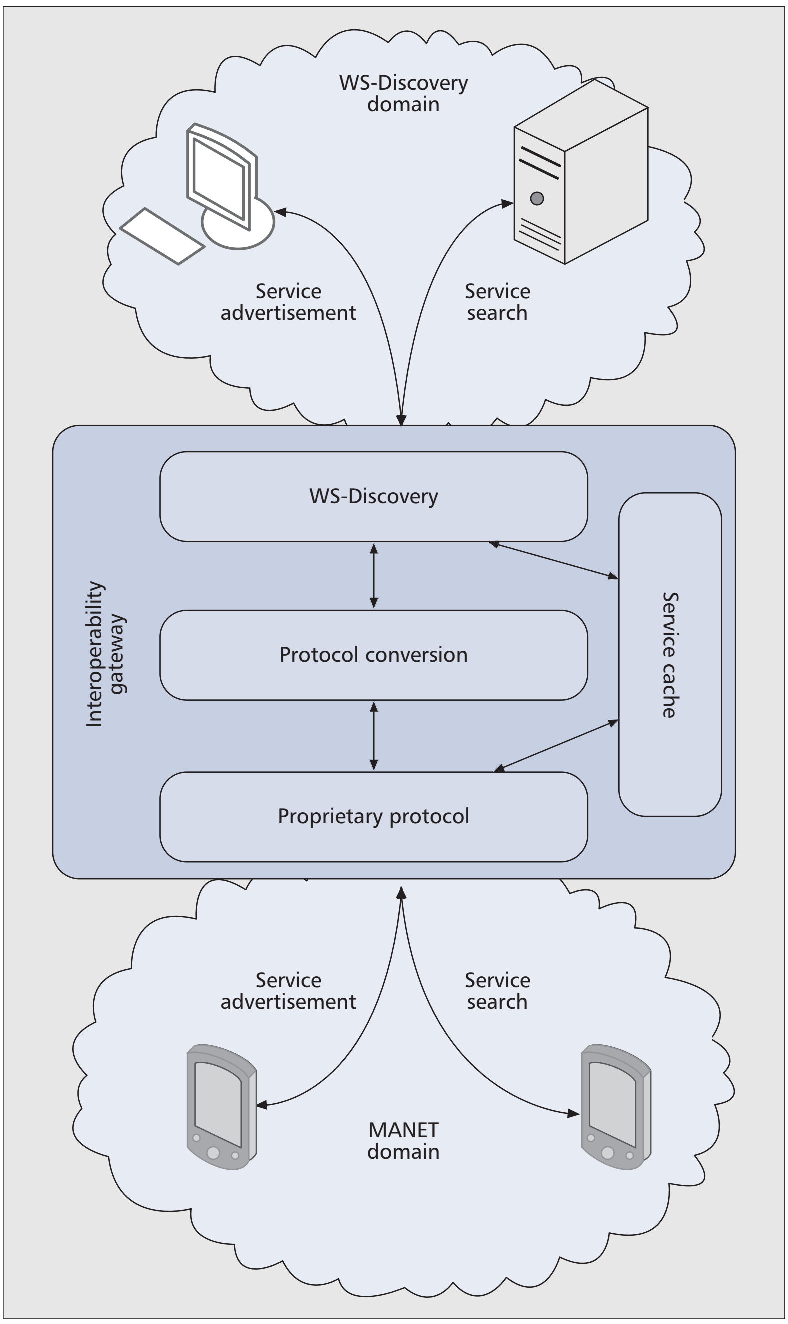


Figure 4. *Interoperability gateway.*

Acknowledgments

Thanks to our colleague at FFI, Joakim Flathagen, for fruitful discussions and feedback when we developed the initial gateway prototype, and for making Fig. 4.

References

[1] P. Bartolomasi *et al.*, "NATO Network Enabled Capability Feasibility Study," v. 2.0, Oct. 2005.

[2] A. Gibb *et al.*, "Information Management over Disadvantaged Grids," Final Report, RTO Info. Sys. Tech. Panel, Task Group IST-030/RTG-012, RTO-TR-IST-030, 2007.

[3] K. Lund *et al.*, "Using Web Services to Realize Service Oriented Architecture in Military Communication Networks," *IEEE Commun. Mag.*, Oct. 2007.

[4] E. Skjervold *et al.*, "Delay and Disruption Tolerant Web Services for Heterogeneous Networks," *IEEE MILCOM '09*, Boston, MA, 2009.

[5] F. T. Johnsen *et al.*, "Interoperable Service Discovery: Experiments at Combined Endeavor 2009," FFI report 2009/01934; http://rapporter.ffi.no/rapporter/2009/01934.pdf

[6] A. N. Mian, R. Baldoni, and R. Beraldi, "A Survey of Service Discovery Protocols in Multihop Mobile Ad Hoc Networks," *IEEE Pervasive Comp.*, Jan.–Mar. 2009, pp. 66–74.

[7] F. Zhu, M. W. Mutka, and L. M. Ni, "Service Discovery in Pervasive Computing Environments," *IEEE Pervasive Comp.*, Oct.–Dec. 2005, pp. 81–90.

[8] M. Amoretti, F. Zanichelli, and G. Conte, "SP2A: A Service-Oriented Framework for P2P-Based Grids," *Proc. 3rd Int'l. Wksp. Middleware for Grid Comp.*, Grenoble, France, 2005.

[9] T. Gagnes, "Assessing Dynamic Service Discovery in the Network Centric Battlefield," *IEEE MILCOM '07*, Orlando, FL, 2007.

[10] V. Srivastava and M. Motani, "Cross-Layer Design: A Survey and the Road Ahead," *IEEE Commun. Mag.*, vol. 43, no. 12, 2005, pp. 112–19.

Biography

FRANK T. JOHNSEN (frank-trethan.johnsen@ffi.no) received his Cand.Scient. degree from the University of Oslo (UiO), Norway, in 2002. Following this he worked with networks and distributed systems at UiO, until he started work as a scientist at the Norwegian Defence Research Establishment (FFI) in 2006. At FFI he is currently working within the area of secure pervasive SOA. His research interests include Web services, quality of service, and middleware.

TRUDE HAFSØE (trude.hafsoe@ffi.no) is a scientist at FFI, where she has been working since 2006. Before coming to FFI she worked with content distribution systems at UiO, from which she received her Cand.Scient. degree. Her research interests are Web services, quality of service, and network protocols.

ANDERS EGGEN (anders.eggen@ffi.no) received his M.Sc. degree in computer science from UiO in 1991. Since 1992 he has been employed by FFI, where he is currently a principal scientist. He has worked on many aspects of communication and application protocol evaluation and design, as well as security and SOA solutions for military systems. He has contributed significantly in technical standardizations activities, and has been an editor of both IETF and NATO standards.

CARSTEN GRIWODZ (griff@ifi.uio.no) is a professor at the Department of Informatics, UiO, and Simula Research Laboratory. He received his Diploma in computer science from the University of Paderborn in 1993, worked at IBM's European Networking Center from 1993 to 1997, and received his doctoral degree in 2000 from Darmstadt University of Technology. His interests lie in the system support for interactive distributed multimedia.

PÅL HALVORSEN (paalh@ifi.uio.no) is an associate professor at the Department of Informatics, UiO, and Simula Research Laboratory. He received his Master's and doctoral degrees in computer science from UiO in 1997 and 2001, respectively. His research activities focus mostly on resource utilization and system support for distributed multimedia systems.

IEEE Communications MAGAZINE

CALL FOR PAPERS

IEEE Communications Magazine encourages the submission of technical papers and articles on topics of interest and relevance to its readers, which includes all members of the IEEE Communications Society. The scope of the articles relates to various types of telecommunications and data communications, including:

- Technology (e.g., optical, wireless, signal processing)
- Networking (e.g., LAN, WAN, ad hoc, automotive)
- Systems
- Services and service management
- Emerging standards
- Development methods
- Deployment experiences and field trial results
- Market trends
- Regulatory and policy issues
- Significant global events and projects

Articles should be tutorial in nature and written in a style comprehensible to readers outside the specialty of the article. Articles should contain a minimum of mathematical content and avoid product marketing content. The author guidelines for additional information on content and style can be found on the magazine web site http://dl.comsoc.org/ci/. The call-for-papers links for current feature topics and regular series can also be found at this web site. Articles not directly aligning with the posted calls for papers are welcome as open call submissions. All papers and articles go through a peer review process to ensure high quality content and industry relevance.

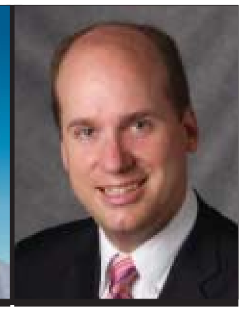
All submissions and their subsequent processing is handled through the on-line Manuscript Central tool: http://mc.manuscriptcentral.com/commag-ieee Author guidelines and forms can also be found on this site.

GUEST EDITORIAL

CARRIER ETHERNET FOR MOBILE BACKHAUL

David Hunter

Alan McGuire



Glenn Parsons

There can be little doubt that the mobile communications industry is growing at a rapid rate fueled by a myriad of new applications that have an insatiable appetite for bandwidth, combined with user expectations of improved speed and coverage. This is putting considerable pressure on the existing network infrastructure and, in particular, the mobile backhaul segment.

Mobile backhaul typically refers to the network between the base station site and the network controller site. However, mobile backhaul can also include a spectrum of networks and network technologies including the radio access network (RAN) and core networks. Network providers are faced with a number of significant challenges in delivering mobile backhaul solutions. As bandwidth demand increases to support new users and mobile applications, the need to invest heavily in new mobile backhaul infrastructure has to be carried out against a backdrop of (at best) modest revenue growth. This is further complicated by the need to support a wide variety of evolving mobile technologies (second generation (2G), 3G, and 4G), which adds complexity to migration strategies. Solutions based on time-division multiplexing (TDM) will not be able to deliver the required cost point, which will necessitate improved efficiency and lower cost of bandwidth. This is driving an evolution toward Ethernet and IP-based backhaul solutions. However, in doing so, there are capabilities inherent to TDM that need to be developed in Ethernet, such as the requirements to support synchronization at cell sites.

In this feature topic we consider the role Ethernet plays in the development of mobile backhaul infrastructure, options for migrating from the existing predominantly TDM-based infrastructure, and the latest developments in Ethernet as a technology that will allow it to do so.

The first article in this feature topic provides an overview of mobile broadband trends, the applications that are pushing the need for more bandwidth and the evolution of radio access technologies toward Long Term Evolution (LTE). Briggs *et al.* consider the options for backhaul access and aggregation, and the suitability of packet technologies in meeting these needs. The authors describe the role Ethernet transport is expected to play and how Ethernet services defined by the Metro Ethernet Forum (MEF) can be utilized.

The second article, by Ghebretensae *et al.*, examines how network providers should migrate from the large installed base of legacy TDM to one based on carrier Ethernet. The authors cite some of the main challenges including the need to significantly increase backhaul capacity to meet the demands of LTE while reducing the cost of ownership. Three migration strategies are described for dedicated RANs and another two for converged networks where providers offer both fixed and mobile traffic. Consideration is also given to network sharing where mobile operators share resources to reduce both operating and capital costs.

The article by Magee considers how synchronization should be provided in packet-based backhaul solutions to support base stations that have stringent frequency, phase, and time requirements, and also considers the challenges and deployment considerations. The move from legacy TDM to packet-based solutions has resulted in the development of new standards for synchronous Ethernet in the ITU-T, and the distribution of frequency, phase, and time in IEEE's 1588v2 protocol, both of which are described. The article indicates that these standards activities can be used in combination to provide a robust synchronization solution before considering the deployment options available to a network provider.

Related to the previous article is that of Ferrant *et al.*, which focuses on one part of the overall synchronization solution, namely the development carried out within the ITU-T on a telecom profile of the IEEE's 1588v2 Precision Time Protocol. This is expected to be the first in a suite of profiles that address particular telecom needs. The authors provide an overview of 1588v2 prior to examining the need to adapt this for the telecom environment to meet specific requirements such as timing protection, load balancing, and traceability of timing flows, which have resulted in the specification of an alternate Best Master Clock Algorithm.

GUEST EDITORIAL

The final article, by Allan *et al.*, looks at the control of large Ethernet networks using shortest path bridging (SPB), a form of link state control. SPB arose from limitations of the various spanning tree protocols in controlling large Ethernet networks, a situation that was highlighted by the increase in scalability afforded with the development of provider backbone bridging. The principles of SPB are outlined, including the two main operating modes and its ability to perform network-wide load balancing for improved network utilization. The authors provide context by illustrating SPB's application in a number of scenarios including data centers, wireless IP backhaul, and multicast clock distribution.

BIOGRAPHIES

DAVID HUNTER [SM] (dkhunter@essex.ac.uk) is a reader in the School of Computer Science and Electronic Engineering at the University of Essex. In 1987 he obtained a first class honours B.Eng. in electronics and microprocessor engineering from the University of Strathclyde, and a Ph.D. from the same university in 1991 for research on optical TDM switch architectures. After that, he researched optical networking and optical packet switching at Strathclyde. He moved to the University of Essex in August 2002, where his teaching concentrates on TCP/IP, performance modeling, and computer networks. He has authored or co-authored over 130 publications. From 1999 until 2003 he was an Associate Editor for *IEEE Transactions on Communications*, and for *IEEE/OSA Journal of Lightwave Technology* from 2001 until 2006. He is a Chartered Engineer, a member of the IET, and a professional member of the ACM.

ALAN MCGUIRE [SM] (alan.mcguire@bt.com) is a principal engineer at BT, where he leads the Ethernet Infrastructure team, which focuses on the design and delivery of optical and Ethernet backhaul and access solutions. He has been an active participant in international standards including the ITU-T, where he has held several editorial positions, and the IEEE. He graduated from the University of St Andrews in 1987 with a first in physics and an M.Sc. in medical physics one year later from the University of Aberdeen. He is both a Chartered Engineer and Chartered Physicist and a member of the IET and the Institute of Physics. He is also a visiting professor to the School of Computer Science and Electronic Engineer at the University of Essex.

GLENN PARSONS [SM] (glenn.parsons@ericsson.com) is a standards advisor at Ericsson, where he is focused on mobile backhaul technologies. Previously, in his 15 years at Nortel he participated in the development, product management, and standards specification of voice and fax messaging protocols, VoIP, Web services, and Ethernet transport. Over the past number of years, he has held several management and editor positions in IETF, IEEE, and ITU-T standards activities. He is currently a member of the IEEE-SA Standards Board committees. In addition, he is involved with Ethernet standardization activities in IEEE, ITU-T, and MEF. In MEF he is assisting the editor of MEF 22.1: Mobile Backhaul. He is a Technical Editor for *IEEE Communications Magazine* and was co-Guest Editor on feature topics on Ethernet in 2004, 2005, and 2008. He holds a B.Eng. degree in electrical engineering from Memorial University of Newfoundland.

Communications Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page qMags

Carrier Ethernet for Mobile Backhaul

Peter Briggs, Ericsson UK
Rajesh Chundury, Ericsson USA
Jonathan Olsson, Ericsson Sweden

ABSTRACT

The wide adoption of mobile broadband services by users of smartphones and other mobile terminals is being enabled by radio access technologies with better performance than many fixed residential broadband lines. Mobile network operators are upgrading their networks to deliver growing packet traffic cost effectively, while maintaining critical operational functions such as base station synchronization and resilience to faults. Carrier Ethernet technology is figuring strongly in the upgrade from the existing TDM mobile backhaul between radio base stations and mobile core networks.

FROM PHONE CALLS TO BROADBAND LIVING

The large-scale adoption of mobile broadband technology is opening new revenue and business opportunities for operators to satisfy their users' expectation of continuous connectivity. Mobile terminal suppliers have already moved beyond providing phones with simple mail and web access to more closely integrate their products with leading business and personal networking services. Current examples include Facebook, Twitter, Fire Eagle, and Google Wave, although this area of web-based applications continues to develop very rapidly.

Hence, the Internet generation is becoming accustomed to the availability of broadband access everywhere they go; indeed, of the estimated 3.4 billion people who will have broadband by 2014, about 80 percent will be mobile. Users already expect one-click access to social media applications for sharing pictures, files, and other media — and because they use these applications to support close-to-real-time social interactions, fast response and high availability are mandatory.

Changing user expectations and behavior coupled with the growing traffic from richer media delivered to new devices presents a number of challenges for the transport of mobile traffic between the mobile network's switch sites and radio base stations, as described in [1]. This article focuses on how operators will introduce Ethernet into their mobile transport infrastructure, enabling them to take full advantage of the mobile broadband revolution.

The relative network load from carrying mobile packet data traffic is already over five times that of voice, Fig. 1 shows how this situation has developed since 2007, when the volume of data traffic is represented as one load unit. In this figure, voice conversations carried as packet payloads are included in *packet data*. (The chart uses average data from Ericsson customers' wideband code-division multiple access (WCDMA)/high-speed packet access (HSPA) networks worldwide.)

Applications contributing to this data load can include voice over IP (14 kb/s), conversational video (77 kb/s), and online-streamed media from YouTube and other sites (145–260 kb/s). These figures represent today's approximate IP traffic, but we expect the media load fraction to increase with the popularity of large-screen mobile devices.

Measurements over the fourth quarter of 2009 in a typical European mobile network show that a minority of subscribers' consumption of *online media* services such as video produces the largest fraction of mobile packet data traffic, albeit by a small margin.

Figure 2 also shows that consumption of online media and web browsing are much more widely used than file sharing, although the latter accounts for 14 percent of the traffic.

The technology enabler of mobile broadband growth is of course the ongoing development of standardized radio access technologies (RATs). These use the finite radio spectrum to deliver increasing capacity between radio base stations (RBSs) of the mobile network and the served population of mobile terminals (so-called user equipment [UE]). Meanwhile, the UE population is being continually refreshed with new models capable of exploiting the latest RATs, thus encouraging even more use of mobile data.

Figure 3 illustrates the growing capability of successive RAT generations, indicating the typical peak downlink bandwidth experienced by network users; the bandwidth achievable (e.g., with Long Term Evolution [LTE]) can be increased even further by allocating additional radio spectrum resources.

Radio access technologies are specified by

the telecommunications standards bodies and associations that are now members of the Third Generation Partnership Project (3GPP), the scope of which also includes the core mobile network and service architectures:
- Global System for Mobile Communications (GSM)
- Enhanced Data Rates for GSM Evolution (EDGE)
- WCDMA
- HSPA
- HSPA evolution: higherspeed version of HSPA
- LTE, in which the radio access provides even higher downlink and uplink rates, and the lower access latency required to exploit them

So not only will the population of mobile broadband terminals served by each RBS site increase, but new RATs enable new ways to consume (or produce) increasing quantities of information. Hence, the radio access network (RAN) that connects RBS sites to mobile network switching nodes acts as a bandwidth funnel, challenged with transporting aggregated user data and network control information for multiple coexisting RAT generations.

The Radio Access Network

The widely deployed second- and third-generation 3GPP mobile systems (GSM and WCDMA) have similar RAN architectures, with two node types: RBSs and RAN control nodes. In GSM RANs these are called base transceiver stations (BTSs) and base station controllers (BSCs); in WCDMA they are the NodeBs and radio network controllers (RNCs).

In the next 3GPP generation, LTE, the RAN control functions are relocated to the RBSs and the mobile core network nodes, creating a RAN with only a single node type, called eNodeB.

Figure 4 shows these RAN architectures with the interconnecting links that carry user traffic and network control information to and from the mobile core network.

For completeness the right of the figure shows some of the key mobile core functions, not discussed further in this article:
- Mobile switching center (MSC)
- Serving general packet radio service (GPRS) support node (SGSN)
- Gateway GPRS support node (GGSN)
- LTE mobility management entity (MME)
- LTE serving gateway (S-GW)

RBS sites are connected to RAN control nodes via Abis interfaces in GSM and Iub interfaces in WCDMA. In LTE RANs the eNodeB BSs connect directly to the mobile core systems via S1 interfaces. Pairs of neighboring LTE RBSs can also optionally be interconnected by a lower-traffic X2 interface to expedite session handover as users move between RBS sites (3GPP specifies up to 32 neighbors); note that this *logical neighbor connectivity* does not imply direct *physical connections* between eNodeBs, since the traffic can be carried via backhaul network nodes.

Mobile BSs need accurate frequency synchronization in order to work together and generate the required radio coverage, with smooth handover of calls. Phase and time synchronization is also required when using time-division duplex (TDD) or for features such as multicast broadcast multimedia services in a single-frequency network (MBSFN), coordinated multipoint (CoMP) transmission and reception, or accurate detection of mobile terminal position. RBS synchronization may be achieved either by using an on-site global navigation satellite system (GNSS) receiver (GPS, GLONASS, or Galileo) or via

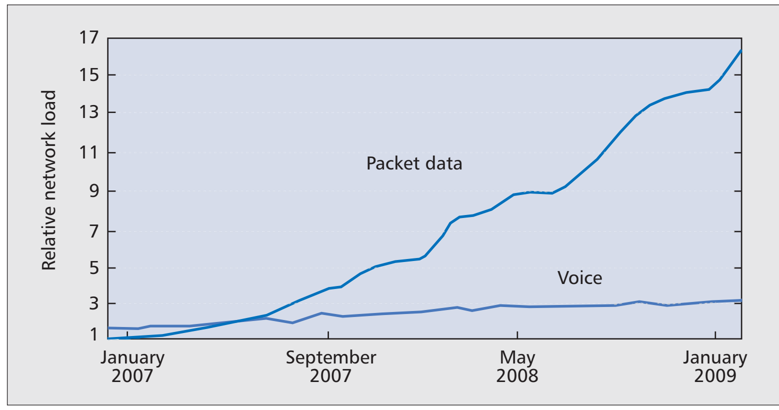


Figure 1. *Historic data traffic growth (source: Ericsson).*

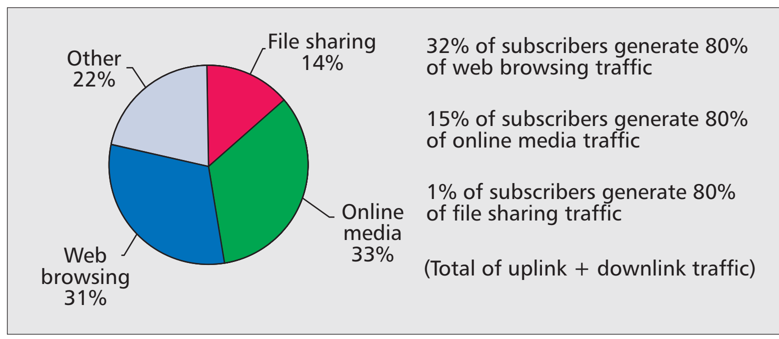


Figure 2. *Breakdown of packet data service traffic (source: Ericsson).*

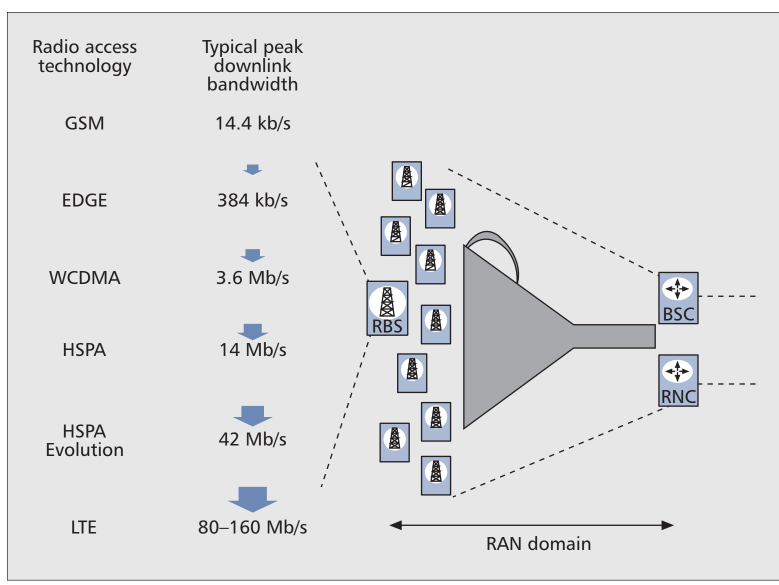


Figure 3. *The bandwidth funnel.*

IEEE Communications | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | qMags

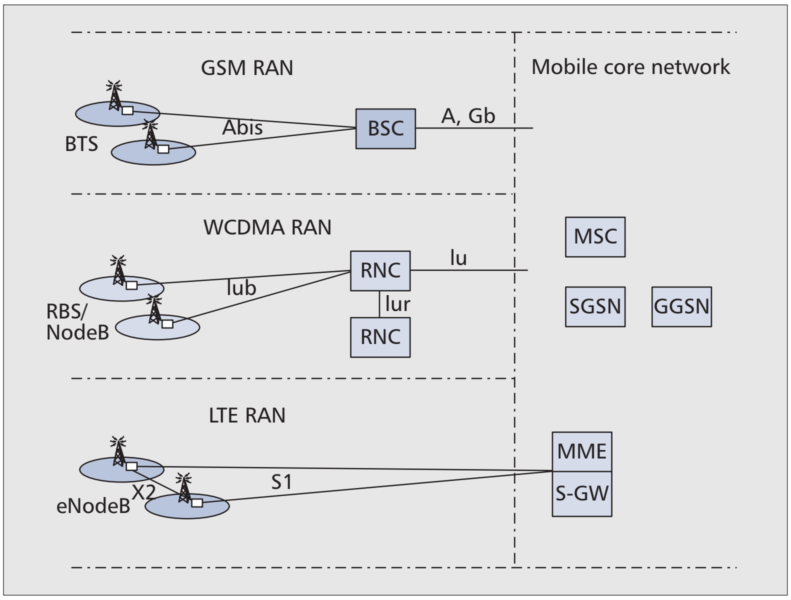


Figure 4. *3GPP RAN architectures.*

the mobile backhaul network from a timing source in the mobile core network.

Other articles in this issue describe mobile network synchronization in more detail, and a historical perspective on telecommunications networks synchronization covering plesiochronous digital hierarchy (PDH), synchronous digital hierarchy (SDH), and asynchronous transfer mode (ATM) can be found in [2].

The RAN specifications thus establish the requirements to be met by the underlying mobile backhaul transport infrastructure.

Mobile Backhaul Transport

The task of mobile backhaul is to provide appropriate transport for RAN traffic between the RBS sites and the other mobile systems shown in Fig. 4.

Connecting a large number of RBS locations typically produces a two-part backhaul design:

- Backhaul access: Connections from individual RBS sites to the mobile core. Similar to the access links serving small businesses, backhaul access can use microwave, copper, or fiber media, depending on local availability and relative costs.
- Backhaul aggregation: Transport of traffic between backhaul access links and the mobile network switch sites. Due to higher traffic levels here and typically also to protect against service outages, this is usually an optical network.

A core transport domain (out of scope for this article) then provides connections within the mobile control node buildings and between them across the network core.

Figure 5 shows how the RAN and backhaul transport segments are related to the RBS and switch sites.

Note that LTE backhaul will typically cover longer distances than that for GSM or WCDMA since, as implied on the right of Fig. 5, LTE switch site functions (MME and S-GW) will normally be located in fewer places and thus be more remote from their RBS sites than the GSM or WCDMA controllers (BSC and RNC).

Transport of Existing Backhaul Traffic

The characteristics of backhaul traffic are determined by the combination of end-user services and RAN protocols, the latter being more influential in GSM and WCDMA than in LTE. An example would be the WCDMA radio link control (RLC) protocol, which ensures reliable transport of data packets to the user terminal through retransmission of errored packets. Successive generations of RAN equipment have been designed to use backhaul transport services commonly available at the time (Table 1).

These transport links are connected to RAN systems via time-division multiplexing (TDM; T1/E1) or Ethernet interfaces and carried over connectivity services in a PDH/SONET/SDH, ATM, or Carrier Ethernet network. Microwave, copper or fiber media are all used for physical delivery of the transport services, as appropriate for each situation.

The main characteristics of the user and control traffic are:

- *User traffic*:
  - –Voice traffic: Low volume but needs low latency to support natural conversation.
  - –Video traffic: High volume, with low packet loss for streamed video, also needs low latency for video calls.
  - –Data traffic: Can be very high volume; throughput is often based on the behavior of TCP, in which case deterministic packet loss can play a part in throughput control. Relatively low latency is important to exploit the higher bandwidth of LTE.
- *Control traffic*:
  - –Network management traffic: Low volume, but high priority — vital to the operation of both the mobile network nodes and the backhaul equipment itself. Although tolerant of high latency, network management traffic must be preferentially delivered during network congestion.
  - –Mobile network control traffic: Low volume, high priority, and low latency to support RLC protocols.
  - –Packet synchronization traffic: If used, this requires low delay variation and low packet loss.

Backhaul OAM and Availability

Like other wide area networks, a mobile backhaul network requires operations, administration, and maintenance (OAM) functions initially to provision and then monitor the connections it provides. Any failures or degradations must be detected and corrected in a timely manner — generally, the allowable repair times are inversely related to the number of affected users, so fast replacement of failed connections (*protection switching*) is often used in the aggregation part of the backhaul network.

The practical upper limits for interruptions in mobile backhaul transport services are determined by the effect on the users of mobile terminals. LTE systems and the appli-

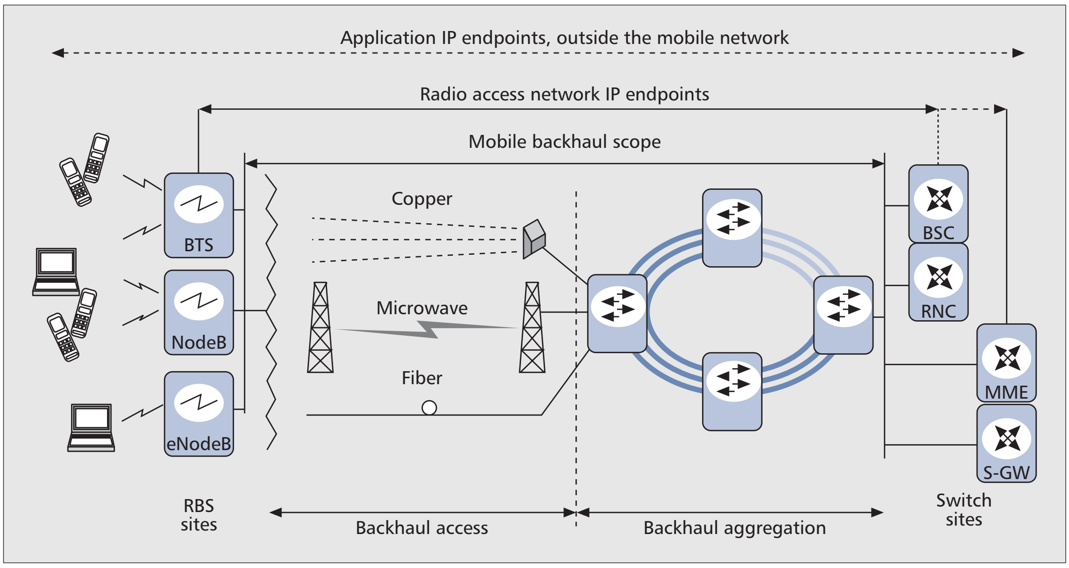


Figure 5. *Mobile RAN and backhaul transport.*

Like other wide-area networks, a mobile backhaul network requires OAM functions initially to provision and then to monitor the connections it provides. Any failures or degradations must be detected and corrected in a timely manner.

cations using them generally have less stringent requirements than GSM or WCDMA (the main performance requirements in LTE backhaul are related to inter-RBS handover and delivery of synchronization). The user impact of LTE backhaul interruptions is dependent on the design of the applications whose traffic is being carried.

Since protection relies on additional network capacity, risk analysis is used to place this so as to meet service availability targets cost effectively. Backhaul access connections to individual RBS sites are typically not protected, unlike the backhaul aggregation network, where a single failure could affect a larger number of users. Here we find carrier-class networking nodes (with redundant power, control, and switching subsystems) interconnected by multiple diverse paths.

These requirements for backhaul of the evolving mobile RAN generations are driving the evolution of future backhaul architectures as described in [3].

Backhaul over Packet Networks

Packet Connectivity and Traffic Performance

Unlike TDM transport, which provides connections of constant throughput (as originally required by the voice traffic for which it was optimized), packet-based transport networks can be a better match for delivery of the variable throughput data traffic required by many mobile broadband services. Although most end-user applications are tolerant of variations in delay and even of some packet loss, a satisfying user experience still requires the network to provide deterministic delivery performance. This can be achieved by using well established packet network design techniques to allocate resources to predefined classes of traffic, coupled with mechanisms to limit the traffic load offered to each class.

Mobile system	Typical backhaul transport
GSM	TDM
WCDMA Rel. 3, 4	ATM
WCDMA from Rel. 5	ATM & IP
CDMA 1X	HDLC/TDM
CDMA 1X EV-DO	IP
LTE	IP

Table 1. *Backhaul transport services used by mobile systems.*

Packet backhaul thus breaks free of the rather restrictive capacity steps of TDM and has the performance needed to deliver the user experience promised by mobile broadband. The performance of a packet transport connection is described by both the information throughput (bits per second) and the impairments it will introduce, explained later in this article.

High-bandwidth packet connections can of course be provided over the same range of physical media that are used for backhaul today: microwave, copper, or fiber. Considerable development is underway [4, 5] to ensure that all these options remain open to the mobile operator, to fulfill the practical needs for cost-effective backhaul access in any situation.

Mobile Backhaul over Metro Ethernet Forum Services

From a layered viewpoint, RAN systems can be seen as entities in a *client* layer, using standard Ethernet interfaces to connections provided by a packet backhaul network *service* layer — whether the latter is provided by the operator's own

Only traffic frames that meet the committed profile qualify for the defined performance, excess information rate frames can be carried by the network if it has sufficient capacity at the time, but with no performance guarantee.

transport equipment and physical media or leased from a connectivity service provider.

The Metro Ethernet Forum (MEF) has defined a set of Ethernet services that can be provided by the operator of a metro Ethernet network (MEN) and furthermore has described how these can be used to provide suitable transport for mobile backhaul [6].

The MEF defines Ethernet virtual connections (EVCs) as transport relationships between two or more user-network interfaces (UNIs). Each EVC has a set of attributes that defines the behavior of the Ethernet service between the UNIs that belong to it.

There are three MEF EVC service types, defined in [7]: point-to-point (E-Line), multipoint-to-multipoint (E-LAN), and rooted-multipoint (E-Tree). An E-LAN service connects multiple UNIs with any-to-any bidirectional connectivity. An E-Tree provides restricted communication, depending on whether its UNIs are individually configured as roots or leaves of the tree. E-Tree roots can communicate bidirectionally with all other UNIs in the E-Tree service instance, whereas leaves are limited to bidirectional communication with roots.

For example, an E-LAN service is a good choice for providing LTE RAN connectivity, since the eNodeB sites can use multipoint connectivity for the X2 interfaces to their radio neighbors and also for the S1 interfaces to the core nodes, MME and S-GW.

As networks are upgraded, RBS sites are typically required to support multiple RATs and so use a combination of service types to meet their different backhaul needs. MEF services may be used to create a variety of connectivity options between sites by defining which VLAN ID value, or set of values, identify each service EVC; the service bandwidth profile; and the service performance. Service performance is specified using four attributes: mean frame delay, frame delay variation, frame loss ratio, and availability.

A set of performance attribute values describes a class of service (CoS) instance with which frames are associated according to the content of one or more fields in the frame (the CoS ID). This is described in MEF 10.2 [8] for each service type defined in MEF 6.1 [7]. Most mobile equipment indicates CoS membership by setting the value of IP quality of service (QoS) (a 6-bit differentiated services code point [DSCP] field in IPv4) but also sets the Ethernet priority value (the 3-bit priority field in 802.1Q), so backhaul network nodes need only examine the Ethernet priority field to associate the traffic with the appropriate CoS Instance.

Bandwidth profiles describe the amount of traffic that can be carried by an Ethernet service instance, thus enabling the transport network to control its own load and protect client equipment, such as the RBS, connected to multipoint services from overload.

Bandwidth profiles are configured using four service attributes: committed information rate, excess information rate, committed burst size, and excess burst size. Only traffic frames that meet the committed profile qualify for the defined performance, excess information rate frames can be carried by the network if it has sufficient capacity at the time, but with no performance guarantee. The frames of such excess traffic can be marked at ingress to indicate their non-qualified status to other transport network nodes. This allows heavily loaded nodes to discard the excess frames in preference to qualified frames.

Hence, carrier Ethernet technologies combined with the MEF definitions provide a broad framework of flexible services that can be tailored to meet the requirements of mobile backhaul. The dynamic properties of Ethernet services can be employed during the migration process to adjust to the changing network requirements.

Packet Synchronization

Detailed RBS synchronization requirements for frequency, phase, or time depend on the RATs in use — GSM, wideband code-division multiple access (WCDMA) frequency-division duplex (FDD), WCDMA TDD), CDMA2000, LTE FDD, or LTE TDD — and on certain optional RAT features (e.g., MBSFN), as described earlier.

Where TDM or ATM backhaul is used, frequency synchronization can be obtained via the underlying SDH/PDH transmission networks, of which this is an inherent feature. Operators must maintain RBS synchronization as packet networks replace TDM for backhaul, even though packet networks do not generally need synchronization between nodes for transmission of traffic between switches or routers (in fact, their clocks can typically free-run at up to 100 parts per million [ppm] away from their nominal frequencies without a problem). RBS synchronization requirements are around 2000 times tighter than these packet network tolerances — at around ±50 parts per billion (ppb). Hence, packet backhaul does not naturally take over the delivery of synchronization from the TDM links it is intended to replace.

Several standard solutions have therefore been designed to deliver frequency synchronization over packet-based RANs and their backhaul networks, from centrally located network sources, typically at the switch sites.

Frequency synchronization may be achieved by transmitting time-stamp messages in packets — as described in IEEE 1588 (PTP) or RFC 4330 (SNTPv4). The performance of packet-based frequency synchronization using either of these alternative methods end-to-end (i.e., without timing support from intervening packet nodes) is similar — and depends mainly on three factors:

- The quality of the RBS oscillator
- The packet delay variation in the backhaul network
- The ability of the RBS clock-recovery algorithm to filter packet delay variation

IEEE 1588 specifies optional functionality in intermediate nodes to improve the accuracy of the timestamps delivered to the RBS equipment.

Alternatively, the signal clock of the physical Ethernet layer can be used to pass frequency synchronization between packet equipment nodes to enable the construction of a synchronization hierarchy similar to those used in TDM networks. This approach is known as synchronous Ethernet and is defined by the International Telecommunication Union — Telecommunication Standardization Sector (ITU-T) in G.8261, G.8262, and G.8264.

Backhaul network equipment that includes explicit features such as the optional PTP intermediate clocks or synchronous Ethernet to support timing distribution can reduce or eliminate the synchronization disturbance caused by packet delay variation. However, this support will not always be available in heterogeneous or multi-operator backhaul scenarios. In many cases, therefore, the best frequency synchronization solution can be the end-to-end packet-based timing-regeneration technologies, independent of the underlying transport layer.

A further alternative is to use GNSS receivers in the RBS sites to provide accurate time references for synchronization. Installing a GNSS antenna and receiver might not be practical everywhere that synchronization is needed, and this alternative relies on continued availability of the GNNS service, outside the control of the mobile network operator. Enhancements to IEEE 1588 standards will reduce this dependency on GNSS by adding intermediate boundary or transparent clocks in the synchronization distribution path. The use of these features in mobile backhaul will require careful analysis of operational and performance aspects.

Packet OAM and Availability

Packet backhaul typically uses recovery mechanisms triggered by OAM fault monitoring features in both Ethernet and IP network layers, configured to implement a designed recovery action for a set of anticipated disruptive events. Standards for fault management and performance monitoring of connections in Ethernet networks are found in IEEE 802.1ag and ITU-T Y.1731, while the MEF is working on frameworks (e.g., MEF 17) for their implementation in metro Ethernet networks.

Ethernet service OAM provides up to eight hierarchical maintenance domains of related maintenance associations (MAs) containing two types of OAM entity: MA endpoints (MEPs) at the periphery and MA intermediate points (MIPs) at intermediate nodes. IEEE 802.1ag defines three OAM messages, for continuity check, linktrace, and loopback, which have been extended for performance monitoring by work in the ITU-T (Y.1731). Ethernet networks can include connectivity protection switching, triggered by the OAM functions, in either linear or ring topologies.

In general, service resilience must be carefully designed into a packet backhaul network using a combination of network and internal node redundancy. Part of the network design process is to decide how each type of expected disruption will be handled by the network.

Practical Transformation

To track the upgrade of RBS systems, today's TDM mobile backhaul infrastructure must be transformed to support IP traffic delivered via Ethernet interfaces. Residual TDM traffic can optionally be emulated over the packet backhaul network.

The decision to *build* (using microwave, optical, or copper physical media) or *lease* backhaul links depends entirely on the environment, regulations, and other restrictions of each geographic area. It is not possible to generalize the outcome of such a network design, except for some observations on common approaches:

- Microwave will probably continue to be the most common technique used for backhaul access links. Multihop microwave is often used too, especially in rural areas.
- Copper (DSL) and fiber (point-to-point and passive optical network [PON]) access can also be used, as determined by local opportunities and costs. The proportion of fiber-connected RBS sites is increasing in higher-traffic areas.
- Optical transport is usually justified by the higher bandwidth and stringent protection required in the backhaul aggregation network.

Replacing TDM by Ethernet backhaul while maintaining user services requires a planned set of migration steps in each network region.

A first step can be to carry only the growing broadband traffic over packet backhaul, continuing to deliver other services via TDM. Either a separate overlay packet backhaul network can be built, or hybrid networking equipment (either microwave or optical) can be used — simultaneously supporting native TDM and native packet traffic. So-called *dual-stack* RAN equipment upgrades are available for just this purpose. In the backhaul aggregation network, where TDM traffic is already carried over optical networks, packet traffic can continue as a separate overlay for the mobile RAN, or join other packet traffic in the operator's metro network if this makes operational sense.

The market for leased transport services is adapting to address mobile backhaul; already service bundles of TDM (T1 or E1) and packet (E-Line) are available, with a roadmap to packet-only services with integrated packet synchronization.

Conclusion

RBS systems providing mobile broadband services need IP connectivity to the mobile switch sites. The TDM transport links that carried mobile voice traffic for many years are now being augmented or replaced by packet transport based on Ethernet technology delivering higher capacity at a lower cost per bit.

Mobile transport networks are being enhanced to provide Ethernet connectivity services according to Metro Ethernet Forum definitions, over microwave, copper, and fiber, while retaining the traditional transport characteristics of deterministic performance, monitoring, and

The market for leased transport services is adapting to address mobile backhaul; already service bundles of TDM (T1 or E1) and Packet (E-Line) are available, with a roadmap to packet-only services with integrated packet synchronization.

resilience to faults to ensure that the user experience lives up to expectations.

Since no single backhaul solution applies in every situation, operators will plan pragmatic transformation steps for each network region, using either self-built or leased capacity. Migration options will include overlay or hybrid networks and dual-stack RAN technology.

So whether backhaul capacity is built or leased, Ethernet will figure strongly in the transformation required to deliver the promise of the mobile broadband life.

References

[1] S. Chia, M. Gasparroni, and P. Brick, "The Next Challenge for Cellular Networks: Backhaul," *IEEE Microwave Mag.*, vol. 10, no. 5, Aug. 2009, pp. 54–66.

[2] S. Bregni, "A Historical Perspective on Network Synchronization," *IEEE Commun. Mag.*, vol. 36, no. 6, June 1998, pp. 158–66.

[3] T. Tjelta *et al.*, " An Evaluation of Future Mobile Networks Backhaul Options" *5th IEEE ICWMC*, 2009, paper #: 978-0-7695-3750-4/09.

[4] S. Little, "Is Microwave Backhaul up to the 4G Task?" *IEEE Microwave Mag.*, vol. 10, no. 5, Aug. 2009, pp. 67–74.

[5] G. Varghese and D. Ghosh, "Wireless Backhaul for LTE — Service OAM Considerations," *3rd IEEE Int'l. Symp. Advanced Net. Telecommun. Sys.*, Dec. 14–16, 2009, pp. 1–3.

[6] MEF, "Mobile Backhaul Implementation Agreement — Phase 1," Tech. Spec. 22, Jan. 2009.

[7] MEF, "Ethernet Services Definitions — Phase 2," Tech. Spec. MEF 6.1, Apr. 2008.

[8] MEF, "Ethernet Services Attributes — Phase 2," Tech. Spec. 10.2, Oct. 27, 2009.

Biographies

PETER J. BRIGGS (peter.briggs@ericsson.com) works in IP and broadband strategy and business development in Ericsson's network business unit, focused on the transformation of user services and overall network architectures for packet technologies. He has long experience in communications engineering from the management of PABX system design, through network design and operation to strategy and CTO roles covering research in the United Kingdom and Europe, communications evolution, and broadband strategies. For many years he has been enthusiastic about the key long-term societal role of communication.

RAJESH CHUNDURY (rajesh.chundury@ericsson.com) oversees strategy and customer requirements for broadband access, transport and IP solutions at Ericsson's Fixed Broadband and Convergence Engagement Practice in North America. He previously worked in the North American Global Services organization where he served as the National Transport Network Design lead for Cingular's UMTS RAN deployments, along with other similar assignments for T-Mobile USA, Sprint-Nextel, and other operators. He joined Ericsson in 1996, spending his first six years with Ericsson Research and Development.

JONATHAN OLSSON (jonathan.olsson@ericsson.com) works in Long Term Evolution product management focused on mobile backhaul and transport functionality. After several years researching multiservice communication networks, he became an active member of the Metro Ethernet Forum where he later became the editor of MEF 22, the Mobile Backhaul Implementation Agreement. In addition to his leadership in the MEF, he has roles covering mobile network design and evolution strategies.

Mobile Broadband Backhaul Network Migration from TDM to Carrier Ethernet

Zere Ghebretensaé, Ericsson Sweden
Janos Harmatos, Ericsson Hungary
Kåre Gustafsson, Ericsson Sweden

ABSTRACT

With the rollout of Long Term Evolution the capacity of the radio access network backhaul needs to be upgraded to 100–150 Mb/s. Next-generation mobile networks, such as LTE Release 10, will increase the requirement for backhaul capacity to gigabits per second. In order to increase network utilization and decrease operating expenses, carrier Ethernet transport infrastructure (MPLS and carrier grade Ethernet) will be deployed and maintained at a lower total cost of ownership than legacy TDM transport infrastructure. This article discusses different migration scenarios from the circuit-switched legacy backhaul networks toward packet-based networks.

INTRODUCTION

Mobile networks are undergoing major changes. The main driving force behind this is the introduction of mobile broadband services. The narrowband circuit-switched data networking that supported the deployment of general packet radio service (GPRS) in second-generation (2G) systems has now, with the deployment of high-speed packet access (HSPA) in third-generation (3G) systems, evolved to broadband data networking that can support various multimedia services. A few years after its deployment, the volume of HSPA packet data traffic has exploded to the point where it has now exceeded circuit-switched voice traffic. This huge take-up of broadband data services has led to a major shift in the composition of mobile traffic from voice-dominated circuit-switched traffic to packet-switched data traffic. The deployment of Long Term Evolution (LTE), which can support a theoretical peak downlink data rate of 330 Mb/s (i.e., peak rate of LTE radio base stations [RBSs] with 4 × 4 multiple-input multiple-output [MIMO] antenna configuration), will further increase the ratio of packet data traffic to TDM traffic in the backhaul. Besides, operators also need to consider next-generation (fourth generation, 4G) mobile systems when planning future backhaul networks. LTE Release 10, which is now being specified by the Third Generation Partnership Project (3GPP) as the 4G mobile system, will support wider channel bandwidth and up to 8 × 8 MIMO antenna configurations in order to reach the targeted peak data rates of 1 Gb/s downlink and 500 Mb/s uplink [1]. Legacy backhaul networks are optimized for circuit-switched voice traffic in which the transmission from RBSs to a base station controller (BSC) is realized using static time-division multiplexing (TDM) circuits. Such networks are, however, not optimal for the delivery of packet traffic; therefore, supporting these huge data traffic rates while maintaining low operations expenditure (OPEX) will be one of the biggest challenges for mobile network operators.

In fixed wireline networks the narrowband data network that was deployed for residential ADSL has now evolved to broadband access, metro, and core networks supporting broadband multimedia services. As a result, fixed-mobile operators (i.e., operators who provide fixed and mobile services) have already adapted their fixed networks to cope with the huge data traffic demand in order to support fixed broadband services such as IPTV, video on demand, and high-speed Internet access. Many of these operators are now in the process of converging their fixed and mobile networks. Fixed-mobile convergence (FMC) is a framework for a common converged network capable of supporting both mobile and fixed services. By employing FMC, fixed-mobile operators will be able to use their fixed and mobile infrastructure base to leverage their service offering and reduce their OPEX. The bottom line is that fixed-mobile and mobile-only operators have to address the same challenge (i.e., supporting high data traffic while maintaining low OPEX). But since their deployed networks are different, operators need to analyze the different migration options and identify the migration steps that optimize the reuse of their existing network while lowering the total cost of ownership.

The migration to packet-based backhaul networks will also have important implications for network sharing, in which the resources of the network are shared among multiple, usually two or three, operators in order to reduce their cap-

 0163-6804/10/$25.00

ital expenditure (CAPEX) and OPEX. Although static infrastructure sharing is possible in TDM networks, large volumes of packet data traffic in the network leads to inefficient use of network resources, since high capacity TDM links have to be maintained for each operator. In contrast, packet networks enable a dynamic resource sharing, in which the un-used resources of one operator can be used by the others leading to an efficient use of the network resources.

The article is organized as follows. After a brief description of mobile backhaul services and transport networking technologies, two categories of radio access networks (RANs), dedicated native RAN backhaul and converged fixed-mobile backhaul networks, are identified. Next, dedicated RAN backhaul networks are discussed, and three migration scenarios toward an IP backhaul network are presented. This is followed by a discussion of fixed-mobile converged backhaul networks after which two migration scenarios are presented. In the last section the implication of TDM for packet backhaul migration to shared networks is briefly discussed and a conclusion drawn.

Evolution of Mobile Backhaul Networks

A mobile backhaul network connects RBSs to RBS controllers. Typically, multiple RBSs are collocated at the cell site, and the traffic from these RBSs is consolidated using site aggregation nodes. The access network, which consists of the *first mile* part of the mobile backhaul, connects the RBSs in the cell site to the aggregation network and has a similar function as the first mile links in fixed access networks. Presently, RBSs are connected using copper and microwave $n \times$ E1/T1 plesiochronous digital hierarchy (PDH) links where n is usually less than 8. These link technologies, however, cannot support LTE and LTE Release 10 capacity requirements; therefore, high-speed optical fiber, microwave, and bonded VDSL links must be used. The access network is connected to the aggregation edge nodes, whose main function is to consolidate the traffic into high-capacity optical links. Because the aggregation network supports a large number of end users, this part of the network is protected from link and node failures. This is usually achieved by deploying synchronous optical network/digital hierarchy (SONET/SDH) ring topology, which has a recovery time of 50 ms. At the switch site the traffic from the different RBSs is segregated to the different RBS controllers.

Mobile Backhaul Services

The role of mobile backhaul is to transport user plane (UP) and control plane (CP) traffic between the RBSs and RBS controllers, while honoring the quality of service (QoS) requirements of the different applications. Related to the QoS requirements, the backhaul must also support clock distribution to RBSs for frequency and phase/time synchronization, and operation, administration, and maintenance (OA&M) for fault detection, service management, and performance monitoring. Over the years, the standardized RBS backhaul interfaces have evolved from TDM and asynchronous transfer mode (ATM) to IP/Ethernet interfaces. The main driver of this evolution is the need to migrate to a low OPEX backhaul capable of supporting the increasing volume of data traffic and compensate for the declining revenue per transported bit over the backhaul. Typically, the IP packets are encapsulated in Ethernet frames, which in turn are transported either natively or over other transport technologies. This combination of low-cost Ethernet interfaces and the statistical multiplexing capability of packet-switched backhaul networks, together with the availability of standardized OA&M capabilities in IEEE 802.1ag connectivity fault management (CFM) and International Telecommunication Union (ITU) Y.1731 performance management (PM), provides the required low-cost backhaul solution. Despite the evolution toward IP traffic over Ethernet interfaces, however, backhaul of legacy TDM and ATM traffic must still be supported for a while longer.

The Metro Ethernet Forum (MEF) has defined three types of Ethernet services that employ Ethernet virtual connections (EVCs): point-to-point E-LINE, multipoint-to-multipoint E-LAN, and point-to-multipoint E-Tree services delivered over carrier Ethernet transport networking technologies. Typically, SDH, Ethernet, and multiprotocol label switching (MPLS) networking technologies are used to transport these Ethernet services in backhaul networks [2].

Ethernet over SONET/SDH

SONET/SDH is a scalable carrier-grade circuit-switched transport technology that supports fast failure recovery time and extensive OAM functionalities. Today, most incumbent operators have a large deployment base of legacy SONET/SDH networks. Legacy SONET/SDH is optimized for transport of constant bit rate voice traffic and is not efficient for transport of burst data traffic. However, the introduction of three significant enhancements, Generic Framing Procedure (GFP), Link Capacity Adjustment Scheme (LCAS), and Virtual Concatenation (VCAT), has transformed SONET/SDH into a flexible multiservice-capable transport technology:

- GFP: Maps Ethernet frames into SDH virtual containers (VCs)
- VCAT: Concatenates SDH VCs for flexible bandwidth provisioning
- LCAS: Provides dynamic hitless capacity adjustment of a VCAT group

Ethernet over next-generation SONET/SDH, which combines the capabilities of GFP, VCAT, and LCAS, is of particular interest to many mobile operators as it allows reuse of the installed base of SDH equipment, and takes advantage of the elaborate management functionalities and fast recovery features of SONET/SDH.

Because the aggregation network supports a large number of end-users, this part of the network is protected from link and node failures. This is usually achieved by deploying SONET/SDH ring topology, which has recovery time of 50 ms.

Dedicated native RAN backhaul networks are optimized for transport of voice, data, synchronization and OA&M traffic. They can be leased or self built and in most cases a mixture of the two, where all or parts of the access network is self built while the aggregation network is leased from a transport provider.

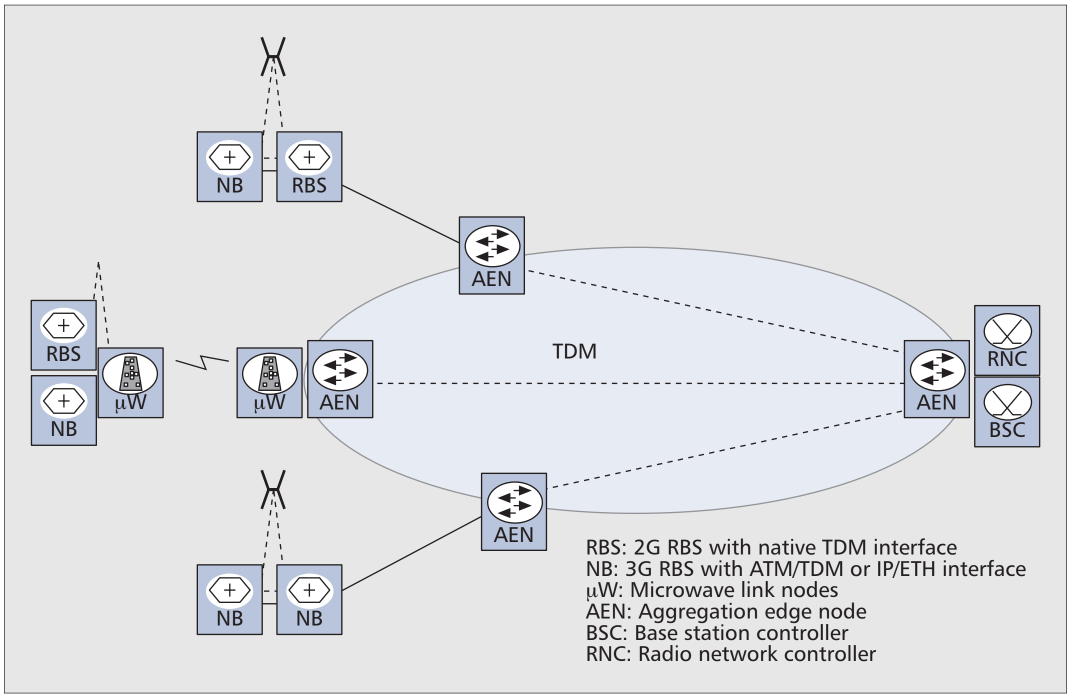


Figure 1. *Reference network for dedicated backhaul scenario 1 in which the legacy optical TDM aggregation network is used to support RBSs with native TDM, ATM, and Ethernet interfaces.*

Native Ethernet

Ethernet has been continuously enhanced to support new features in the form of virtual LAN (VLAN)-aware Q bridges, provider bridges (PBs), and provider backbone bridges (PBBs) [3]. With the completion of the PBB — Traffic Engineering (PBB-TE) specification in IEEE 802.1 Qay, Ethernet has now evolved from its origins in LAN technology to WAN transport technology supporting carrier-class features including standardized services, scalability, reliability, QoS, and OAM. This means that Ethernet E-LINE, E-LAN, and E-tree services as used in mobile backhaul can be supported using native Ethernet transport technology.

Ethernet over MPLS

MPLS is a highly scalable packet forwarding technology, which supports QoS, traffic engineering (TE), and fast recovery from link and node failures. Using pseudowire, IP/MPLS supports wire emulation for carrying ATM, Ethernet, TDM, and SONET/SDH services over packet-switched networks. Ethernet E-Line and E-LAN services are emulated using virtual private wire service (VPWS) and virtual private LAN service (VPLS), respectively. Many operators have now deployed MPLS to implement unified metro and core networks supporting different types of services.

TDM to Packet-Switched Backhaul Migration Scenarios

Mobile backhaul networks can be classified according to either their ownership: self-built and leased backhaul; or the type of services provided: dedicated native and converged fixed-mobile backhaul networks. Usually, an operator's network can include both of these in different parts of the network. In this article we focus on the latter classification and make a distinction between dedicated native and converged backhaul networks.

Migration Scenarios for Dedicated Native RAN Backhaul

Dedicated native RAN backhaul networks are optimized for transport of voice, data, synchronization, and OA&M traffic. They can be leased or self-built and, in most cases, are a mixture of the two, where all or part of the access network is self-built, while the aggregation network is leased from a transport provider. In the access network part, single- or multihop microwave links and optical fiber links are used to connect to the aggregation edge node in the aggregation network. In most cases the aggregation network consists of protected SONET/SDH rings, but other types of TDM-based optical transport technologies are also used to interconnect the access network to the RBS controllers. Depending on the volume of packet data traffic over the backhaul, three different migration scenarios can be identified.

Dedicated Backhaul Scenario 1

Dedicated backhaul scenario 1 is the first step to transform the operator's legacy TDM backhaul networks to support both TDM and packet traffic. In this scenario the operator uses its existing TDM infrastructure to support both TDM and packet data traffic. The reference network for scenario 1 is shown in

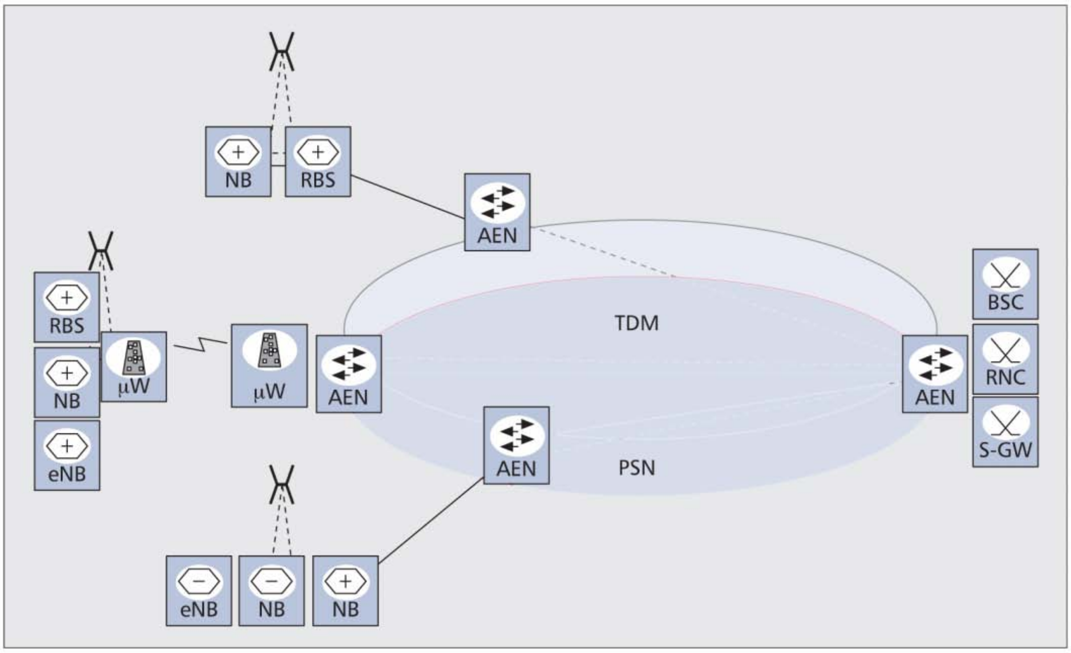


Figure 2. *Reference network for dedicated backhaul migration scenario 2, which supports TDM and Ethernet traffic using MSPP to create separate TDM and packet-switched aggregation networks.*

In the access network, the hybrid transport capability of microwave and optical fiber links can be upgraded to a higher bit rates for sites with high capacity requirements. Again, the underlying TDM transport network can be used to synchronize the RBSs.

Fig. 1. In the access network the hybrid, TDM, and Ethernet transport capability of both microwave and optical links is used to support TDM and Ethernet traffic. In the aggregation network the operator maintains its legacy TDM network and builds a high-capacity packet overlay network to support data traffic. For example, if the operator's aggregation network consists of SDH, an Ethernet overlay over SDH virtual containers (VCs) is created using GFP mapping. The inherent support for TDM clock distribution over an SDH network is also used to synchronize sites with RBSs. For example, to comply with Global System for Mobile Communications (GSM) and wideband code-division multiple access (WCDMA) frequency specifications and guarantee proper network function, the RBSs must maintain a stable and controlled radio frequency over the air interface. The frequency synchronization requirement for these RBSs is in the range of 50 ppb, and a clock delivered by TDM over an SDH transport network can typically achieve an accuracy of 16 ppb, which, with added wander and a holdover budget, is well within the requirements for the air interface.

For an operator with a TDM transport network, this is the obvious first step to transform a legacy TDM network to carry data traffic; thus, this migration scenario is already taking place in many operators' networks. Since there is no multiplexing gain, the bandwidth utilization is not optimal, so this scenario is appropriate as long as the volume and growth of packet traffic is low or moderate.

Dedicated Backhaul Scenario 2

In scenario 2 the aggregation part of the network is built of two separate transport networks, TDM and packet-based. As the volume of packet data traffic increases, operators may choose to offload low-priority high-bandwidth data traffic to a separate packet-switched aggregation network using the multiservice capability of their transport network. In most cases the operators' legacy networks consist of multiservice provisioning platforms (MSPPs) with digital cross-connection (DXC) and packet switching functionalities capable of interfacing both TDM and packet-switched networks. Using MSPPs, separate aggregation networks are deployed to support the TDM and Ethernet backhaul traffic as shown in the reference network in Fig. 2. The actual deployment of scenario 2 will depend on the availability of fibers, wavelengths, type of transport network, MSPP functionalities, and so on. For example, operators with SDH legacy networks may, depending on the expected volume of data traffic, deploy 1 GbE or 10 GbE packet-switched aggregation networks to offload low-priority data traffic using either a separate wavelength or separate fiber.

In the access network the hybrid transport capability of microwave and optical fiber links can be upgraded to higher bit rates for sites with high capacity requirements. Again, the underlying TDM transport network can be used to synchronize the RBSs. For sites with RBSs that have only Ethernet interfaces (e.g., sites with HSPA), the RBSs' synchronization signal from the TDM network can still be terminated in the sites. Alternatively, GPS or packet-based synchronization methods can be used. At the switch site, the TDM and packet traffic is connected to the appropriate controller nodes. In switch sites where several Ethernet aggregation networks are terminated, the site switch must be able to handle the traffic from all the RBSs connected to these networks.

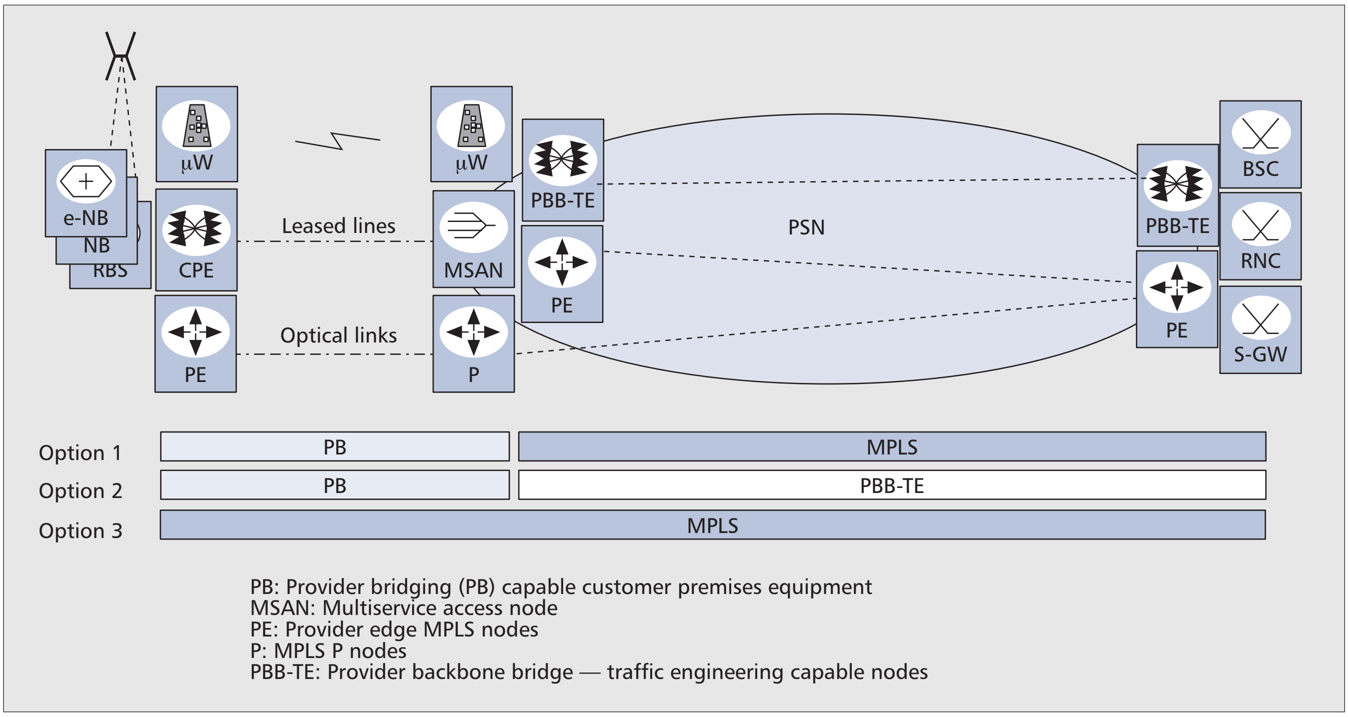


Figure 3. *Dedicated backhaul scenario 3, packet-based transport technology delivering Ethernet services and the different options for transport networking technologies: PB, PBB-TE, and MPLS.*

Dedicated Backhaul Scenario 3

In scenario 3 the mobile backhaul network only transports Ethernet traffic. This scenario applies to both greenfield operators' new network rollout and when the TDM connectivity of scenario 2 networks is retired. As shown in Fig. 3, different combinations of Ethernet (PB and PBB-TE) and MPLS transport networking technologies can be used in this scenario. The main issues in this case are the continued support of legacy RBSs with TDM and ATM interfaces, and the synchronization of the RBSs, which can be addressed in a number of ways. One straightforward method to minimize TDM support is to upgrade the RBS equipment to use IP/Ethernet interfaces. For example, GSM RBSs can use the Abis over IP solution, in which the TDM traffic that carries voice, data and signaling is mapped into IP packets using a minimum of repacking and reformatting. Similarly, RBS ATM interfaces can be upgraded to IP/Ethernet. These upgrades will enable RBSs to share low-cost Ethernet transport services and make use of statistical multiplexing gain by deploying a cell site aggregating node.

Operators who plan to make use of the multistandard radio (MSR) capability of RBSs will also benefit by upgrading their TDM and ATM RBSs to IP/Ethernet interfaces since it will enable them to create unified interfaces and a common backhaul transport solution. MSR enables the reuse of the same hardware radio unit for different radio access technologies (i.e., for GSM, WCDMA, and LTE) operating in the same frequency band.

A second alternative is to use circuit emulation and pseudowire services to support TDM and ATM traffic by emulating point-to-point and point-to-multipoint connectivity over packet-switched networks. Circuit emulation and pseudowires emulate the essential attributes of a telecommunications service, such as T1 and E1 leased lines, over a packet-switched network by supporting the minimum necessary functionality to emulate the wire or circuit. By incorporating a pseudowire or circuit emulation interworking function at the cell sites and controller sites, TDM and ATM traffic is carried over the packet backhaul network. The Internet Engineering Task Force (IETF) has specified pseudowire (PWE3) and circuit emulation service (CES) for transport of TDM and ATM services over packet-switched networks [4–6]; similarly, the MEF has also specified CES of TDM services over a carrier Ethernet network in MEF-8. TDM services have strict timing requirements; therefore, when emulating TDM services, delay in the backhaul can be reduced by encapsulating a single TDM frame into a single Ethernet frame. But doing so reduces the bandwidth efficiency of the network. Therefore, operators need to make sure that the latency and efficiency trade-off is set properly to support TDM services while still maintaining high bandwidth efficiency [7].

For synchronization of the RBSs in scenario 3, any of the packet-based synchronization methods, such as Network Time Protocol (NTP) and Precision Time Protocol (PTP), IEEE 1588 v2, or Ethernet physical layer timing infrastructure-based ITU — Telecommunication Standardization Sector (ITU-T) G.8261/G.8262/G.8264 synchronous Ethernet (SyncE), can be used. In NTP or IEEE 1588 v2, the timing information is provided by protocol-specific packets with hardware-based timestamps in combination with algorithms to determine phase information used to lock a

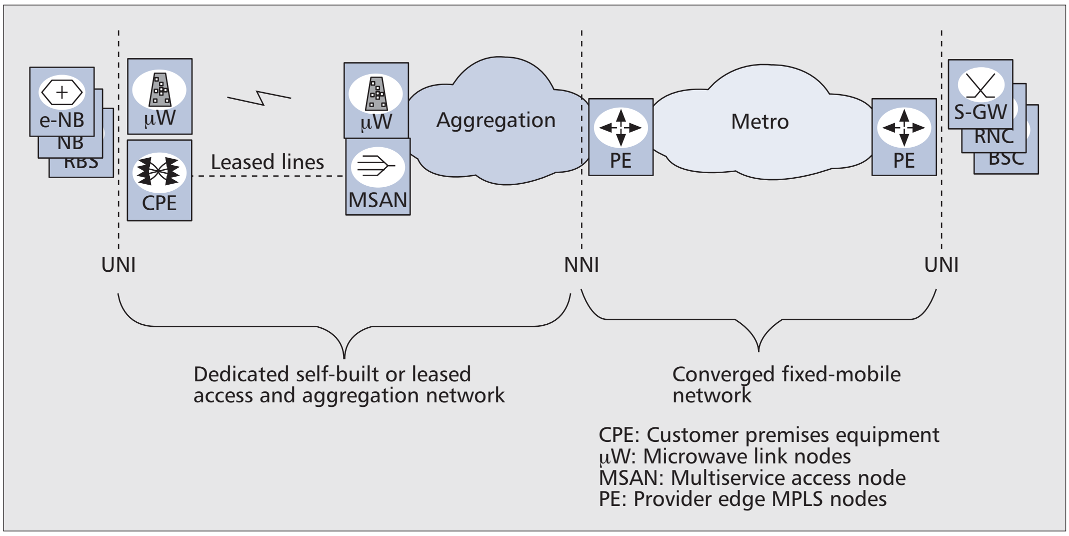


Figure 4. *Converged scenario 1, in which the operator's metro network supports fixed and mobile traffic, while the access and aggregation parts can be self-built or leased from a transport provider.*

The Network to Network Interface between the Aggregation and the metro network is Ethernet interface with a C-VLAN or S-VLAN awareness. The converged metro network can be deployed using Ethernet Provider Bridge (PB) or IP MPLS.

local oscillator. Both NTP and IEEE 1588 v2 are sensitive to packet delay variation (PDV); therefore, to ensure that the impact of the network remains as small as possible, the PDV should be kept to a minimum. The PDV will depend on QoS configuration, link speed (store-and-forward delay), maximum transmission unit (MTU) size, and so on. It is essential that synchronization traffic is handled as highest-priority traffic and by strict priority. The buffer depth of a queue handling synchronization traffic should be kept to a minimum since synchronization mechanisms are generally better at handling lost synchronization packets than synchronization traffic with large PDV.

Migration Scenarios for Converged Fixed-Mobile RAN Backhaul

Converged networks are run by incumbent operators that provide both fixed and mobile services. For a combined fixed-mobile operator, the huge take-up of packet-switched mobile traffic presents an opportunity to leverage its fixed transport services by reusing transport links and networking technologies deployed in its fixed access and core networks. In this case the mobile backhaul service, which includes Ethernet and emulated TDM services, is part of its IP/Ethernet business services. The reference network architecture for the fixed-mobile convergence network consists of first mile links, an aggregation network, and a metro network. Depending on the level of convergence, several migration scenarios can be identified; in this article we limit our discussion to two migration scenarios for backhaul over fixed-mobile converged infrastructure.

Converged Network Scenario 1

This scenario reflects the early stage of convergence, and the only converged part of the network that carries both fixed and mobile traffic is the metro network. As shown in Fig. 4, in the first mile part of the network, dedicated self-built or leased microwave links and point-to-point fiber is used, while the aggregation network is usually leased lines connecting to the converged metro network. The network-to-network interface (NNI) between the aggregation and metro networks is an Ethernet interface with C-VLAN or S-VLAN awareness. The converged metro network can be deployed using Ethernet PB or IP MPLS.

Converged Network Scenario 2

This scenario covers a completely converged network, which means the first mile links, aggregation network, and metro networks are all used to support both fixed and mobile services. The main advantage of this scenario is that the different link technologies (e.g., xDSL, GPON, EPON, 10G-EPON, and the emerging 10GPON) developed to support broadband access for residential users can now also support mobile services [8]. Similarly, microwave links can also be used when available. As shown in Fig. 5, different combinations of Ethernet (PB and PBB-TE) and MPLS transport networking technologies can be used in this scenario. TDM and ATM RBS interfaces are preferably replaced by IP/Ethernet interfaces, since the fixed broadband access is optimized to support Ethernet services; otherwise, circuit emulation and appropriate pseudowire services that span the whole backhaul network must be used. The support of packet-switched technology from the cell site to the switch site ensures high bandwidth utilization by means of statistical multiplexing gain at the cell and switch sites.

Network Sharing

Another driver for TDM to packet migration is the need for efficient network sharing. In shared networks the resources of the mobile network are shared among multiple, usually two or three, operators in order to reduce OPEX and CAPEX. In the case of fluctuating packet traffic

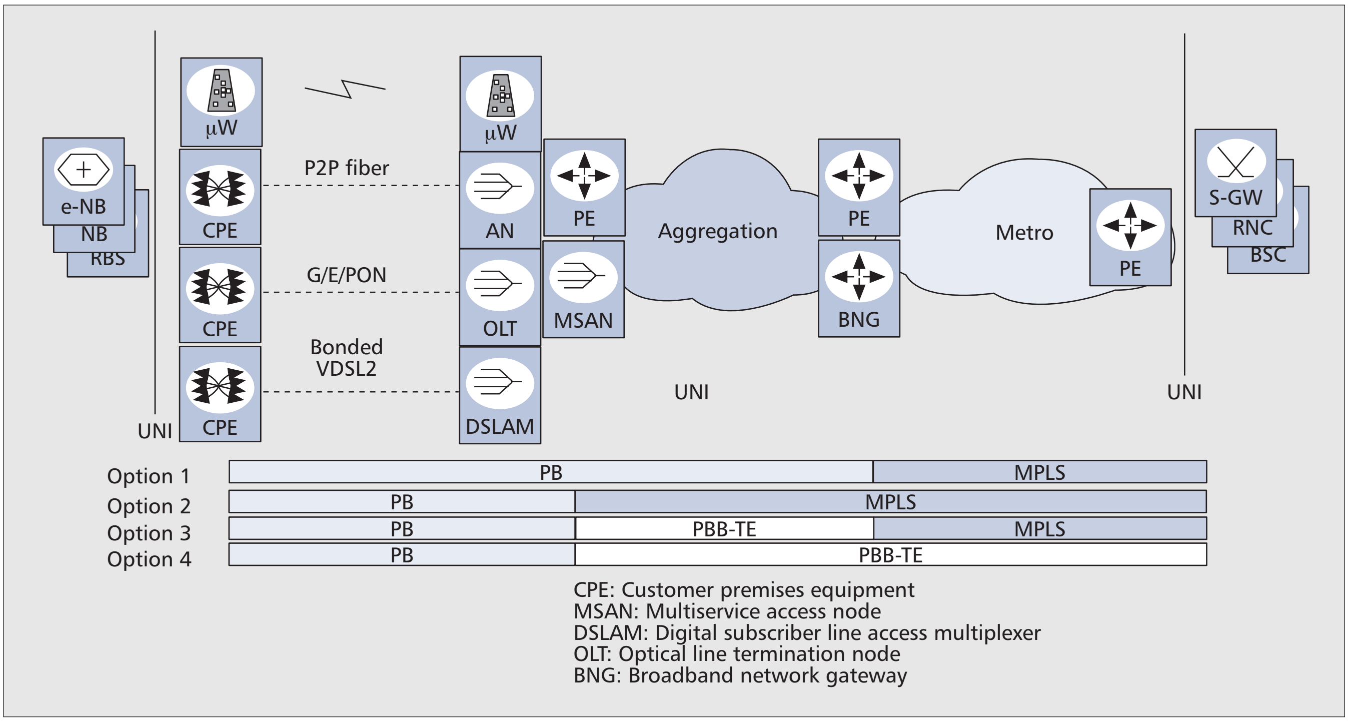


Figure 5. *Converged fixed-mobile scenario 3, packet-based transport technology delivering Ethernet services and the different transport networking technology options. Note that other options are also possible.*

with high average traffic volume, if the unused resources of one operator can be reused by another, better network utilization can be achieved. In other words, less bandwidth will be needed to handle the same amount of total traffic if sharing is possible.

In TDM backhaul networks only network infrastructure sharing is possible; for example, when the same network infrastructure (e.g., SDH links) is used by different operators, properly configured add/drop multiplexers (ADMs) groom low-capacity links from different operators to a common aggregation link. However, due to the static configuration of the TDM aggregation nodes, the unused bandwidth of one operator cannot be used by other operators, so there is no bandwidth sharing; similarly, no statistical multiplexing gain can be achieved in TDM networks. If the network supports a large amount of data traffic, this leads to inefficient network operation, since high-capacity TDM links have to be maintained to each operator, which results in high OPEX.

Pure packet backhaul networks facilitate efficient network sharing, since strict bandwidth guarantees and effective usage of free resources as well as fairness can be guaranteed at the same time. According to the MEF service specification, each operator can define its committed information rate (CIR), which determines the guaranteed capacity of the operator. However, in contrast to a TDM network, this bandwidth is not necessarily *dedicated* exclusively to a certain operator. For example, two operators, A and B, share the network resources with their respective committed information rates. Now during periods of low traffic from operator A, the unused bandwidth in the network can be used by operator B who might already have reached its guaranteed capacity. Operator B uses this *borrowed* bandwidth to transport lower-priority best effort traffic marked with drop precedence. Any time operator A wants to use its guaranteed capacity or in case of network congestion, this extra traffic from operator B is dropped first, preventing operators from starving each other while providing fairness and efficient use of the network resources.

Conclusions

The landscape of mobile networking is changing very fast. A few years after the deployment of HSPA, the volume of mobile packet data traffic has exploded to the point where it has now exceeded circuit-switched traffic. Furthermore, with the coming deployment of LTE and 4G mobile systems, the ratio of packet traffic over TDM traffic in the backhaul is expected to increase even more. To cope with the changed traffic composition, operators need to migrate their legacy TDM networks to packet-switched backhaul networks capable of supporting high volumes of packet traffic while maintaining low OPEX. Inspection of the different migration options shows that there is no silver bullet solution or single migration path that fits all types of networks. Operators need to make careful analyses of their deployed networks, present and future traffic demands, link technologies, and capability to support TDM, packet, and hybrid traffic.

Figure 6 summarizes the migration options for TDM backhaul networks and TDM RBS interfaces to packet-switched networks and IP/ETH RBS interfaces. As a first migration phase, operators have already deployed packet overlay on TDM backhaul networks to support low volumes of data traffic; as the volume of data traffic increases, operators may deploy separate packet-switched networks to offload packet traffic, maintaining hybrid TDM and packet-switched networks. With the deployment of LTE and LTE Release 10, operators may opt to migrate to a pure packet-switched backhaul network by retiring their TDM networks while upgrading their ATM and TDM RBSs to IP/Ethernet interfaces. This option has efficient network resource utilization and lower OPEX for leased lines. Alternatively, operators can deploy a packet-switched backhaul network, and use pseudowire and CES interworking functions to support their ATM and TDM RBSs.

For operators with fixed-mobile converged networks, the migration steps should be aligned with the degree of convergence assumed in the network. Fixed-mobile network convergence at the metro, aggregation, and access networks will enable operators to leverage on their xDSL, GPON, EPON, 10EPON, and emerging 10GPON access networks to support mobile services as well. The multiple backhaul solutions described provide operators with a choice of backhaul migration paths to suit their individual situations, decreasing operating expenses while increasing the transport capacity required to support graceful TDM to IP migration.

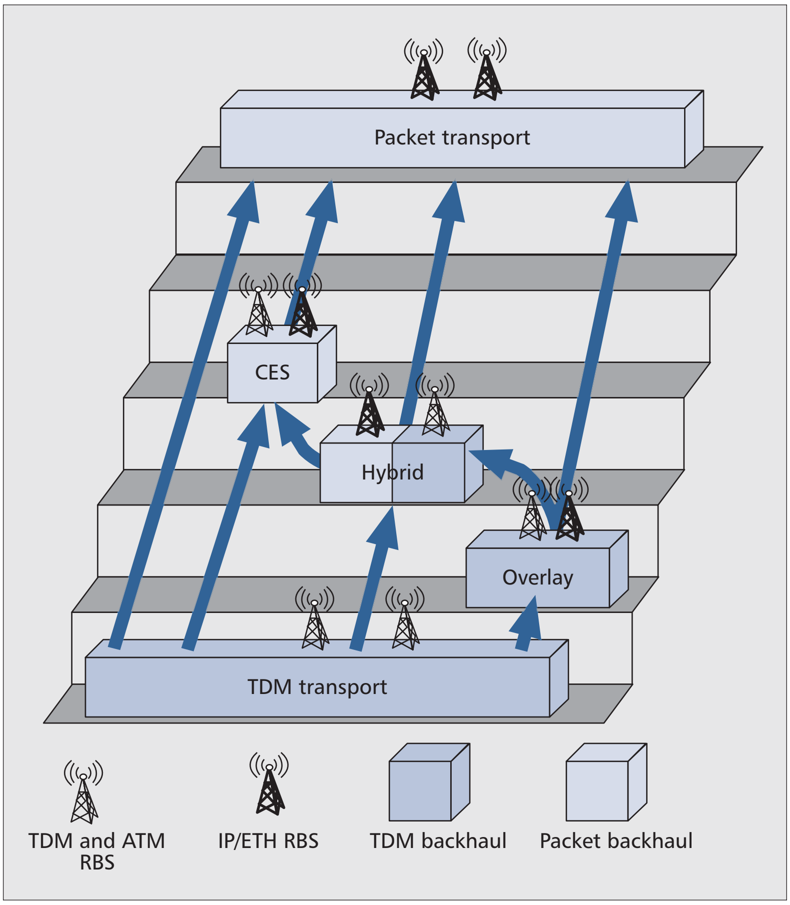


Figure 6. *Transformation options of TDM backhaul network and TDM RBS interface to packet-switched network and IP/ETH RBS interface.*

References

[1] 3GPP TR 36.913 900, "Requirements for Further Advancement for E-UTRA LTE Advanced," 2009.
[2] S. Chia *et al.*, "The Next Challenge for Cellular Networks: Backhaul," *IEEE Microwave*, Aug. 2009.
[3] K. Fouli *et al.*, "The Road to Carrier-Grade Ethernet," *IEEE Commun. Mag.*, Mar. 2009.
[4] IETF RFC 3985, "Pseudowire Emulation Edge to Edge (PWE3)."
[5] IETF RFC 4717, "Encapsulation Methods for Transport of ATM over MPLS."
[6] IETF RFC 5086, "Circuit Emulation Services over Packet-Switched Networks (CESoPSN)."
[7] X. Li *et al.*, "Carrier Ethernet for Transport in UMTS Radio Access Network: Ethernet Backhaul Evolution," *IEEE VTC-Spring*, 2008.
[8] S. Sherif, *et al.*, "On the Merit of Migrating to a Fully Packet-Based Mobile Backhaul RAN Infrastructure," *Proc. IEEE GLOBECOM*, 2009.

Biographies

ZERE GHEBRETENSAE (zere.ghebretensae@ericsson.com) graduated from the Institute of Technology, Linköping, Sweden, in 1989 with an M.Sc. degree in technical physics and electronics. Since then he has worked in Televerket Radio and Telia Research on radio channel modeling and optical fiber transmission. He joined Ericsson Research in 2000, and has worked in optical networking, access and mobile backhaul network architecture. He was a work package leader in the FP6 European research project MUSE and has participated in OIF, ITU, and IEEE standardization work.

JANOS HARMATOS (janos.harmantos@ericsson.com) received his Ph.D. degree in electrical engineering from Budapest University of Technology and Economics, Hungary, in 2005. He is currently working in the Traffic Analysis and Network Performance Laboratory of Ericsson Hungary Ltd as a research engineer. He is working on mobile backhaul architecture solutions and planning of power-efficient mobile networks, as well as in the area of residential networking.

KÅRE GUSTAFSSON (kare.gustafsson@ericsson.com) graduated from the Royal Institute of Technology, Stockholm, Sweden, in 1980 with an M.Sc. degree in applied physics, and has held various positions at Ericsson since that time. He joined Ericsson Research in 2002. He led a subproject for fixed-mobile convergence in the FP6 European research project MUSE and is since 2008 project leader for a research project on mobile broadband backhauling.

Synchronization in Next-Generation Mobile Backhaul Networks

Anthony Magee, ADVA Optical Networking

ABSTRACT

As mobile backhaul networks migrate from legacy time-division-multiplex-based to packet-switched network infrastructures, the transport technology needs to address synchronization requirements once inherently provided by the network. This article explores key technology topics highlighting challenges to be considered in rolling out mobile backhaul synchronization solutions.

LEGACY NETWORKS AND MOBILE AIR INTERFACES

Frequency is the key synchronization component required in legacy mobile base stations, which operate in frequency-division duplex (FDD) mode, using specific frequency assignments for upstream and downstream channels.

FDD base stations need an accurate and stable clock reference to derive a stable operating frequency on the air interface. If the reference clock drifts, interference may be caused, ultimately leading to dropped calls.

Legacy base stations use the inherent capability of the backhaul network, which is time-division multiplex (TDM)-based, to transport timing information from the primary reference clock (PRC) (Fig. 1). The TDM-based E1/T1 traffic interface is used to deliver frequency in the form of 2.048 Mb/s or 1.544 Mb/s, respectively. The rate of the E1/T1 interface is locked to a PRC, which is typically traceable to a Stratum 1 timing source in the plesiochronous digital hierarchy (PDH) or synchronous digital hierarchy (SDH) core of the network, and eventually used to lock the reference clock at the base station. The base station multiplies this clock using phase locked loop (PLL) modules to achieve the desired carrier frequency. Since all base stations have access to a common reference frequency, they are stable with respect to each other. If one base station drifts, so do its neighbors.

Code-division multiple access (CDMA) synchronization has often been solved by the use of a clock recovery function via the Global Positioning System (GPS) being deployed to the cell site. There has been discussion on the scalability of GPS deployment to the cell site in terms of cost and resilience to jamming. Therefore, synchronization via network timing solutions is under investigation for CDMA deployments.

Voice calls were the initial driver for mobile networks. Therefore, bandwidth demand per mobile user was limited. However, with second-generation (2G) technology, features such as short messaging service (SMS) text messaging became available, adding to bandwidth demand. Early implementations of the Global System for Mobile Communication (GSM) provided a bandwidth of 9.6 kb/s/channel.

Consumer Internet awareness and capabilities available from desktop computers drove demand for interactive services via mobile phones. Users adopted picture messaging and became accustomed to accessing services while on the move.

In the late 1990s it was predicted that new modulation schemes and technologies would be required to meet the explosion in data bandwidth expected in mobile communications. Peak bandwidth available to users is currently in the region of 3–14 Mb/s, although this is predicted to increase to 70–100 Mb/s. Video streaming to mobile devices is now accelerating as smart phones increasingly come preloaded with links to video websites; this is a cause of concern for mobile operators as it stretches the mobile backhaul infrastructure.

Observing the challenge in mobile networks, the International Telecommunication Union (ITU) and other standards bodies were motivated to define and harmonize a common set of specifications for new international mobile telecommunication standards known as IMT2000. In Europe the IMT2000 label is known as the Universal Mobile Telecommunications System (UMTS). The equivalent in the United States is referred to as CDMA2000 and Evolution Data Optimized (EV-DO). Also, the European Telecommunications Standards Institute (ETSI) established a project within the Third Generation Partnership Project (3GPP) called Long Term Evolution (LTE), aimed at producing globally acceptable standards for 3G GSM focusing on FDD and time-division duplex (TDD) modes of operation [1–3].

TDD is built on the concept of time-division multiplexing. With TDD, the air interface is split in time rather than frequency, with a single mobile channel being used for transmit and receive directions. This lowers the radio frequency component cost in the mobile handset. In

0163-6804/10/$25.00 © 2010 IEEE

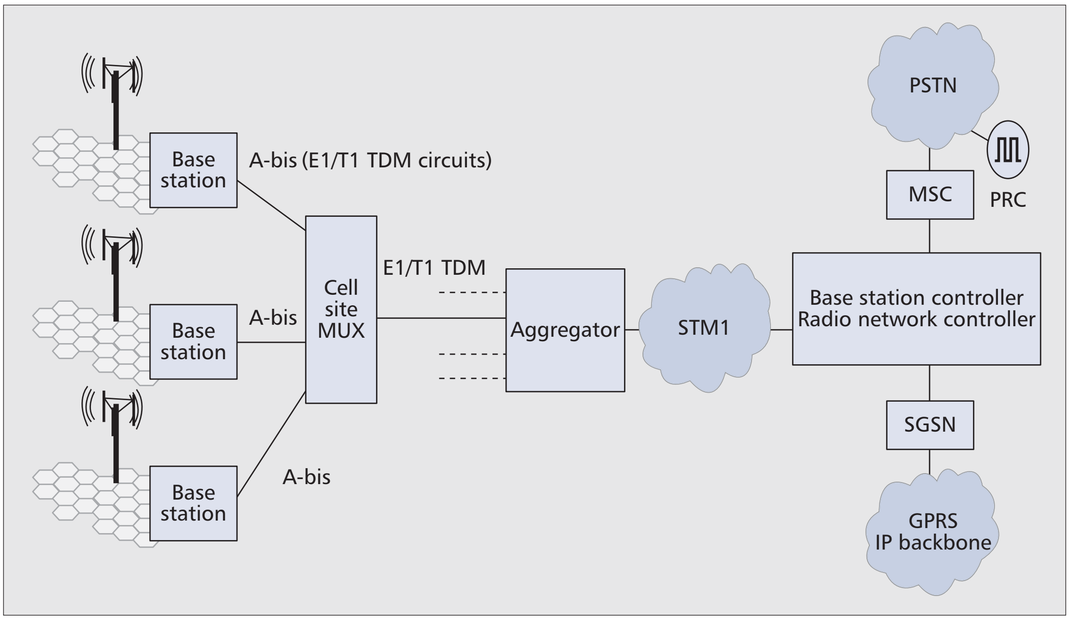


Figure 1. *Legacy mobile backhaul network.*

The frequency requirement at the air interface is an accuracy of 50 ppb from nominal, and time alignment accuracy is in the region of 1 μs (±500 ns). Standards bodies are discussing requirements such as aligning time of day to 200 ns enabling location based services.

between transmit and receive operations transition gaps are required, which, if minimized, can make the channel more efficient. Base stations that are allocated active slots for transmission and reception within a radio network need to be time-aware with respect to each other so that transmission/reception windows are aligned, avoiding interference and dropped calls.

The base stations in a TDD network must be accurately aligned with respect to time of day. They each must know the time relative to the master clock accurately to about 1 μs. UMTS, CDMA2000, and WiMAX are technologies that are specified for use with both FDD and TDD technology.

Air standards that use TDD mode refer to phase[1] and time alignment in addition to frequency. Phase alignment can be implied as a result of time of day alignment, where clocks within the network have the same epoch, often aligned to Coordinated Universal Time (UTC) [3].

Synchronization Requirements

Although applications including performance monitoring, account billing, video distribution, and power systems are candidates for adoption of synchronization, the requirements discussed in this material relate to the mobile network.

Table 1 summarizes the frequency requirement in terms of parts per million (ppm), parts per billion (ppb), and time-of-day requirement in microseconds from the most common air interface specifications, including UMTS (ETSI) [1, 2], CMDA2000 (3GPP2) [3], and WiMAX standardized as IEEE 802.16-2009 [4].

As shown in Table 1, the frequency requirement at the air interface is an accuracy of 50 ppb from nominal, and time alignment accuracy is in the region of 1 μs (±500 ns). Standards bodies are discussing requirements such as aligning time of day to 200 ns, enabling location-based services.

To meet 50 ppb accuracy at the air interface, the budget apportioned for frequency accuracy in the backhaul network is viewed as approximately 16–20 ppb [5].

Synchronous Ethernet and IEEE 1588-2008

To address the challenges posed by migration of mobile backhaul to a packet network architecture, many standards bodies including the ITU-T, IEEE, Metro Ethernet Forum (MEF), and Internet Engineering Task Force (IETF) are defining specifications to enable synchronization information to be transported accurately.

A fundamental standards activity in this area is synchronous Ethernet as defined by the ITU-T in Study Group 15, question 13 (Q13/15). Synchronous Ethernet as indicated in Fig. 2 is a suite of recommendations that allows an Ethernet network to distribute frequency traceable to a PRC node by node in a similar manner to SDH/PDH mechanisms. The clock is separated from the data plane and is constantly transferred across the physical network. Some interfaces such as 1000BASE-T require careful planning and configuration of the master/slave auto-negotiation parameters. Synchronous Ethernet is able to operate over many media with the exception of intermittent physical coding sublayers such as 10BASE-T.

ITU-T G.8261, "Timing and Synchronization Aspects in Packet Networks" [5], discusses the challenges in synchronization over packet-switched networks and provides guidance on synchronization solutions deployment.

ITU-T G.8262, "Characteristics of Synchronous Ethernet Equipment Slave Clock (EEC)" [6], defines performance characteristics of Ethernet equipment clocks (EECs). Within this Recommendation mask plots are provided specifying the accuracy of wander and jitter using mean time interval error (MTIE) and time

[1] *The use of the term* phase *has been discussed in various standards bodies, and there is agreement that phase alignment is inherently provided via time synchronization.*

IEEE 1588v2 PTP is a protocol that measures the path delay between a master and a slave. It uses this information to allow the slave to offset a time-of-day register so that it is aligned to the time of the master. Time stamps are used within fields of the protocol to allow the appropriate calculations to be made.

Technology	Frequency	Phase/time
UMTS LTE/3GPP FDD [1]	50 ppb wide area 100 ppb med/local 250 ppb home	None
UMTS LTE/3GPP TDD [2]	As above	2.5 μs
CDMA2000 [3]	50 ppb	Shall be less than 10 μs Should be less than 3 μs When unlocked, <10 μs for 8 hrs Base stations with multiple CDMA channels pilot time tolerance should be <1 μs
WiMAX [4]	1 ppm worst case Clause 8.1.8	5 μs Clause 8.4.5.1
Location-based services for 3GPP		200 ns Not currently standardized

Table 1. *Air interface synchronization requirements.*

deviation (TDEV) values over various observation intervals.

ITU-T G.8264, "Distribution of Timing through Packet Networks" [7], specifies the use of synchronization status messaging (SSM) within synchronous Ethernet networks. SSM allows the source traceability of the clock to be signaled to downstream devices.

If the clock distribution suffers an interruption, such as physical link failure, the device should indicate to downstream devices the quality level of the clock being used in place of the incoming reference. This way, the device receiving the clock can make a decision perhaps to use a locally attached secondary clock source.

Synchronous Ethernet accuracy is a function of the source reference clock provided by the PRC (typically in the region of a few ppb). When intermediate nodes are locked, the clock is propagated with a long-term frequency deviation of typically less than 1 ppb from that of the source.

When a device becomes unlocked, limits are placed on clock stability for a given period after loss of the reference. To reduce the effects of reference loss, practice is to provide holdover in an equivalent manner to that provided by legacy TDM networks. In case of prolonged holdover, a tightly specified, free-running oscillator is often supported to limit the maximum wander of the clock in the network. Standard practice is to support ±4.6 ppm frequency accuracy in certain parts of the network, while in the central office supporting multiple base stations, accuracy in the region of ±0.016 ppm (or Stratum 2 capable) may be needed to maintain ±15–16 ppb at the base station.

Mobile backhaul networks are required to support the distribution of frequency as well as time to support FDD and TDD mobile air interfaces. IEEE 1588-2008 (also known as 1588v2) [8], which defines the Precision Time Protocol (PTP), is one such method for distribution and alignment of time within a network. Other time-based protocols exist, such as Network Time Protocol (NTP) as defined by the IETF, although 1588v2 is currently favored for this application.

It is widely acknowledged that NTP accuracy is currently limited to a few milliseconds due to implementation at upper layers, whereas 1588v2 has the potential to reach a few hundreds of nanoseconds of time deviation in a real network due to its interaction with lower layers and the use of hardware-based time stamps. However, if NTP was implemented using hardware support, it may be able to reach accuracies similar to that attainable via 1588v2.

IEEE 1588v2 PTP is a protocol that measures the path delay between a master and a slave. It uses this information to allow the slave to offset a time-of-day register so that it is aligned to the time of the master. Time stamps are used within fields of the protocol to allow the appropriate calculations to be made. Time stamp accuracy is an important factor in the level of accuracy that the protocol can achieve. Performing the time stamp somewhere in software at the 1588v2 instance leaves the protocol at the mercy of other protocols transmitting sublayer packets such as operations, administration, and monitoring protocol data units (OAMPDUs), OAM packets, and PAUSE frames, each of which cause 1588v2 packets to be delayed as they traverse the stack. For this reason, it is strongly recommended that 1588v2 is implemented using hardware-based time stamping at the lowest level possible between the physical layer (PHY) device and the media access control (MAC) device [8].

IEEE 1588v2 PTP describes various modes of clock operation at various locations within a network. Figure 3 shows the main components, including the grandmaster clock (GM, the ultimate source of time within a domain), the boundary and transparent clock placed at intermediate nodes, and the ordinary clock (OC), which can be a master or slave clock. IEEE 1588v2 recommends in clause 6.2 that network components such as bridges should be replaced by PTP boundary or transparent clocks whenever possible [8].

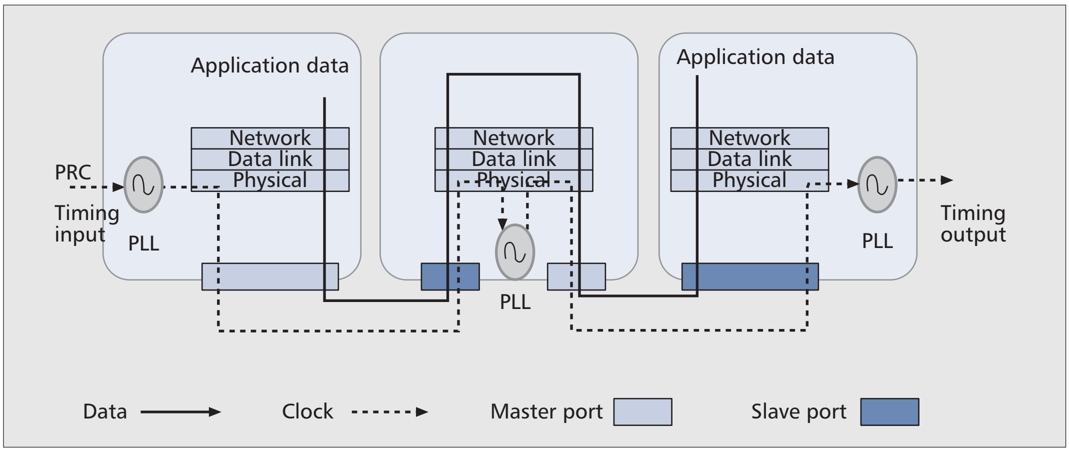


Figure 2. *Clock and data path in synchronous Ethernet.*

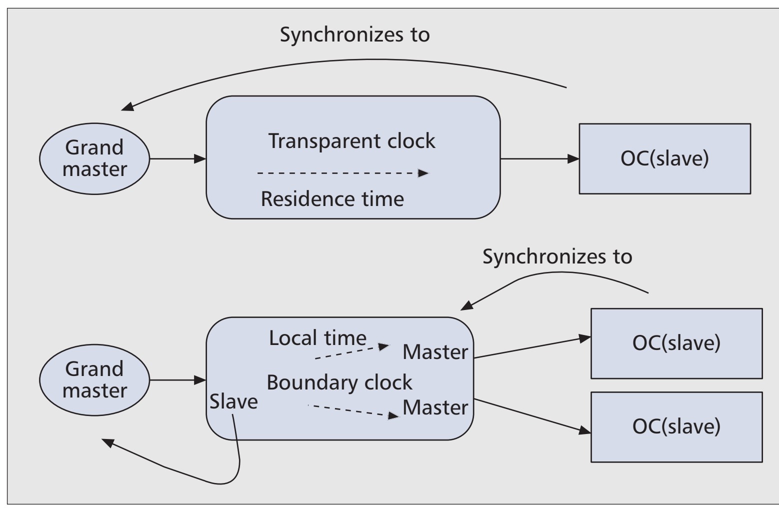


Figure 3. *IEEE 1588v2 overview.*

Boundary clocks are useful to partition the network, separating different measurement delay methods, and also to allow scalability in the master-slave hierarchy. Transparent clocks offer improved or faster convergence in the event of protection switching cases. Some deployment models may favor a particular type of clock. Boundary clocks and transparent clocks are both likely to be required in the long run in order to support flexible deployment models.

1588v2 is considered a suitable technology for use in transporting of clock information that can be used to derive frequency. This allows flexible infrastructure options enabling distribution of frequency over 3rd party networks.

For further discussion regarding use of 1588v2, the measurement technique, and protocol for mobile backhaul needs, refer to [9], which describes work undertaken in ITU-T SG15/Q13 in developing a telecom profile of 1588-2008.

Other standards bodies working on synchronization include the MEF mobile backhaul implementation agreement (MEF 22), and the IETF TICTOC group, which considers 1588v2 over multiprotocol label switching (MPLS) networks.

Additional work is taking place within the IEEE 802.1 Audio Video Bridging Task Group of the 802.1 Working Group. This group is working on IEEE 802.1AS, which defines a profile of 1588-2008 and supports any two endpoints to align their time to a desired accuracy of 1 μs over at least seven nodes [10]. Furthermore, the IEEE 802.3 Working Group has initiated the IEEE P802.3bf TimeSync Task Force to analyze time stamping in support of 802.1AS.

SYNCHRONIZATION CHALLENGES

The challenges associated with synchronous Ethernet and 1588v2 should be considered prior to wide deployment of either technology.

Synchronous Ethernet requires each node in the path to be synchronous Ethernet enabled. If any node does not support the transfer of the clock in a synchronous manner and in compliance with the ITU-T Recommendations, the deployment will not support frequency distribution and accuracy as specified by the ITU.

IEEE 1588v2 also requires important considerations. The accuracy of IEEE 1588v2 is dependent on a number of potential sources of error. One important source of error is the inaccuracy incurred as packets traverse non-IEEE 1588v2 enabled Ethernet nodes with internal packet buffers that vary in size over time. In order to overcome the effects of this packet delay variation (PDV), it is recommended within 1588v2 that all intermediate nodes are 1588v2-aware and can work in either a boundary clock or transparent clock (measuring the residence time of packets and adding this to the correction field of the 1588v2 event messages) in order to meet the more stringent time alignment accuracies. Given the hardware time stamping recommendations, it is unlikely that devices already deployed in the mobile backhaul segment are able to support the accuracy required.

Transparent clocks are recommended within clause 6.5.4 of IEEE 1588-2008 to have their local oscillator syntonized or rate aligned to that of the GM oscillator. One method of achieving this is to use synchronous Ethernet to rate align the oscillators. Other techniques exist, but they are subject to the operation of the protocol and therefore packet delay variation. Note, however, that it is not essential to syntonize a transparent clock in

Although 1588v2 can in principle work in an end-to-end manner and without support from intermediate nodes, studies have shown that using synchronous Ethernet at the physical layer in combination with 1588v2 improves the performance and minimizes time deviation.

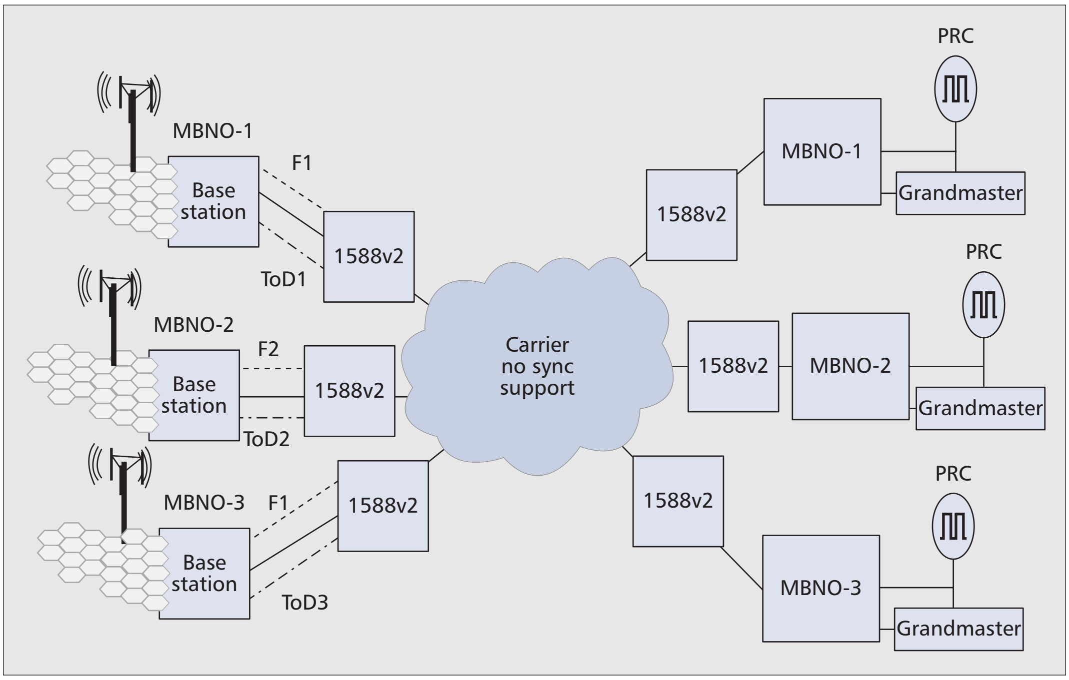


Figure 4. *Mobile synchronization: the ultimate goal.*

order to meet the time alignment specifications we know of today. One proposal is to use a free running oscillator in region of 4.6 ppm [10].

It should also be noted that errors can occur due to the sampling effect of the signal indicating the passing of key 1588v2 packets (event messages) through the PHY device used on Ethernet interfaces, which causes a snapshot of the time-stamp counter [10]. It is likely that the receive clock on the interface between the PHY and MAC will be the clock used to sample the start of packet frame indication on reception and transmission. The typical frequency of this clock is 25 MHz for 100 Mb/s interfaces and 125 MHz for 1 Gb/s interfaces. At 25 MHz, an uncertainty of 40 ns is introduced, narrowing to 8 ns for 125 MHz.

Another source of error in a 1588v2 network is associated with fractional frequency offset (FFO) and residence time (RT). This error (FFO × RT) is accumulated across a network at each node. This can be minimized by using synchronous Ethernet and 1588v2 in conjunction with one another.

The practical size of an accurate 1588v2 network is a matter of debate. Although 1588v2 can in principle work in an end-to-end manner and without support from intermediate nodes, studies have shown that using synchronous Ethernet at the physical layer in combination with 1588v2 improves the performance and minimizes time deviation [11]. In a fast moving environment with ever demanding accuracy requirements it is important to combine technologies in the network together to future-proof the synchronization capability.

In addition to the use of synchronous Ethernet to support greater accuracy with 1588v2, using synchronous Ethernet to distribute frequency at the physical layer also has the advantage of providing a more prolonged holdover availability for time-of-day synchronization in the event that the 1588v2 protocol suffers impairments or loss of the communications channel as a result of faults above the physical layer.

The topics above highlight the reasons why many believe that a complete synchronization solution can be made more robust if synchronous Ethernet and 1588v2 are deployed together and that the technologies are complimentary to each other [12–14].

Another important consideration is to maintain a stable oscillator frequency in the slave. This is important to maintain 50 ppb at the base station. The challenges associated with minimizing oscillator drift as solutions migrate from PLLs to time locked loops are described by Dr. Nigel Hardy in [15].

SYNCHRONIZATION DEPLOYMENT OPTIONS

It is worth considering the options available to a carrier in order to achieve a flexible rollout of synchronization solutions in the backhaul network.

Some believe that the ultimate goal, shown in Fig. 4, is to provide frequency distribution and time-of-day alignment across a carrier's network without any special support for synchronization in the carrier's network. Given that in most regions a number of mobile operators have footprints in a given geographic location, it is desirable that multiple mobile operators can transport their PRC traceable frequency and time of day across the network. 1588v2 deployed at either side of the carrier's network in theory can support this goal, transporting frequency and time of day. This capability will be of benefit in deploying frequency distribution in cases where synchronous Ethernet is not readily available.

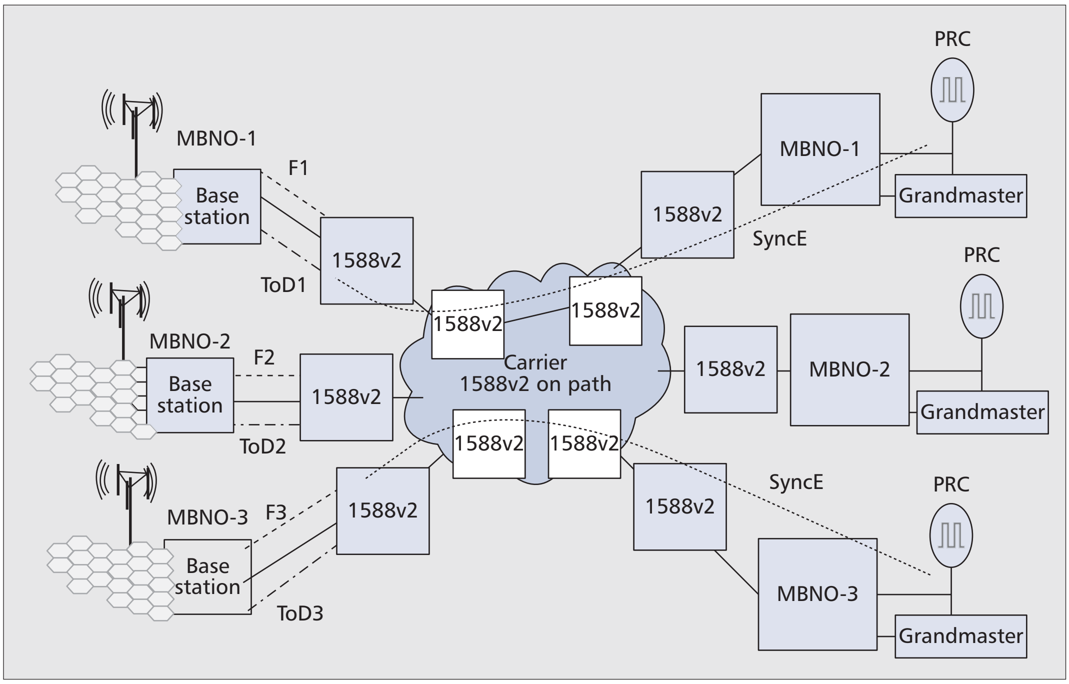


Figure 5. *1588v2 with on path support and synchronous Ethernet.*

Another option to minimize the effects of PDV is to deploy 1588v2 within intermediate nodes within the carrier's network. By deploying boundary clocks and transparent clocks, PDV can be virtually eliminated.

As 1588v2 accuracy is heavily dependent on packet delay variation (PDV) and is affected by transients such as rerouting of packet flows or protection switching events within the network, one might question the practical ability to achieve the desired accuracy in time alignment and frequency stability via 1588v2 end to end without support from the carrier's network. Thus, carriers and mobile operators may be cautious about adopting such topologies until the technology is proven over a wide variety of networks, and time alignment and clock recovery algorithms have been proven over a number of possible events that can occur within a network (note that at the time of writing the ability to support time alignment via 1588v2 for applications such as femtocells via digital subscriber line [DSL] or time of day over wireless infrastructures is still an area of investigation).

A possible workaround is to incorporate traffic management into the network, and place the delay-sensitive traffic in a high-priority class of service, thus minimizing PDV experienced by 1588v2. Even this may not be enough for the highest accuracy, or the outcome may be uncertain until the network has already been built.

Another option to minimize the effects of PDV is to deploy 1588v2 within intermediate nodes within the carrier's network. By deploying boundary clocks and transparent clocks, PDV can be virtually eliminated.

Since this option requires the carrier to provide on-path support, thus touching each node within their network, it makes sense to consider what else may be required to avoid deploying a solution that fails to deliver the highest levels of accuracy.

As discussed in the previous section, it is believed that 1588v2 can be improved via the use of synchronous Ethernet and that this enables a longer holdover capability for time alignment in case of faults occurring in the protocol layer. Thus, there is a strong indication that it makes sense to deploy synchronous-Ethernet-capable equipment at the same time as touching all nodes within the network.

Synchronous Ethernet availability may provide the ability to optimize the clock recovery function in order to derive time alignment. Figure 5 shows the combined deployment of 1588v2 and synchronous Ethernet within the carrier's network.

Practically, the rollout of a synchronization service to meet LTE base station requirements will likely be based on evolution rather than revolution, with a carrier deploying synchronous Ethernet early on, then laying 1588v2 over the top of that, and eventually migrating to 1588v2 alone once algorithm development has progressed sufficiently. The equipment deployed in the backhaul network needs to be flexible enough to cater for an evolutionary upgrade path.

The next factor that should be considered is the requirement to support multiple time domains, allowing multiple mobile operators to be supported.

In the case of multiple mobile operators using the services of a carrier, the carrier may need to support a separate time domain for each mobile operator. This largely depends on the application required by the mobile operators, and the levels of interoperability and agreement between the mobile operators and carriers. In the worst case the carrier needs to support the distribution of multiple synchronization flows, especially for 1588v2.

Questions that need to be considered before deployment are:

- For legacy base stations and FDD environments: Do base stations owned by different operators need different frequency refer-

Long term, 1588v2 is considered an appropriate technology for delivering frequency and time of day alignment, although there are valid reasons for deploying 1588v2 and Synchronous Ethernet combined in a solution. These are viewed by many as complimentary technologies.

ences, or can they share a common PRC source?
- For base stations in TDD environments: Do base stations owned by different operators need different time sources/epochs, or can they share a common source such as UTC?

Other considerations are scaling and architecture of the network. The number of slave clocks that need to be supported and the rate of 1588v2 sync messages can lead to significant bandwidth consumption within the network. Options to mitigate this, especially in tree or cascaded network structures, include the use of multicast rather than unicast, and also the appropriate use of peer-to-peer transparent clocks and/or boundary clocks. A peer-to-peer transparent clock is potentially more efficient than the end-to-end transparent clock since the GM can avoid being subjected to a `Delay_Req` (delay request message) from each slave. The boundary clock supports even greater potential for reducing messaging bandwidth, although this could have an impact on the number of time domains that may be supported across an intermediate node. Since a boundary clock instance contains a slave that aligns to a single master, and uses that single time source to distribute to downstream slaves, multiple time domains require multiple boundary clocks.

Use of boundary and transparent clocks may come under scrutiny in future as potential issues are explored for both modes of operation. Fears that time offset and wander accumulation over a chain of cascaded boundary clocks will be a significant source of error are under investigation within some industry groups, and violations of layer boundaries and the impact of security are causing some to question the use of transparent clocks.

Two further important aspects include asymmetry and synchronization layer congruency. Asymmetry can cause a static error to be introduced in a time alignment solution, so it is important to engineer the network to avoid asymmetry in the PTP traffic path or factor asymmetry into calculations. Regarding synchronization layer congruency, there may be cases where synchronous Ethernet is used at the physical layer and 1588v2 at a protocol layer, but due to network topology and switching events, the timing traceability at the different layers takes different paths. In order to ease fault finding and provide stable management of the solution, it may be beneficial to maintain congruency between the various synchronization layers.

Finally, a significant challenge in deployment of a synchronization service is that of the service level agreement (SLA). In order for a mobile operator to have confidence in the synchronization service and a carrier to prove capability for that SLA, the ability to configure the synchronization domain, monitor the quality in some way, and provide fault finding and reporting tools must be in place.

Summary and Conclusions

Next-generation mobile backhaul networks need to support synchronization in order to serve base stations requiring both frequency and time-of-day (phase) alignment.

Frequency needs to be supported to ±16 ppb, and time alignment needs to be accurate within the region of 1 μs, although the trend will likely be toward tighter accuracies.

Long term, 1588v2 is considered an appropriate technology for delivering frequency and time-of-day alignment, although there are valid reasons for deploying 1588v2 and synchronous Ethernet combined in a solution. These are viewed by many as complementary technologies.

Flexible deployment models supporting multiple mobile operators are likely to drive the need for multiple time domains in the carrier's network. In order to achieve widespread use of these synchronization techniques in such an environment, SLA definitions need to be agreed and developed, along with the tools to measure and verify performance against an SLA.

References

[1] ETSI-TS 125 104, "Universal Mobile Telecommunications System (UMTS); Base Station (BS) Radio Transmission and Reception (FDD) (3GPP TS 25.104 version 8.6.0 Release 8)," Mar. 2009.
[2] ETSI-TS 125 402, "Universal Mobile Telecommunications System (UMTS); Synchronization in UTRAN Stage 2 (3GPP TS 25.402 version 8.0.0 Release 8)," Jan. 2009.
[3] 3GPP2, C.S0010-B Version 2.0, "Recommended Minimum Performance Standards for cdma2000 Spread Spectrum Base Stations," Feb. 2004.
[4] IEEE Std 802.16-2009, "IEEE Standard for Local and Metropolitan Area Networks — Part 16: Air Interface for Broadband Wireless Access Systems," 2009.
[5] ITU-T G.8261, "Timing and Synchronization Aspects in Packet Networks," Apr. 2008.
[6] ITU-T G.8262, "Timing Characteristics of Synchronous Ethernet Equipment Slave Clock (EEC)," Aug. 2007.
[7] ITU-T G.8264, "Distribution of Timing through Packet Networks," Oct. 2008.
[8] IEEE Std 1588-2008 "IEEE Standard for a Precision Cock Synchronization Protocol for Networked Measurement and Control Systems," July 2008.
[9] J.-L. Ferrant *et al.*, "Development of the First IEEE1588 Telecom Profile to Address the Mobile Backhaul Needs," *IEEE Commun. Mag.*, this issue.
[10] G. Garner, "Time Stamp Accuracy needed by IEEE 802.1AS," Contribution to Joint Session between IEEE 802.3 and IEEE 802.1, Jan. 2009.
[11] J. Lewis, "Hybrid Mode Synchronous Ethernet & IEEE-1588 in Wireless TDD Applications," *NIST WSTS*, Mar. 2010.
[12] M. Gilson and S. Taylor, "SyncE & PTP — Deployment Perspective," *NIST WSTS*, Mar. 2010.
[13] S. Jobert, "Challenges with PTPv2 Slaves Performance Testing and Network PDV Characterization," ITSF, 2009.
[14] R. Megino, "T-Mobile view on Next Generation Mobile Network Synchronization," *ITSF*, 2009.
[15] N. Hardy, "Oscillator Performance Criteria for use in Time Lock Loop Applications," *ITSF*, 2009.

Biography

Anthony Magee (amagee@advaoptical.com) has a first class honors degree in engineering and is a principal engineer in ADVA Optical Networking's Advanced Technologies group. He is responsible for synchronization and packet encapsulation techniques for carrier Ethernet. He attends and monitors activities within IEEE 802.3/802.1, ITU, and IETF.

Enter A Subscription

when you renew your IEEE Communications Society membership for 2011

World Leading Publications From IEEE Communications Society

IEEE Communications Society Digital Library

65,000+ Articles • 59,700+ Authors
http://dl.comsoc.org

Electronic access to:

- IEEE Communications Magazine
- IEEE Wireless Communications Magazine
- IEEE Network Magazine
- IEEE Transactions on Communications
- IEEE Communications Letters
- IEEE Journal on Selected Areas in Communications
- IEEE Transactions on Wireless Communications
- IEEE/ACM Transactions on Networking
- IEEE Transactions on Network and Service Management
- IEEE/OSA Journal of Optical Communications and Networking
- The Journal of Communications & Networks
- IEEE Communications Surveys & Tutorial

Communications Society conference proceedings:

ICC, GLOBECOM, WCNC, MILCOM, INFOCOM, NOMS, IM, CCNC, SECON, PIMRC, DYSPAN.

- Regular package (Product#: ONL731):
  Current year & previous three years, about 25,000 papers.
  Member: US $134 Student Member: US $67
- Plus package (Product#: ONLY748):
  Regular package plus back to 1953, over 65,000 papers.
  Member: US $249 Student Member: US $125

IEEE Communications Society Publication Subscriptions

IEEE Network: The Magazine of Global Internetworking
6x/yr, electronic $32, print & electronic, $66

IEEE Wireless Communications
6x/yr, electronic $32, print & electronic, $66

IEEE Communications Letters
12x/yr, electronic $35, print & electronic, $82

IEEE Transactions on Communications
12x/yr, electronic $40, print & electronic, $96

IEEE Journal on Selected Areas in Communications
10x/yr, electronic $40, print & electronic, $96

IEEE Transactions on Network & Service Management
4x/yr, electronic, $32

IEEE Transactions on Multimedia
6x/yr, print & electronic, $50

IEEE/ACM Transactions on Networking
6x/yr, electronic $42, print & electronic $70

IEEE Transactions on Wireless Communications
12x/yr, electronic $42, print & electronic $87

IEEE Transactions on Mobile Computing
12x/yr, print & electronic $51

IEEE/OSA Journal of Lightwave Technology
12x/yr, print & electronic $80

IEEE/OSA Journal of Optical Communications and Networking
12x/yr, electronic $56, 4x/yr, print & electronic $75

IEEE Transactions on Smart Grid
4x/yr, electronic $40, print & electronic $60

Digital Subscription to IEEE Communications Magazine is included with IEEE Communications Society Membership

Above Special Rates are for IEEE Communications Society Members. Subscribe Today!

IEEE

To become a member or order subscription, use IEEE web account at www.ieee.org, or
call +1 800 678 IEEE (USA and Canada), +1 732 981 0060 (worldwide)
Fax: +1 732 562 6380 • email: member-services@ieee.org

IEEE COMMUNICATIONS SOCIETY

Development of the First IEEE 1588 Telecom Profile to Address Mobile Backhaul Needs

Jean-Loup Ferrant, Calnex
Mike Gilson, BT Innovation & Design
Sebastien Jobert, France Telecom
Michael Mayer, Ciena Corporation
Laurent Montini, Cisco Systems
Michel Ouellette, Huawei
Silvana Rodrigues, IDT
Stefano Ruffini, Ericsson

ABSTRACT

This article describes the work performed by ITU-T SG15Q13 for defining the first telecom profile based on the use of IEEE Std 1588-2008. The first profile is specifically developed for the distribution of frequency using unicast IPv4 transmission, and required adaptation of the IEEE1588 protocol to make it suitable for the telecom environment. The objectives, reasons, and results of this adaptation are explained in this article. Since the distribution of phase/time is also gaining importance in telecom, the article briefly discusses the objectives, reasons, and upcoming work for the definition of another profile that will leverage other functions and clocks defined in IEEE1588.

INTRODUCTION

Within the telecom industry there is considerable interest in using packet-based timing mechanisms such as the IEEE 1588 Precision Time Protocol (PTP) to distribute frequency, phase, and time. Packet-based timing solutions such as PTP provide an evolutionary step in the development of next-generation synchronization architectures. One of the first applications for the use of PTP in telecom is in mobile backhaul. Reference [1], published in this issue, provides an overview of mobile backhaul technologies and the radio interface synchronization requirements. It discusses some of the differences with respect to synchronization between frequency-division duplex and time-division duplex mobile systems. It also provides an overview of synchronous Ethernet (SyncE) [2–6] and IEEE 1588 [7] as technologies to be used for the distribution of frequency.

This article describes the work performed by International Telecommunication Union — Telecommunication Standardization Sector (ITU-T) Study Group 15, Question 13 (SG15Q13) on defining the first telecom profile based on the use of IEEE 1588-2008. The first profile is specifically developed for the distribution of frequency using unicast IPv4 transmission, and required adaptation of the IEEE 1588 protocol to make it suitable for the telecom environment. PTP, as specified in [7], does not describe the clock recovery algorithm implementation and limitations or the features to be used in a telecom environment. A PTP profile allows specifying the set of PTP features that are necessary and applicable to the telecom environment. The first telecom profile [8] is being released within ITU-T, with the objective of facilitating end-to-end frequency distribution and synchronization.

This profile is seen as the first in a suite of profiles that will address mobile backhaul needs. As part of this work, ITU-T is also addressing other aspects not in the scope of a PTP profile: packet clock specifications, packet delay variaton (PDV) network limits, and so on, which are fundamental for the telecom environment. This article provides the current status and progress within ITU-T in those areas.

The focus of this article is on a solution based on the exchange of PTP messages between a packet master and packet slave where there is no additional sync support from the network. Such a solution is envisaged to be applicable to a system requiring frequency synchronization, but most likely not applicable to a system with stringent phase/time synchronization requirements.

 0163-6804/10/$25.00 © 2010 IEEE

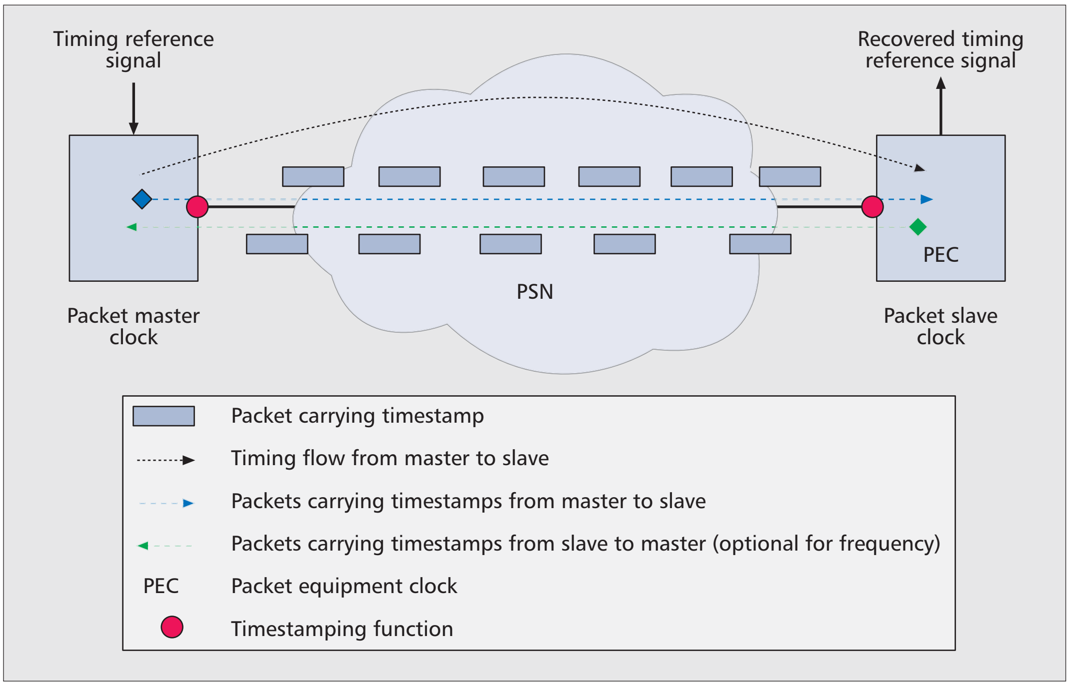


Figure 1. *General principle of packet-based timing methods.*

The term packet-based method identifies a generic class of methods that rely on the exchange of timing information carried within packets. The two main network elements in a packet-based method are the packet master and packet slave.

Packet-Based Solutions to Support Mobile Backhaul Synchronization Needs

In the initial transition phase from time-division multplexing (TDM) to a packet network, and when the distribution of timing over the physical layer is not available (e.g., equipment not supporting synchronous Ethernet or physical limitations across administrative domains), the distribution of timing via packets-based solutions (e.g., PTP) is seen as an alternative option.

Although the initial PTP profile covers only end-to-end frequency synchronization, some of the basic principles presented in this section provide background material covering both frequency and time synchronization. Future profiles will address the need to distribute time synchronization.

Packet-Based Timing Methods

The term *packet-based method* identifies a generic class of methods that rely on the exchange of timing information carried within packets.

The two main network elements in a packet-based method are the packet master (master) and packet slave (slave). As described later in this article, additional functions may be implemented in the network to enhance performance, especially when distributing phase/time. As shown in Fig. 1, the master generates timing packets based on a timing reference signal (frequency reference). The slave recovers the timing reference signal from the packet timing flow. Between the master and slave, the packet-switched network will generate various types of impairments (packet loss, packet delay variation, route change, etc.).

The general principle behind packet-based methods is to compare the time of arrival of a packet (as measured by the local clock in the slave) with the expected time of arrival of the packet generated by the master (i.e., time of departure of the packet from the master plus an offset related to the transfer delay across the network).

One method to transmit the times of departure and arrival is based on the use of timestamps, and protocols such as Internet Engineering Task Force (IETF) Network Time Protocol (NTP)/Simple NTP (SNTP) and IEEE 1588 PTP adopt this approach between a master and slave for synchronizing clocks in the network. The protocols can be used to distribute frequency and/or time. In addition, the protocols can be used in one-way or two-way mode as shown in Fig. 1.

The comparison of local time of arrival with the content of the timestamp as generated by the master corresponds to a measurement of a time offset in a one-way packet transfer that is analogous to the phase error measurements obtained with current physical layer one-way synchronization. As such it is capable of supporting frequency transfer but not precise time transfer.

In contrast to one-way operation, two-way timestamp operation implies packet timing flows in both directions. The use of two-way is necessary to estimate the network transfer delay between master and slave, thus allowing the distribution of time. The use of two-way operation has also been proposed in case of frequency distribution as it would enable the distribution of additional information that can be used in the frequency recovery process.

From an ITU-T perspective, the clock supporting packet-based methods in a packet slave is called a packet-based equipment clock (PEC), and its characteristics are under study.

IEEE Communications | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | qMags

When phase/time synchronization is required, the estimation of the network transfer delay between Master and Slave is also impacted by any asymmetry in the network. In fact the basic assumption in any two-way time transfer method is that the delays are symmetric in the forward and reverse direction.

Characteristics and Performance of Packet-Based Timing Methods

Packet-based timing methods distributing frequency are essentially based on adaptive clock recovery methods [3], and the performance of the recovered timing signal is generally impacted by the PDV in the network.

If the delay through the packet network is constant, there is no variable delay added by the network on the time of arrival of packets at the packet slave (PEC clock). If the delay varies, it may be perceived by a clock recovery process as a change in phase or frequency of the timing reference signal at the packet master. Specific algorithms need to be implemented at the slave side to compensate for PDV. The characteristics of the oscillator implemented at the slave clock also need to be considered.

When phase/time synchronization is required, the estimation of the network transfer delay between master and slave is also impacted by any asymmetry in the network. In fact, the basic assumption in any two-way time transfer method is that the delays are symmetric in the forward and reverse direction.

As described later in this article, additional requirements on the intermediate network nodes (e.g., IEEE 1588 boundary clock and/or transparent clocks) between the packet master and packet slave can also be considered to enhance the performance of these two-way time transfer methods, and their use in telecom networks is being studied.

The timing target performance of some applications is described in [3].

Packet Timing Flow Characteristics

A timing flow represents the timing information (e.g., via timestamps) that may be carried across a packet network.

Packet equipment is designed for optimizing bandwidth use based on statistical multiplexing. The statistical nature of transmission implies that all flows (either data or timing) interfere with each other to some degree, regardless of the priorities (high or low) assigned to the flows.

The forwarding/routing and priority treatment of packets is typically done using various packet processing techniques, which are implemented in either hardware (e.g., network processing unit [NPU], application-specific integrated circuit [ASIC]) or software (e.g., CPU). These techniques have generally been designed to support a variable set of (sometimes complex) network services, but not necessarily the requirements of timing flows. Depending on the vendor implementation, each equipment will generate unique delay variation characteristics, and it can be quite challenging for an operator to characterize the PDV of the network, especially in a multivendor network. This complex scenario is a difficult challenge since timing flows are extremely sensitive to PDV.

It is recommended that network and timing distribution be engineered closely. For instance, assigning the highest priority to a timing flow and/or engineering a specific path to minimize the packet network impairments can be part of such a process. What constitutes a well engineered network to transport timing is currently under study. In particular, some packet metrics are being considered to help characterize packet delay and packet delay variation.

Evolving the Synchronization Architecture

The distribution of frequency using packet-based methods is changing the architecture of the synchronization network to migrate from a link-based approach (e.g., synchronous digital hierarchy [SDH], SyncE) to a path-based distribution mechanism. However, as mentioned above, the use of a well engineered network is essential to prevent network impairments and allow the proper operation of timing methods such as PTP. Such an engineering approach to the transport of synchronization may not be easy, but the advantage provided is that timing can be distributed through the network without having to use node-by-node support (link-based).

Given the drivers to deploy PTP within mobile networks, the relationship between fixed and mobile networks and their respective operators may well influence the development of the network architecture. For example, the entire backhaul network may be owned by the mobile network operator (MNO), the MNO can purchase Ethernet leased lines from another carrier operator that does not provide synchronization, or the MNO may buy some form of managed service that includes the provision of synchronization. All these scenarios and the commercial relationship will thus influence development of the network architecture. This needs to be factored into the limitations of offering PTP technology and the characteristics of the network over which it is transported.

In the case of a fully MNO-owned network, the PTP master may be deployed at the MNO's central office (or anywhere within the backhaul network) since the MNO has control over the end-to-end architecture. This may also be the case when purchasing some form of backhaul capacity from a fixed (carrier) provider. An example is depicted in Fig. 2a. Note that in such a model the MNO slave might be integrated in the base station. It shall be noted that especially when leasing capacity and, due to commercial aspects, the characteristics of this backhaul may raise some issues when the tolerance of the packet clock requires particular PDV conditions. In fact, the production of PDV is very dependent on the architecture, and standardized PDV measurement methodologies for synchronization are still under study.

In the case of a managed service, synchronization might be provided. In this case timing can be provided to the base station via physical layer synchronization delivered from the service delivery platform (e.g., dedicated 2048 kHz signal). In the situation where it is not possible to provide timing to the service delivery platform via physical layer (e.g., via SyncE), a PTP slave might be integrated in this node, and timing could be delivered via PTP from a PTP master that is owned by the carrier, as shown in Fig. 2b.

The end-to-end architecture models illustrated in Fig. 2 will influence where the masters are

Communications | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | qMags

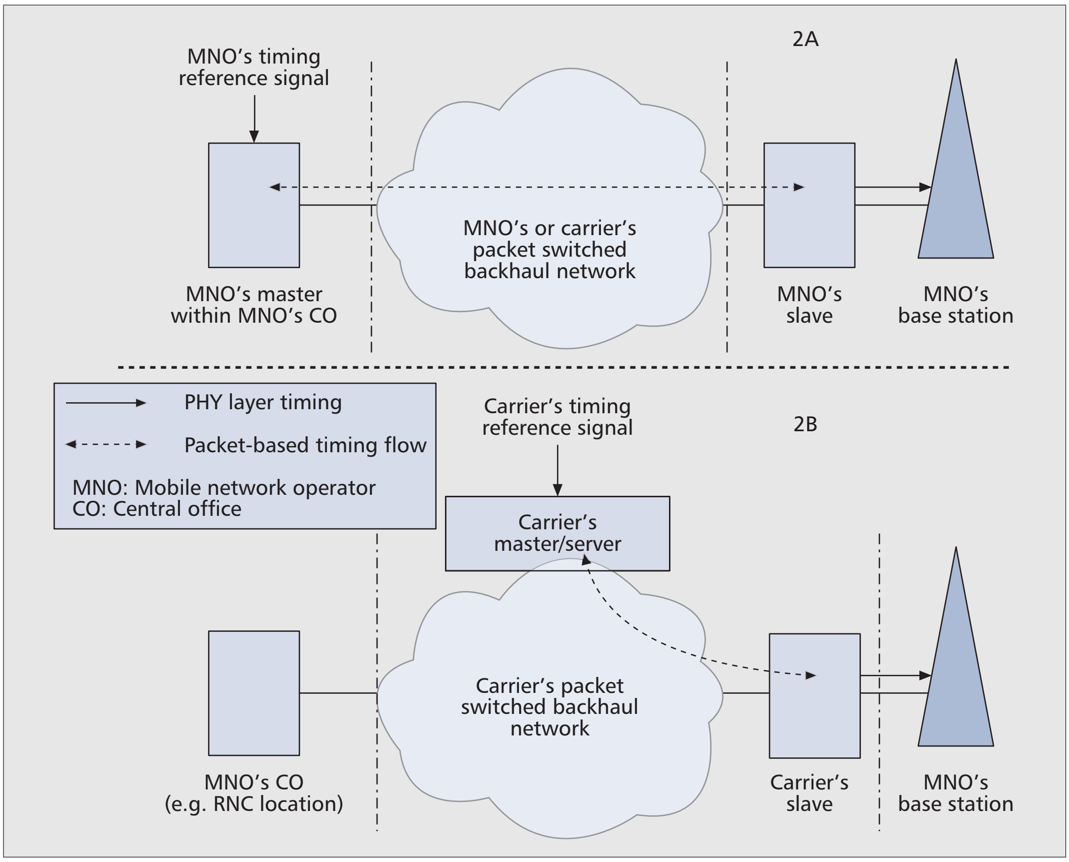


Figure 2. *Deployment of PTP master and slave in a multi-operator context (MNO/carrier).*

To support the different requirements for these applications, the new version introduced the concept of profile where a set of allowed IEEE Std 1588-2008 functions may be selected and specified for a particular use.

to be located, as well as the redundancy aspect of their location. The end-to-end PTP profile is only one part of the synchronization architecture, and this part will be entirely compatible with either an asynchronous or synchronous packet-switched network. Furthermore, the timing signal reference for a PTP master may well be generated from a synchronous network such as SyncE. This can be used to interwork between a synchronous network and an asynchronous network, such as the hybrid Ethernet/packet equipment clock (EEC/PEC) device shown in Fig. 3.

In reference to Fig. 3, it should be noted that, unlike the PECs, the EECs, which are part of the SyncE reference chain, will not be impacted by network traffic. Hence, a high level of performance can more easily be achieved by a SyncE solution, which has an established architecture and performance.

IEEE 1588-2008 as a Solution

This section provides an introduction to PTP defined in IEEE 1588-2008, and the activities that have been required in order to adapt this protocol to the telecom environment.

IEEE 1588 Protocol Overview

The first version of IEEE-1588, "IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems," was published on November 2002. It was designed to synchronize real-time clocks in a distributed network and developed to support the timing requirements of industrial automation and test and measurement applications in a local area network (LAN) environment.

However, interest in expanding the application scope of PTP resulted in the development of a second version of IEEE 1588 containing features that enable the protocol to be used over a wide area network. The second version targeted several applications such as military, power generation and distribution, consumer electronics, and telecommunications. To support the different requirements for these applications, the new version introduced the concept of the *profile* where a set of allowed IEEE 1588-2008 functions may be selected and specified for a particular use.

The ITU-T is currently releasing such a profile for telecom applications. ITU-T Recommendation G.8265.1, "ITU-T PTP Profile for Frequency Distribution without Timing Support from the Network (Unicast Mode)" [8], is intended to be used by telecom applications that need frequency synchronization only, in an environment where the network does not provide any support for PTP, such as boundary clocks or transparent clocks. This first profile is applicable to unicast mode only.

Additional profiles are also expected to be developed later by the ITU to address multicast transmission as well as transport of accurate time and phase. The latter case should involve the use of boundary or transparent clocks, possibly in combination with physical frequency methods such as SyncE. Indeed, while a mode without

It is important to note that while the profile defines the configuration parameters, attribute values and optional features of PTP necessary to ensure protocol interoperability, specification of the profile itself does not guarantee that the performance requirements of a given application will be met.

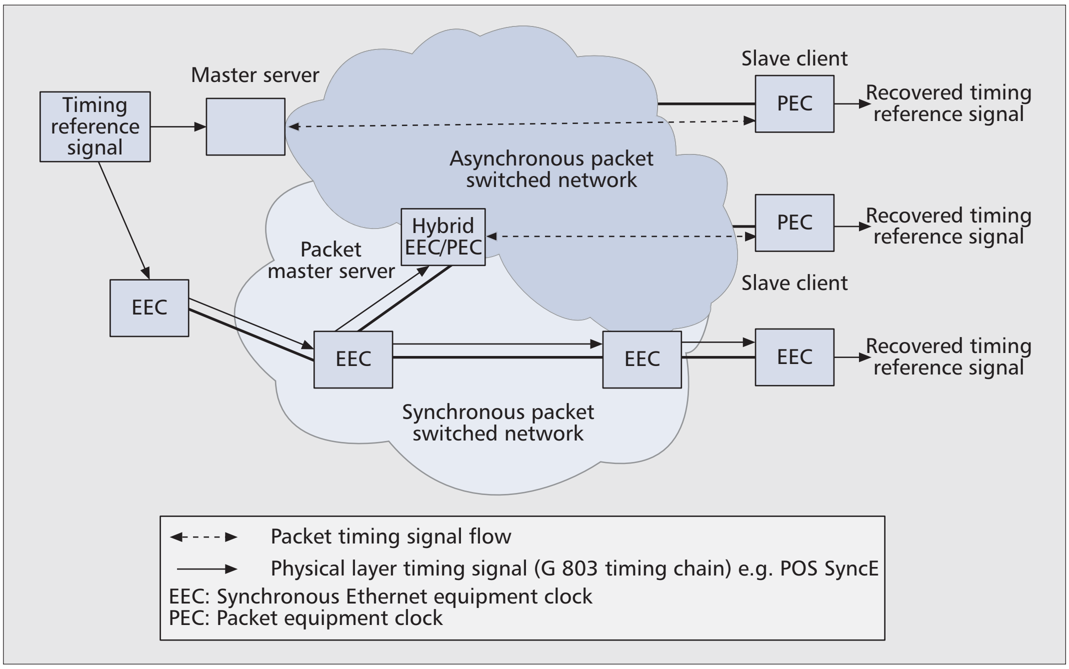


Figure 3. *Frequency distribution based on a packet-based, physical layer and combined mode.*

timing assistance from the network may in some cases enable delivering frequency, it is likely that this mode will not be appropriate in general for accurate phase and time delivery.

The basic synchronization principle in PTP is achieved by the transmission of messages between a slave clock and a master clock as illustrated in Fig. 4.

IEEE 1588 defines one-step and two-step clock operations. In a one-step clock the master sends `Sync` message with the precise timestamp t_1 embedded in the message. In a two-step clock, the `Sync` message contains an approximation of the timestamp t_1, and a follow-up message contains the precise timestamp t_1 of the corresponding `Sync` message. Note that in one-way mode, only these messages are used to exchange timing information.

In two-way mode, in addition to the previous depicted messages, the slave sends a `Delay_request` message to its master and computes timestamp t_3. The master sends a `Delay_response` message with the time of reception t_4 of the `Delay_request` message.

Assuming a symmetric network, the offset and propagation time can be measured as follows:

$$\text{Offset} = [(t_2 - t_1) - (t_4 - t_3)]/2$$
$$\text{Propagation time} = [(t_2 - t_1) + (t_4 - t_3)]/2$$

The offset is the slave time (i.e., the slave clock reading) at an instant of time minus the master time (i.e., the master clock reading) at the same instant of time. It represents the time difference between the two clocks.

IEEE 1588 Features of the Frequency-Only Telecom Profile

The profile contained in ITU G.8265.1 adheres to the general rules for profiles outlined in IEEE 1588. Some choices for particular functionality are outlined below and are based on the specific requirements of the telecom environment. It is important to note, however, that while the profile defines the configuration parameters, attribute values, and optional features of PTP necessary to ensure protocol interoperability, specification of the profile itself does not guarantee that the performance requirements of a given application will be met. Those performance aspects are currently under study at ITU-T and will be part of a different Recommendation.

IEEE 1588-2008 defines several mappings to transport PTP messages, and the default profiles specified in Annex J of IEEE1588 use multicast. However, the standard allows the use of unicast provided the behavior of the protocol is preserved, and defines several optional PTP features.

G.8265.1 contains a general description of unicast and multicast modes. The first version of G.8265.1 defines an Annex for unicast mode where PTP messages are mapped over UDP/IP to facilitate the use of IP addressing. G.8265.1 uses the unicast message negotiation optional PTP feature as this profile is focused on unicast mode. It is planned for G.8265.1 to define another Annex for multicast mode.

PTP is always executed within a PTP domain, representing a logical separation of IEEE 1588 clocks synchronized with each other using PTP. In G.8265.1, for the unicast profile, domains are established by the use of unicast messages, which isolate the different PTP communications. G.8265.1 utilizes the PTP domain number range from 4 to 23.

PTP is a two-way time transfer protocol initially designed to deliver time synchronization, but one-way is also allowed to be used according to IEEE 1588; as for frequency delivery, there is no need to compensate for the propagation delay. Therefore, G.8265.1 allows the use of both one-way and two-way modes. The mode used depends on the slave implementation.

G.8265.1 only uses ordinary clocks, and allows the use of both one-step and two-step clock behaviors.

Alternate Best Master Clock Algorithm and Master Selection Process — IEEE 1588-2008 defines a Best Master Clock Algorithm (BMCA) used to determine the state of each port. The BMCA operates to elect a single PTP clock within a PTP domain as the grandmaster, which is the unique active master in the PTP domain. The standard allows a profile to specify an alternate BMCA that can be used by a profile to determine the state of each PTP clock (e.g., master or slave).

Key telecom-specific requirements include timing protection (slave clocks are visible to multiple master clocks), load balancing (the timing distribution can be shared among multiple active grandmaster clocks in a network), and the need to associate traceability information to the PTP timing flow. Based on examination of these requirements, the ITU-T concluded that clock selection in G.8265.1 would be provided by a master selection process outside of PTP itself, and decided to specify an alternate BMCA.

The alternate BMCA specified in G.8265.1 ensures that masters are always active, and slaves are always slave-only clocks. As IEEE 1588-2008 requires that only one master be active in a PTP domain, each master is isolated in a separate PTP domain by network isolation due to the unicast communication. Moreover, grandmasters do not exchange messages.

G.8265.1 introduces the concept of a telecom slave, consisting of multiple PTP slave-only clock instances. Each instance is configured to communicate with a single grandmaster in a separate PTP domain. This allows the slaves to be connected to several grandmasters, and therefore to participate in different PTP domains.

Figure 5 depicts a logical model of a telecom slave. This model is not intended to imply a specific implementation. In this specific example three different grandmasters are active in the network. As noted above, separate domains are established by unicast communication between the grandmaster and the slave-only clocks. As an example, the telecom slave is locked to the primary grandmaster 1, and grandmasters 2 and 3 will be redundant. If grandmaster 1 fails, the telecom slave will switch to either grandmaster 2 or 3 depending on the master selection process.

The master selection process is based on the quality level (QL)-enabled mode specified in ITU-T Recommendation G.781, "Synchronization Layer Functions" (SDH and SyncE), with some adaptations in order to cover the differences between a physical timing signal and a packet timing signal. The following parameters contribute to the grandmaster selection process:

- Packet timing signal fail (PTSF): Failure condition of the PTP packet-timing signal
- QL: Traceability information of the grandmaster clocks carried by PTP `Announce` messages; interoperable with SDH/SyncE traceability values

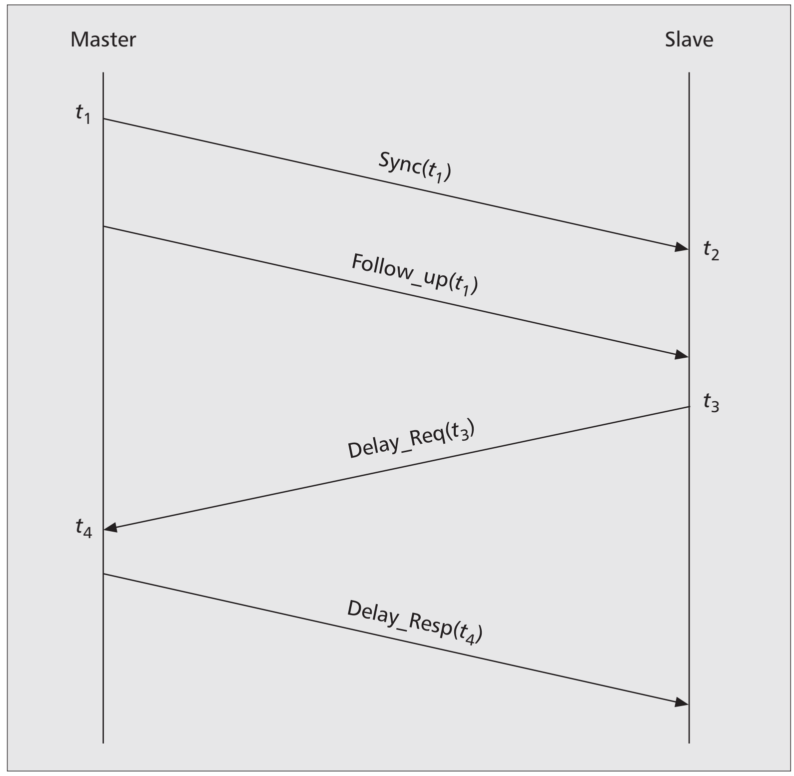


Figure 4. *Basic PTP synchronization.*

- Priority: A value associated with each master, locally maintained in the telecom slave

The reference with the highest QL that is not in a failure condition is selected. In case of similar QL, the reference with the highest priority is selected.

Additional Considerations for IEEE 1588 Deployment

The development of the telecom profile begins to specify the parameters and features specified within IEEE1588 that will allow PTP to be used in carrier networks for distribution of frequency. While much progress has been made in this area, as noted above, it is important to note that other aspects considered to be outside of the scope of the profile itself are, in some cases, critical to the eventual deployment of packet-based timing solutions. At a high level, deployment of packet-based timing requires careful consideration of network design and management in order to meet the end goal of providing traceable timing. The behavior of packet networks under various load conditions will require care in engineering the network to ensure that a packet network does not introduce excessive PDV.

Once deployed, the issue of monitoring and maintaining performance exists. Work is ongoing at ITU-T to develop a standardized set of metrics that will allow measurement of basic packet delay metrics in order to estimate the effective performance that may be produced from a slave clock. At present, only proprietary metrics exist, but this is an area of increased interest in standards.

The structure and characteristics of the slave clocks in the network are another critical area

Work is currently underway to attempt to define clock performance parameters for use in packet timing distribution. Additionally, future work may look at the applicability to telecom environments of other PTP clocks that have been defined for LAN environments.

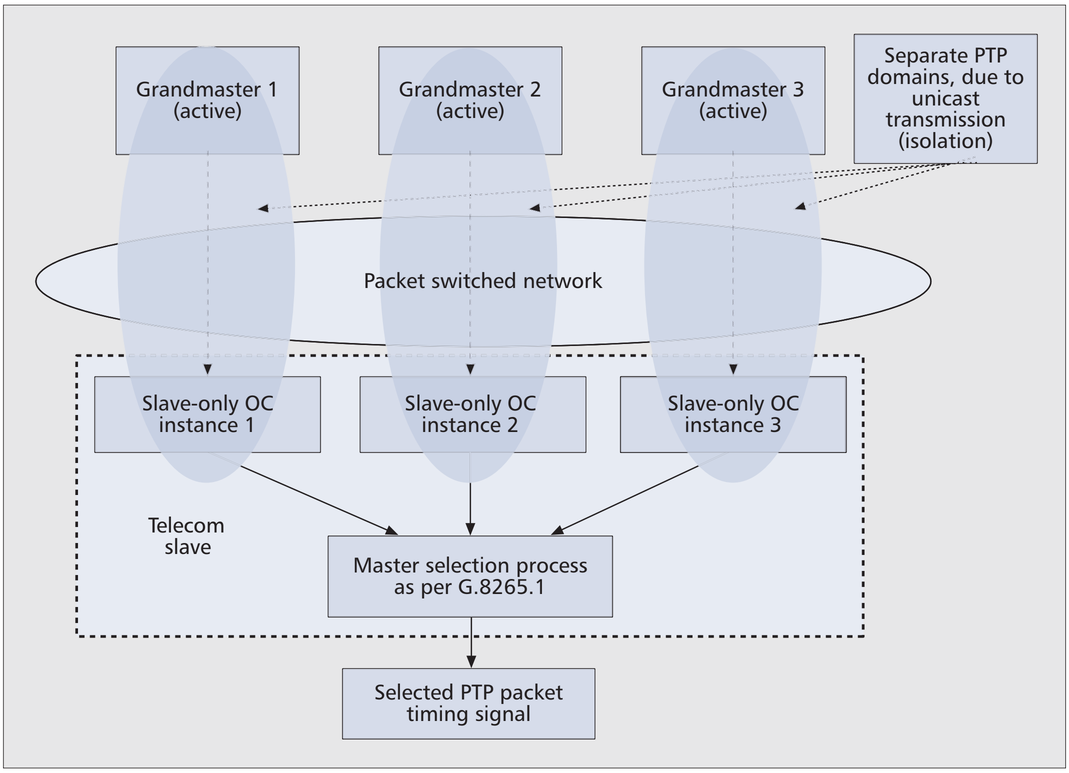


Figure 5. *Multiple slave-only OC instances in a telecom slave.*

that is outside the scope of G.8265.1, but will have a strong impact on the performance achievable and hence the performance that may be available to support a revenue generating service. For the first version of the profile, the clocks supported are referred to as *ordinary* clocks. These simply terminate the packet flows containing PTP messages and then generate the appropriate reference. Performance aspects of the clocks are outside of the scope of the profile. Work is currently underway to attempt to define clock performance parameters (holdover, etc.) for use in packet timing distribution. Additionally, future work may look at the applicability to telecom environments of other PTP clocks (e.g., boundary and transparent clocks) that have been defined for LAN environments. The impacts on existing network architectures and topologies need to be carefully considered.

Future Profiles

While frequency distribution addresses many current applications, certain technologies require the use of time and phase in addition to frequency. Code-division multiple access (CDMA) and Long Term Evolution time-division duplex (LTE-TDD) are examples of such technologies. The ITU-T is currently planning to develop additional profiles that will utilize IEEE 1588 for the delivery of accurate phase/time and frequency. The accuracy levels for time and phase are application-dependent, and in some cases the frequency stability requirements to attain highly accurate and stable time distribution may involve the use of expensive oscillators. Proposals have been made to the ITU-T to utilize PTP for phase/time distribution in conjunction with synchronous Ethernet to enable the physical layer to provide access to highly stable network clocks, as shown in Fig. 6.

As noted earlier, for phase/time distribution the performance achievable using packet-based methods is affected by network asymmetry. To achieve accurate phase/time synchronization, timing support from the network may be needed as it can significantly reduce the impact of packet network impairments.

In this case the intermediate nodes provide support for PTP, via the implementation of IEEE 1588 boundary or transparent clocks. Each intermediate node is therefore *IEEE 1588 aware*. Hence, the timing flow containing phase or time information generated by a master is processed in each node. Support of boundary and transparent clocks will require development of additional telecom profiles, and clock or node characteristics and metrics.

In the original LAN environment targeted by IEEE1588, implementing and deploying boundary or transparent clocks is possible, as there is full control over the equipment and topology that may be deployed. However, deploying boundary or transparent clocks in telecom networks may be operationally challenging and need appropriate engineering considerations from the timing and networking standpoints.

Conclusion

The development of the PTP profile for telecom is a very important step in the ongoing development of the next-generation network synchronization architecture. Considerable development

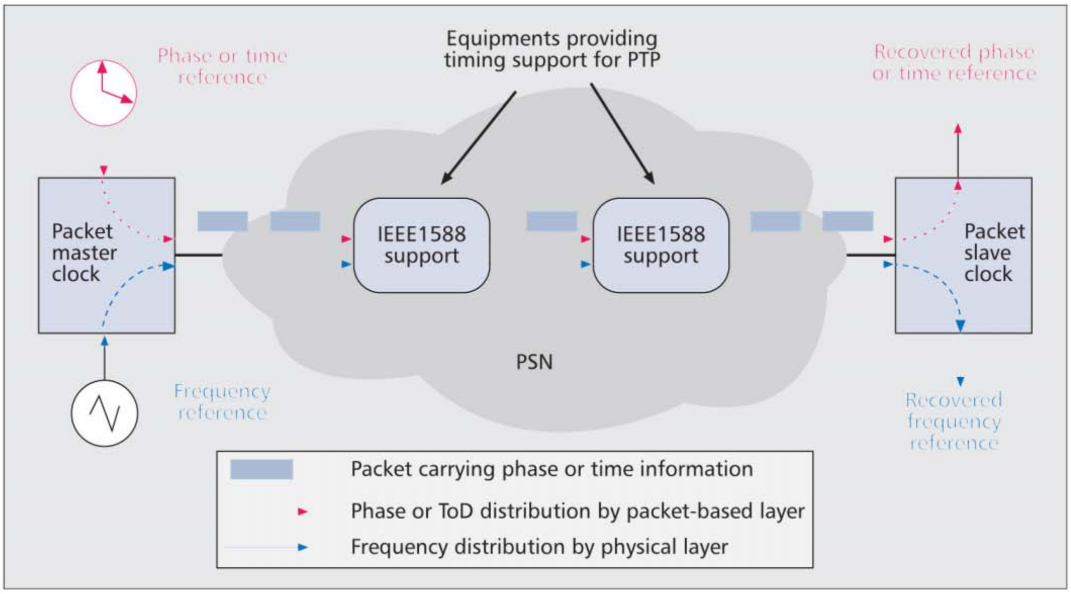


Figure 6. *Illustration of a packet-based method with timing support from the network combined with frequency distribution delivered at the physical layer.*

It is critical to understand that to achieve expected application performance, the packet-based timing method requires the network to be well engineered accordingly, as PDV may impact the achievable synchronization performance.

work has now taken place within the ITU, and the point has been reached where an end-to-end frequency-based profile is nearing completion. This article has highlighted mobile backhaul with its requirements as one of the initial application areas for the profile.

Within the initial profile a number of key decisions have been made in terms of mode of operation (i.e., unicast), aspects of its reach (i.e., specific domains of operation), and master clock selection (i.e., the Alternate Best Master Clock algorithm).

It is critical to understand that to achieve expected application performance, the packet-based timing method requires the network to be well engineered as PDV may impact the achievable synchronization performance.

A number of key areas still require work. These include network design and management, monitoring and metrics, and the implementation of slave clocks, among others. In addition, work to extend its use to transport phase/time information will also be important. Such requirements may well make use of synchronous Ethernet and IEEE 1588 PTP hop-by-hop implementations.

References

[1] A. Magee, "Synchronization in Next-Generation Mobile Backhaul Networks," *IEEE Commun. Mag.*, this issue.

[2] J. L Ferrant *et al.*, "Synchronous Ethernet: A Method to Transport Synchronization," *IEEE Commun. Mag.*, Sept. 2008.

[3] ITU-T G.8261, "Timing and Synchronization Aspects in Packet Networks," Apr. 2008.

[4] ITU-T G.8262, "Timing Characteristics of Synchronous Ethernet Equipment Slave Clock (EEC)," Aug. 2007.

[5] ITU-T G.8264, "Distribution of Timing through Packet Networks," Oct. 2008.

[6] ITU-T G.781, "Synchronization Layer Functions," Sept. 2008.

[7] IEEE 1588-2008, "IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems," July 2008.

[8] ITU-T G.8265.1, "ITU-T PTP Profile for Frequency Distribution without Timing Support from the Network (Unicast Mode)," June 2010.

Biographies

JEAN-LOUP FERRANT (jean-loup.ferrant@calnexsol.com) graduated from INPG Grenoble, France, joined Alcatel in 1975, and worked on analog systems, PCM, and digital cross-connects. He has been working on SDH synchronization since 1990, and on SDH and OTN standardization for more than 15 years in European Telecommunications Standards Institute (ETSI) TM1 and TM3, and ITU-T SG13 and SG15. He has been rapporteur of SG15 Q13 on network synchronization since 2001. He was an Alcatel-Lucent expert on synchronization in transport networks until he retired in March 2009. He is still rapporteur of SG15 Q13, sponsored by Calnex.

MIKE GILSON (mike.gilson@bt.com) is a technical specialist with BT Innovation & Design at Adastral Park. He joined BT in 1983, and has played a major role in developing and implementing BT's time and synchronization strategy from 1988 to the present day. He has worked on synchronization standardization in ITU, ETSI, and IETF, but is currently only actively contributing to the ITU and ITSF NIST synchronization fora. He is a member of both the ITSF and NIST Steering groups. He has a B.A. (Hons) degree in business studies and is a member of the Institute of Engineering and Technology (IET).

SÉBASTIEN JOBERT (sebastien.jobert@orange-ftgroup.com) is an R&D engineer, Network Synchronization, France Telecom. He received an engineering Master's degree from Paris University Pierre et Marie Curie in 2005. After two years working on VoIP and data synchronization aspects, he joined France Telecom R&D/Orange Labs in 2006, where his current research domain is focused on time and frequency synchronization in telecom networks, and provides expertise to the France Telecom Group in this domain. He takes an active part in standardization through ITU-T SG15 Q13 dealing with network synchronization and IETF. He is also a member of the ITSF steering group.

MICHAEL MAYER (mmayer@ciena.com) graduated from Queens University (B.Sc.EE). In 1982 he joined Bell-Northern Research and later Nortel. He is currently with the CTO Department of Ciena Corporation. Since 1995 his area of focus has been systems design and standardization, where he has been actively contributing to standards in the areas of synchronization, OTN, SONET/SDH, and ASON control plane architecture. Within the ITU, he has been a technical editor of several ITU-T Recommendations covering the ASON optical control plane and synchronization related topics.

LAURENT MONTINI (lmontini@cisco.com) is technical leader in the Office of the CTO at Cisco Systems. Since 2005 he focused his work on frequency and time distribution over

packet network, especially in the telecom environment. He participates in standardization efforts including IEEE 1588, ITU-T SG15 Q13, and IETF TICTOC. He joined Cisco in 1997 and for six years has been a consulting engineer for mobile operators promoting network convergence for backbone and radio access networks where the emergence of pseudo-wire techniques raised the timing issue.

MICHEL OUELLETTE (michel.ouellette@huawei.com) is a technical advisor in Huawei's IP Network Solutions and Clock Laboratory, where he focuses on mobile backhaul networks, and the development and analysis of packet network architectures/protocols for accurate frequency and phase/time distribution. He participates in ITU-T and IEEE 802.3 standardization. Prior to joining Huawei, he spent 12 years at Nortel focusing on ATM/TDM pseudowires, synchronous Ethernet, clock algorithms for base stations, TCP/IP active queue management, and ATM switching. He has been granted 10 patents and has published several international journal papers. He received his B.A.Sc. and M.A.Sc. from the University of Ottawa in 1995 and 1997, and attended l'Ecole Nationale Superieure des Telecommunication.

SILVANA RODRIGUES (silvana.rodrigues@idt.com) is director of system engineering at Integrated Device Technology (IDT). She graduated in electrical engineering from Campinas University, Brazil. She started her career at the Telecommunication Research Center in Brazil; she worked at Zarlink where she held several positions as analog designer, engineering manager, and synchronization system architect. She participates in several standards groups; she is the editor of ITU-T Recommendations G.8263 and G.8262, and secretary for IEEE 1588.

STEFANO RUFFINI (stefano.ruffini@ericsson.com) is Expert R&D at Research & Innovation, Ericsson. He joined Ericsson in 1993 and has been working on synchronization aspects for about 15 years. He has represented Ericsson in various standardization organizations (including ETSI, ITU, 3GPP, and IETF) and is currently actively contributing to ITU-T SG15 Q13 (serving as associate rapporteur and editor) and IETF TICTOC. He is one of the Ericsson experts involved in the definition of equipment and network synchronization solutions.

Shortest Path Bridging: Efficient Control of Larger Ethernet Networks

David Allan, Ericsson
Peter Ashwood-Smith, Huawei
Nigel Bragg, Ciena
János Farkas, Ericsson
Don Fedyk, Alcatel-Lucent
Michel Ouellete, Huawei
Mick Seaman, Consultant
Paul Unbehagen, Alcatel-Lucent

ABSTRACT

This article provides an overview of IEEE 802.1aq shortest path bridging and outlines some application scenarios that will benefit from the new capabilities SPB offers. SPB is built on the IEEE 802.1 standards, and inherits unaltered the existing OAM and data plane scalability enhancements, such as the MAC-in-MAC forwarding paradigm. SPB introduces link state control for bridge networks, thus improving control plane scalability, network bandwidth utilization, and control of the forwarding paths. Furthermore, SPB minimizes latency by forwarding frames on the shortest path. Network-wide load balancing is also supported by spreading the traffic on multiple equal cost paths in a user controllable manner. Thus, SPB provides enhanced control for Ethernet networks in metro, RAN backhaul, or data center environments.

INTRODUCTION

The traditional Ethernet bridging and data link layer has served well for several decades now and forms a nearly universal layer 2 infrastructure for many protocols including IP/multiprotocol label switching (MPLS). Furthermore, Ethernet bridging used as a service layer provides end-to-end Ethernet connectivity, and is complemented by robust and fully featured operations, administration, and maintenance (OAM) capabilities.

Much of the success of Ethernet is due to its plug-and-play nature, which is highly valued by consumers and operators alike. This desire to preserve simplicity also led to the adoption of relatively simple control planes in the form of the Spanning Tree Protocol (STP) and Rapid Spanning Tree Protocol (RSTP). These protocols automatically maintain the active topology required for learning bridges, while requiring minimal configuration and storage. Multiple Spanning Tree Protocol (MSTP) was added to allow multiple spanning trees to be constructed over small meshes of Ethernet switches.

The primary shortcoming of spanning trees is that they do not utilize all links in certain topologies. The resulting connectivity is inefficient for traffic that does not terminate at the root of the spanning tree. Furthermore, spanning tree protocols implement the transactional distance-vector class of algorithm instead of a proper topology database, which adversely impacts the convergence time of an Ethernet network after a topology change. To date, this has constrained the scale and utility of pure Ethernet networks.

The addition of 802.1ah provider backbone bridging (PBB) [1] to Ethernet highlighted the need for a new control protocol to address these issues. Any such control protocol would be required to maintain all of the key architectural properties of Ethernet specified by IEEE 802.1 and make minimal changes of detail only where essential so that the huge body of existing work and implementations could be leveraged. A further key requirement was to use minimum cost paths in an arbitrary mesh and in ways not limited to a single shortest path between any two points.

Shortest path trees (SPTs) can be built to both support Ethernet 802.1Q and leverage the new medium access control (MAC)-in-MAC encapsulation (first specified for PBB) for larger networks. MAC-in-MAC is ideally suited for isolation of layer 2 virtual private networks (VPNs) and, when combined with the 802.1aq shortest path bridging (SPB) [2, 3] control plane, truly extends the scalability of Ethernet by several orders of magnitude. It was also a requirement to support the VLAN and provider bridge (PB)

0163-6804/10/$25.00 © 2010 IEEE

data path, albeit at smaller scales, so that existing inexpensive packet forwarding application-specific integrated circuits (ASICs) could also provide shortest path functionality.

Link state techniques remain the approach of choice for distributed routing. SPB uses the International Organization for Standardization (ISO) Intermediate System to Intermediate System (IS-IS) routing protocol [4]. IS-IS is ideally suited to support the additional functionality, and only requires modest extensions for layer 2 [3]. Furthermore, IS-IS has a long track record of large-scale deployment with robust behavior. With its adoption, it was believed that 1000-node or larger pure Ethernet networks were not out of the question. Early small pre-standard live deployments and much larger emulations bear this out. The extended version is referred to as ISIS-SPB.

SPB is specified as an amendment to IEEE 802.1Q [5] standard and meets the backward compatibility requirements inherent to such amendments; hence, SPB is able to interoperate with MSTP, RSTP and STP. Furthermore, SPB is equipped with the OAM specified by 802.1ag connectivity fault management (CFM) [6].

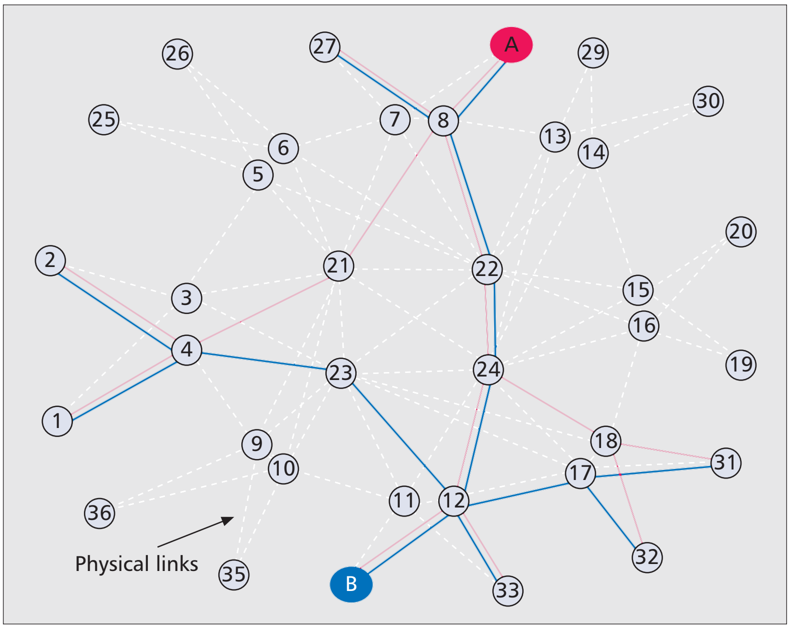


Figure 1. *The symmetry of two shortest path trees, A and B: the path between A and B is the same in both trees, and both reach same subset of members.*

Shortest Path Bridging Principles

Shortest path bridging [2] provides frame forwarding on the shortest path within an SPT region of a network by using ISIS-SPB on all SPT bridges to control the forwarding paths. An SPT region is seen as a single bridge from outside the region. ISIS-SPB uses the standard IS-IS procedures to construct and update the link state database in each SPT bridge. SPB extensions to IS-IS are minimal, being only the procedures and fields (TLVs) strictly necessary for the link state control of a bridged network.

ISIS-SPB sets up and maintains at least one SPT for each bridge, which connects to every other bridge in an SPT region. Each bridge roots at least one SPT, and a bridge only sends frames on one of its own SPTs. Thus, SPB implements source rooted trees, which is ideal for multicast forwarding and applications using it, such as IPTV.

SPTs have to meet two congruency criteria. Forward and reverse paths must be the same between any two bridge pairs, and unicast and multicast paths also have to be congruent. This is key for the support of MAC address learning, OAM fate sharing, and preserving frame ordering guarantees during the learning process. Therefore, SPT bridges have implemented a tie-breaking extension to the algorithms used for shortest path calculation, which makes the results independent of the order of computation, and so deterministic. Figure 1 shows this result visually for two trees (A and B). Varying the tie-breaking rules enables the generation of multiple equal cost trees (ECTs) in order to implement load balancing (described and shown later in detail).

It is essential to avoid loops in an Ethernet network so as to prevent the multiplication of looped broadcast, flooded unknown destination, and multicast frames. 802.1aq incorporates both loop mitigation and loop prevention techniques. Mitigation exists to minimize the damage done if a loop should form, while prevention focuses on ensuring that loops do not happen in the first place. Ingress filtering mitigates loops by auditing the port of arrival of a frame, to ensure that it arrives on the port lying on the path from the source SPT. As not all possible loops can be eliminated by the ingress filtering, ISIS-SPB also implements a loop prevention mechanism for multicast forwarding. Neighbor bridges exchange digests of the topology database to check whether they have the same view of the physical topology, and by inference also have agreement on the distance to all SPT roots. An SPT bridge only installs changes to multicast forwarding to a peer when their digests match. Traffic unaffected by a topology change sees no interruption in forwarding.

SPB takes additional advantage of the topology auto-discovery built into IS-IS, extending this to include discovery of service membership. To achieve this, virtual LAN identifiers (VIDs) or PBB service identifiers (I-SID) configured on SPT bridges are carried in ISIS-SPB TLVs. In this way, ISIS-SPB can configure frame forwarding specific to the registered services without extra discovery mechanisms or further signaling.

SPB has two operating modes, as illustrated in Fig. 2. The two modes are distinguished by the method used for tree identification. Shortest path bridging VID (SPBV) uses a VID to identify an SPT. A backbone MAC address is used as the identifier of an SPT in shortest path bridging MAC (SPBM), which uses the 802.1ah encapsulation. However, both modes operate using the same SPB principles.

A VID per bridge supports a VLAN in an SPBV region. A distinguished VID, the base VID, is used to identify a VLAN in management operations in both SPBV and SPBM modes; in SPBM this VID is also used in the forwarding

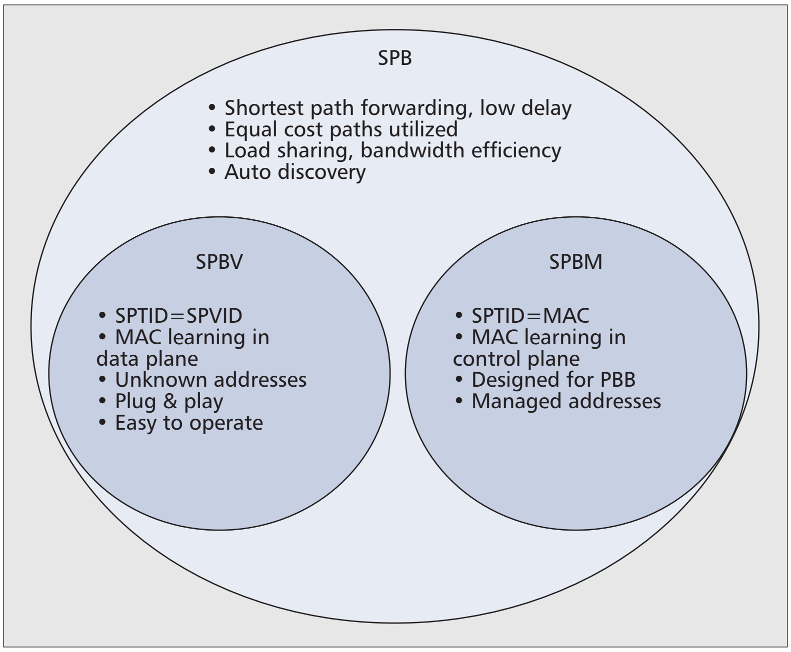


Figure 2. *SPB operating modes.*

plane. Furthermore, the frames belonging to a particular VLAN are tagged with the VLAN's base VID outside of the SPT region.

SPBV

SPBV uses a shortest path VID (SPVID) to identify each SPT within an SPT region. The assignment of an SPVID to a base VID, and hence VLAN, for each SPT bridge, and maintenance of the associated VID translation tables are performed automatically by ISIS-SPB.

A frame arriving at the boundary of an SPT region has its VID swapped to the SPVID corresponding to the SPT rooted at the boundary bridge for transmission within the region. A frame leaving the SPT region has its SPVID swapped back to the base VID for that VLAN.

It is also a key characteristic of SPBV operation that it keeps the MAC address learning in the data plane as specified in [5]. In other words, SPT bridges maintain the practice of learning the reachability of a MAC address from the source address field of the frames. The set of SPVIDs belonging to the same VLAN operate in shared learning mode to do this.

In this way, SPBV delivers the benefits of SPB while retaining the traditional Ethernet support for plug-and-play operation, and it is thus easy to manage.

SPBM

Shortest path bridging MAC operation uses MAC addresses to identify the SPT, thus increasing scalability by several orders of magnitude. Also, the 24 bits of the 802.1ah header service instance (I-SID) provide highly scalable support for service virtualization (i.e., L2VPNs). The ISIS-SPB for MAC mode distributes the B-MAC addresses of the bridge rooting the SPTs for all client VLANs supported within the SPB region. The predecessor of SPBM is provider link state bridging (PLSB), described in detail in [7].

SPBM is designed to be applied in a network using MAC-in-MAC encapsulation, as defined in [1]. With SPBM, all B-MAC addresses are known and distributed by ISIS-SPB, so B-MAC learning can be turned off in the SPB region, resulting in frames with unknown B-MAC addresses being discarded.

Instead of distributing the multicast addresses in SPBM, they are computed locally by creating source-specific multicast trees. A 20-bit identifier called shortest path source ID (SPSourceID) identifies the source SPT bridge for multicast forwarding. A group MAC address is formed by concatenating the SPSourceID with the 24-bit I-SID value.

SPBM permits two user selectable multicast models, one of which builds source-specific multicast tree state into the network, and another that uses head-end replication and unicast transmission of frames.

SPBM matches the inherent scalability of MAC-in-MAC, making it appropriate in larger fully managed networks.

Support for MEF Services

SPB supports the E-LINE, E-TREE, and E-LAN services defined by the Metro Ethernet Forum (MEF) [8].

These services are supported by SPBV mode in the same way they are supported by an MSTP controlled network.

SPBM takes advantage of the I-SID service identification feature of PBB. SPBM assigns multicast attributes to each SPT bridge participating in a service instance, which specifies whether that bridge is a transmitter, a receiver, or both for that particular I-SID. ISIS-SPB then propagates this information to all bridges within the region, which can then set up the required local multicast frame forwarding accordingly.

In SPBM an E-LINE is a point-to-point connection and is implemented by unicast forwarding. The multicast attributes are both *off* for an E-LINE I-SID, because no group MAC (i.e., multicast) address needs to be set in the FDBs.

An E-LAN is created when both of the multicast attributes are *on*, as then each backbone edge bridge (BEB) that is a member of this particular service is both a transmitter and a receiver as well.

Two I-SIDs are used to implement an E-TREE service, which may have multiple roots. One I-SID is used for forwarding from the leaves toward the roots; thus, leaf BEBs are configured as transmitters and root BEBs are configured as receivers. On the other I-SID, the leaf BEBs are receivers only, but root BEBs are both transmitters and receivers to allow forwarding from a root to all other members of the E-TREE, other roots as well as the leaves.

Use of Equal Cost Paths

Shortest path forwarding allows for inherent simplification and improvement in utilization of a mesh network, because all paths may be used, and none need to be blocked for loop prevention. It is possible to get even greater utilization by allowing the simultaneous use of multiple equal cost shortest paths. 802.1aq initially allows

IEEE Communications | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | qMags

for 16 relatively diverse tunable shortest paths between any pair of nodes, while providing a framework for future innovations. This is achieved by manipulation of tie breaking between multiple equal cost shortest paths.

When presented with equal cost paths, the SPB congruency properties can only be achieved if all bridges make the same path choices, requiring independence of the computation order and of the network position of the computation. Each bridge has a bridge ID, which ensures uniqueness. A PATHID is specified as a lexicographically sorted list of the bridge IDs the path traverses. Thus, all nodes implementing the same logic choose the same path from the available options (e.g., the one having the lowest PATHID). By using a set of globally defined transformations of the bridge ID prior to sorting, different paths are selected. 802.1aq specifies 16 bridge ID transformation methods and makes possible the application of further tie breaking methods to choose from the nearly infinite set of other possible permutations.

The initial 16 equal cost trees (ECTs) have unique attributes. First, the path congruence property means that 802.1aq actually supports equal cost routing for multicast and broadcast traffic as well as unicast. Second, since assignment of traffic to a path is done at the ingress to the 802.1aq network, the operator has the ability to forgo random assignment and place the traffic based on estimates of utilization, which can be considered a very lightweight form of traffic engineering (TE). For example, all traffic for one subnet, both unicast and multicast, in a data center can be constrained to one set of shortest paths, while traffic for a different subnet can be constrained to a second.

In order to foster continued exploration in this area of 802.1aq traffic placement, a framework has been created to allow future extensions to the equal cost modes. Research is already demonstrating techniques for even greater mesh utilization and controllability.

Figure 3 gives a concrete visual representation of 802.1aq's multipath routing features. Shown are all 16 individual shortest paths superimposed on the network between a given pair of nodes (1 and 20). The distribution compares favorably with the hop-by-hop hash-based approaches of IP ECMP; however, unlike ECMP, 802.1aq permits per service assignment by the network operator. In addition to tuning the distribution of offered load by adjusting the assignment of service to route, the operator can also tweak the ECT behaviors, giving a high degree of traffic control without resorting to full TE protocols.

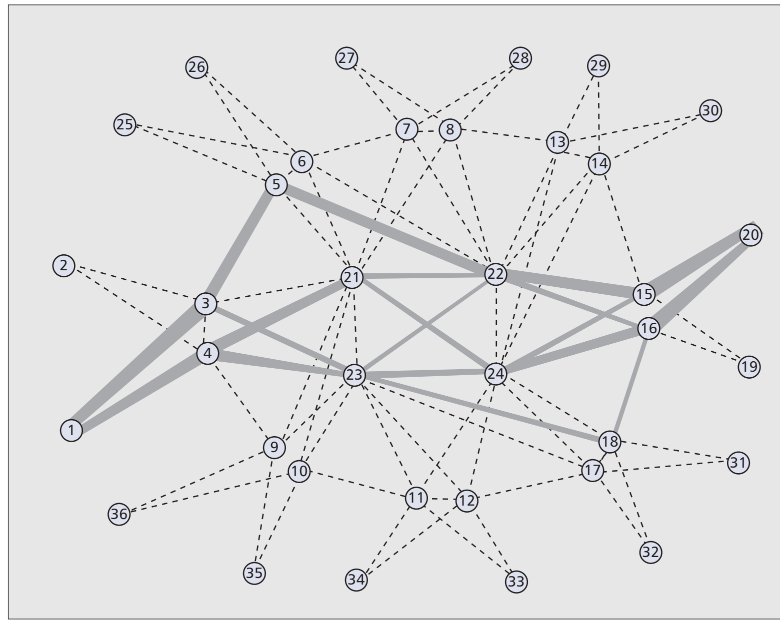


Figure 3. *802.1aq network emulation view of all 16 shortest paths between a pair of nodes. An operator can select any of them per service.*

Application Scenarios

A number of scenarios are now described in the following sections in order to illustrate the application benefits of SPB.

Campus Backbone

Campus environments today contain large numbers of Ethernet switches providing access to an even larger set of IP subnets to corporate networks. These environments have predominantly resisted the use of STP beyond the first layer of access switches due to the traditionally complex nature of configuration, the scalability limitations of STP, and the lack of natural traffic spreading, which results in unused links.

802.1aq addresses these issues while offering a vastly simpler management model and extending the size of the campus Ethernet domain significantly. Provisioning of a VLAN endpoint requires only the single-touch provisioning of a service identifier at the new point of attachment to the VLAN. The use of the SPBM service identifier (I-SID) gives the network the theoretical ability to scale to 16 million such VLAN services, a substantial advance on the 4094 VLANs of IEEE 802.1Q.

SPBM also provides added security by hiding the true infrastructure from the users, while Ethernet OAM provides the most complete and capable set of tools in the industry. These tools can provide insight into network operation that was previously unavailable to network staff.

Data Center

Many requirements of large data centers are satisfied by 802.1aq. The scale and scope of these makes this application the industry benchmark for layer 2 networking. 802.1aq's L2VPNs are an excellent fit with the enterprise data center applications' requirements.

The first requirement is the ability to keep the very large number of virtual machine (VM) MAC addresses hidden from the forwarding tables of transit switches, thereby reducing the cost of those switches. This is a natural consequence of the MAC-in-MAC data path of SPBM.

The second is the ability to easily support VM movement from one physical server to another, by growing or shrinking an E-LAN as necessary and dynamically leveraging SPB's endpoint-only based provisioning model. Providing E-LANs, and consequently IP subnets, that are

IEEE Communications | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | qMags

Substantial complexity has hitherto been needed to achieve loop freeness when interworking the Ethernet and PW control planes, particularly for multi-homed Ethernet clients. With 802.1aq, this is no longer needed.

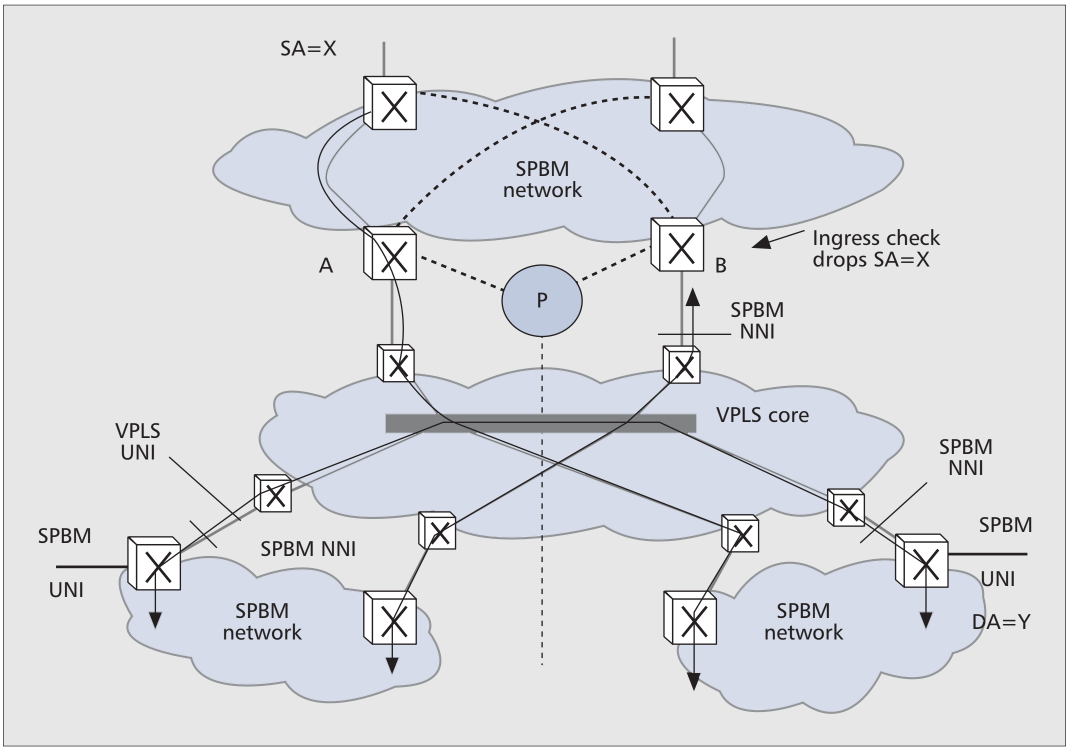


Figure 4. *SPBM overlay of bridged Ethernet emulation.*

decoupled from the physical topology is key to energy reduction in a data center. By allowing busy VMs to be logically clustered together, idle servers can be freed so that they can be shut down. Geographic independence of access to an E-LAN means a VM can be moved with a minimum of administrative overhead (e.g., using only Address Resolution Protocol [ARP] mechanisms, and hence no reprovisioning at layer 3).

A third requirement of a data center is to make better use of very regular mesh connectivity (e.g., fat trees). The current 16 ECTs of 802.1aq allow for vast improvements in utilization, and are easily extensible and customizable going forward.

A fourth requirement of a data center is for cluster-based load spreading. The perfect logical layer 2 VPNs created by 802.1aq allow layer 2 ARP-based multicast load balancing techniques to operate transparently, while also allowing easy growth or shrinkage of the clusters to any locations within the data center.

A fifth requirement of a data center is control over how traffic is placed. The head-end assignment of a subnet to a set of ECTs allows a degree of TE to be performed without the need for traditional heavierweight protocols.

Finally, speed of detection of and recovery from a network failure is critical to the data center environment, as mission-critical information and applications are being hosted. The use of IS-IS for 802.1aq intrinsically provides fast restoration. SPBM can be further improved by setting up a multicast community, which 802.1aq inherently and readily supports, for use by ISIS-SPB itself, to allow very fast data plane broadcast of link state updates. Early prototypes of this mode have validated the expected performance improvements.

Carrier Ethernet Services

MEF carrier Ethernet services [8] are the fastest growing provider service offering. Delivery of these frequently uses virtual private LAN service (VPLS) [9, 10]. SPBM can offer significant operational simplification to VPLS deployments, in addition to the scalability benefits of using PBB overlaid onto VPLS.

VPLS offers an emulation of Ethernet over MPLS, using split horizon forwarding from the network edge (PE) with a mesh of PWE3 pseudo-wires for each service instance. Initially deployed *flat*, VPLS had scaling limitations due to the meshed pseudo-wires. The hierarchical VPLS (H-VPLS) model mitigates this, using hub-and-spoke backhaul onto S-PE gateways at the edge of the split-horizon core. However, the desired reduction in mesh size causes customer MAC address scaling issues and complexity in tandem sparing.

To address the C-MAC scaling problems, there has been significant interest in PBB [1] to *front-end* H-VPLS [11, 12], because then only B-MACs are switched at the core gateways, C-MACs being confined to the network edge. SPBM builds on this PBB-VPLS model, directly inheriting its C-MAC hiding properties, and adding further attributes that make it an improved and more scalable *wrapper* around H-VPLS than even PBB.

Substantial complexity has hitherto been needed to achieve loop freeness when interworking the Ethernet and PW control planes, particularly for multihomed Ethernet clients. With 802.1aq, this is no longer needed: the Ethernet control plane can coordinate all aspects of loop avoidance directly, and has the scale to overlay the VPLS LAN service.

IS-IS already has capabilities that facilitate this: IS-IS supports LAN segments as native topology components, by modeling them as a star topology, with each node creating a single synthesized adjacency to a virtual IS-IS speaker called a Pseudonode. This minimizes the number of IS-IS adjacencies required.

This is illustrated in Fig. 4, showing SPBM regions surrounding a VPLS core, and using dual-homing for resiliency and load sharing. The SPBM gateways all use the IS-IS pseudonode model to construct the *node* P representing the entire VPLS core.

First, observe that all paths across the VPLS core will pass through the pseudonode P, but all tree construction will use the SPBM tie-breaking algorithm for ECT calculation. Thus, there can be only a single path per B-VID between any two SPBM bridges.

Second, while VPLS floods and learns, SPBM teaches by filtering. For example, a packet to a B-MAC previously unknown to VPLS (Y) will enter VPLS through bridge A (the shortest path from the source [X] to pseudonode P). VPLS, being a learning technology, will flood the packet. However, the SPBM filtering surrounding VPLS will only permit the frame to re-enter the SPBM domain at the closest interface to Y. One go and return transaction between X and Y is sufficient to teach VPLS the unicast path between SPBM domains as determined by ISIS-SPB.

Next, SPBM learns remote C-MAC addresses against a nodal B-MAC and not a pseudo-wire, so no remote *C-MAC unlearning* will be needed after a fault, because the B-MAC of the endpoint will remain unchanged.

Resiliency can be enhanced and VPLS MAC table flushing eliminated by splitting VPLS LAN segments (e.g., along the gray dashed line transecting P). VPLS will now offer two services, each single-homed onto each SPBM area. Fault recovery times will now be dictated by SPBM reconvergence only, because SPBM has alternate paths through VPLS, so VPLS restoration will no longer be needed to restore connectivity.

Finally, this SPBM wrapper model will be significantly more operationally scalable. VPLS needs to offer only one or two services per SPBM B-VID, not a service per end customer. Service instantiation, a single touch per point of presence (PoP), is handled by the SPBM I-SID. The core VPLS connectivity operates at the infrastructure level, as it should for scalability. This holds directly only for unicast, because customer multicast would have to be broadcast over this infrastructure. However, Ethernet already has a way to achieve per service multicast at the backbone layer. When VPLS PEs implement the standard Ethernet MMRP functionality, SPBM can signal the required multicast trees into VPLS automatically. This can also be achieved if the VPLS PE integrates SPBM interworking directly.

Wireless IP Backhaul

The radio access network (RAN) for fourth-generation (4G) wireless significantly extends the requirements on the wireline technology over previous generations. These requirements include:

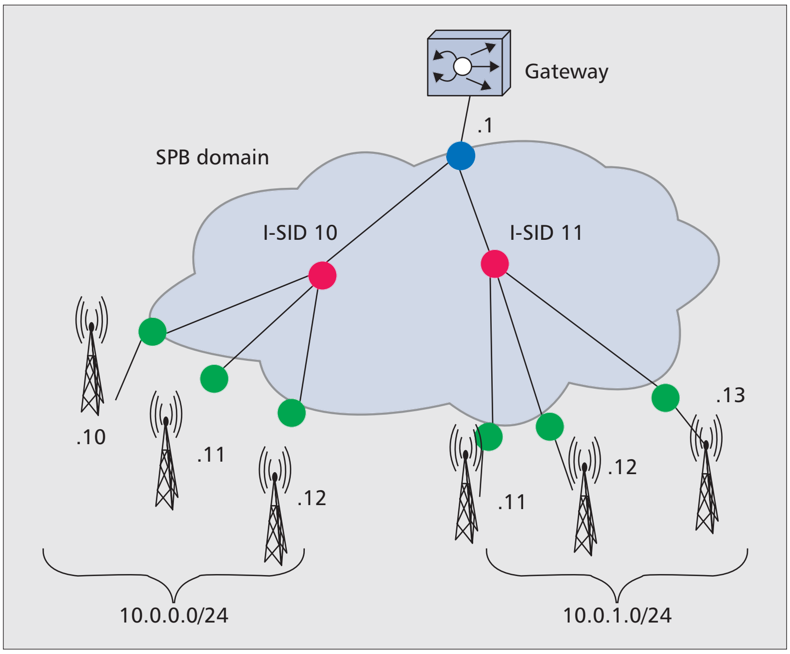


Figure 5. *Use of 802.1aq L2VPN in the RAN backhaul.*

- A flat end-to-end IP/Ethernet approach
- Direct communication between a NodeB (radio base station) and the access gateway
- Direct communications between NodeBs
- Ethernet interfaces to NodeBs/gateways
- Ethernet OAM
- Support for L2VPNs of all types; E-LAN, E-TREE, and E-LINE

IEEE 802.1aq can satisfy these requirements at low operational cost points (Fig. 5). This section will describe how SPBM may be used to offer a complete backhaul infrastructure for a flat RAN.

The approach to using SPBM for campus networks, described in a previous section, can be reapplied to IP subnet management for wireless backhaul. SPBM provides a flexible method with which to manage the virtual layer 2 VLAN topology used for IP subnets between base stations and gateways.

This model mimics the use of IP addressing and subnet management traditionally used over independent LAN topologies, as campus networks have done for years.

Multiple VLAN topologies can be created within the same backhaul environment, one for each IP subnet. The backhaul network can then manage the base stations on each subnet. This allows regional base stations to talk directly to each other on the LAN, with the advantage that all communications between them will take the shortest path (with the choice of path engineerable with the ECT features previously discussed).

By analogy with a campus environment, this is similar to the way in which desktops are placed on common IP subnets within a VLAN, and a router provides routing between VLANs and the rest of the network. At the same time all machines on the same VLAN are configured to be on the same subnet and can therefore communicate directly with each other without needing to be routed back at a central point in the

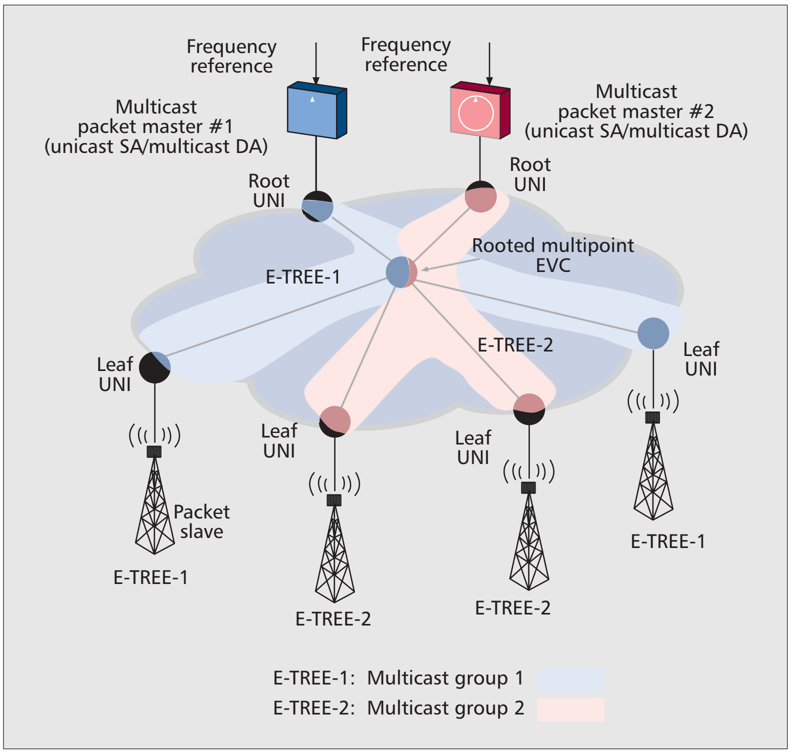


Figure 6. *Multicast clock distribution over 802.1aq E-TREE.*

topology. This distributed communication model allows for direct communication between Long Term Evolution (LTE) X2 interfaces, for example, and prevents a routing point from becoming a bottleneck for traffic that would otherwise need to be hair-pinned up and down through the backhaul network.

Flat RAN and Clock Distribution over 802.1aq E-TREE

In addition to requiring either a layer 2 or 3 VPN technology for backhaul, various mobile technologies require clock frequency distribution and synchronization (and in some cases phase/time synchronization) without relying on GPS or time-division multiplexing (TDM) approaches. The International Telecommunication Union — Telecommunication Standardization Sector (ITU-T) is currently developing and specifying an IEEE 1588 frequency profile (and in the future a time profile), which will allow a packet master clock to exchange 1588 timing messages with a number of packet slave clocks (e.g., Global System for Mobile Communications [GSM]/Universal Mobile Telecommunications System [UMTS]/LTE base stations). In the ITU-T frequency profile the network itself is seen as a cloud. However, the properties of the cloud can have an impact on the quality and management of clock distribution.

The important characteristics for clock distribution (and frequency distribution) are:

- The same timing message (carrying frequency information via timestamps) sent by the master can be used by several slaves.
- The messages sent by the master can be sent using unicast or multicast transmission.
- Timing messages can be exchanged in one or both directions, and when exchanged in both directions the path between the master and slave must be congruent (symmetrical) to permit one-way delay to be estimated. Without this path symmetry, complex corrections must be applied. This symmetry property is a fundamental requirement for phase/time sync.
- The messages must be delivered over the network path that offers the lowest latency and delay variation.
- The amount of sync traffic must be kept as small as possible.
- Timing loops are not permitted, to prevent slaves from becoming master, self-clocking, and so on.

Figure 6 shows packet masters (root user-network interfaces [UNIs]) providing a frequency reference to a number of packet slaves (leaf UNIs). In the example described below, a rooted multipoint EVC is used to provide a synchronization service using multicast transmission. Each master distributes timing messages to slaves via a multicast group (e.g., E-TREE-1 or E-TREE-2). The characteristics listed above blend very well with 802.1aq, MEF E-TREE services and rooted-multipoint EVCs as an example. For proper multicast clock distribution, the following constraints must also be respected:

- Exchange of slave-to-slave (leaf-to-leaf) timing messages is not permitted.
- Exchange of master-to-master (root-to-root) timing messages is not permitted (except in special cases), which can be implemented by an independent E-TREE for each root.
- The distribution of timing messages can be unidirectional master-to-slave (root-to-leaf) or bidirectional (root-to-leaf and leaf-to-root).

SPBM offers the ability to easily create E-LANs for the L2VPN requirements of mobile backhaul, while simultaneously allowing for easy creation of E-TREEs with the desirable properties outlined above for 1588v2 multicast clock packet distribution.

Summary and Future Work

SPB provides a new control plane for Ethernet networks while keeping the simplicity of Ethernet frame forwarding. SPB incorporates Ethernet OAM and the earlier scalability enhancements offered by the MAC-in-MAC encapsulation first introduced in PBB. SPB provides key features for efficient frame forwarding and improved network utilization, and furthermore makes possible the deployment of even larger Ethernet networks.

Further opportunities exist to exploit the merits of link state control in Ethernet, and there are several areas being explored for future IEEE 802.1aq extensions.

The first of these is referred to as IP shortcuts and is based on a very simple concept: IP routing normally forwards to the MAC address of the next hop router. This work would allow an IP router to forward to the last-hop MAC address, thus providing simple IP-VPNs on top of an SPB network.

The second area being explored is hierarchical 802.1aq to allow for even greater scalability and build on the IS-IS hierarchical models.

The third area of active exploration is increased network utilization with new equal cost shortest path algorithms and forwarding paradigms.

SPB provides a new control plane for Ethernet networks while keeping the simplicity of Ethernet frame forwarding. SPB incorporates Ethernet OAM and the earlier scalability enhancements offered by the MAC-in-MAC encapsulation first introduced in PBB.

References

[1] IEEE 802.1ah, "IEEE Standard for Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks — Amendment 6: Provider Backbone Bridging," Nov. 2007.

[2] IEEE 802.1aq D3.0, "IEEE Draft Standard for Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks — Amendment 9: Shortest Path Bridging," June 2010.

[3] IETF, "IS-IS Extensions Supporting IEEE 802.1aq Shortest Path Bridging"; draft-ietf-isis-IEEE-aq-00.txt

[4] ISO/IEC 10589, "Information Technology — Telecommunications and Information Exchange Between Systems — Intermediate System to Intermediate System Intra-Domain Routing Information Exchange Protocol for Use in Conjunction with the Protocol for Providing the Connectionless — Mode Network Service (ISO 8473)," 2nd ed., 2002.

[5] IEEE 802.1Q, "IEEE Standard for Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks," 2005.

[6] IEEE 802.1ag, "IEEE Standard for Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks — Amendment 5: Connectivity Fault Management," Dec. 2007.

[7] D. Allan *et al*., "Provider Link State Bridging," *IEEE Commun. Mag*., Sept. 2008.

[8] MEF, "Ethernet Services Definitions — Phase 2, Approved Draft 6," MEF D00057_006, Feb. 2008.

[9] IETF RFC 4762 "Virtual Private LAN Service (VPLS) Using Label Distribution Protocol (LDP) Signaling," Jan. 2007.

[10] IETF RFC 4761 "Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling," Jan. 2007.

[11] IETF, "Extensions to VPLS PE model for Provider Backbone Bridging," draft-ietf-l2vpn-pbb-vpls-pe-model-01

[12] IETF, "VPLS Interoperability with Provider Backbone Bridges," draft-ietf-l2vpn-pbb-vpls-interop-00

Biographies

David Allan (david.i.allan@ericsson.com) is senior program manager for standards and strategy at Ericsson, and a former distinguished member of technical staff at Nortel. He has been active in data telecommunications standards for the past 15 years including WG chair roles in the Broadband Forum and IETF. He has been active for over 25 years as an architect, design engineer, and developer of real-time systems in diverse areas of technology ranging from process control and avionics to financial transaction processing. His current role at Ericsson is focused on carrier infrastructure based on MPLS and Ethernet. He has a B.Eng. (1978) from Carleton University in Ottawa.

Peter Ashwood-Smith (peter.ashwoodsmith@huawei.com) is a Huawei Fellow with B.S and M.S. degrees in computer science from the University of Toronto. He has worked on the design, standardization, implementation, deployment, and support of many modern routing protocols (e.g., early label switched networks including ATM, MPLS, and GMPLS). He has approximately 30 networking related patents. His interests and research are now tending in the direction of computation-based/non-label-switched protocols.

Nigel Bragg (nbragg@ciena.com) has degrees from Trinity College, Cambridge, and Southampton University. He has spent 20 years in the telecommunications industry, initially in a contract product development environment, and for the following 13 years with Nortel. He has contributed to a wide variety of projects during that time, ranging from design leadership on a high-performance voice switching system to many aspects of data switching and routing, to automated power optimization in optical line systems. For the last six years he has focused on carrier packet transport, and is a co-inventor of PBT and PLSB. He has recently joined Ciena.

Janos Farkas (janos.farkas@ericsson.com) has been a research engineer at Ericsson Research since 1999. He is an active contributor to the IEEE 802.1aq Shortest Path Bridging project. He has focused on carrier transport networks based on Ethernet in the past few years. His former research activities include IP QoS solutions for mobile access networks and network traffic management. He has 18 patent applications and 12 research papers in the area of telecommunications networks. He has a Master's degree in electrical engineering from Budapest University of Technology and Economics

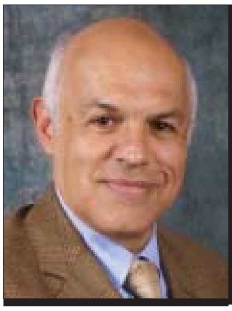
Don Fedyk (Donald.Fedyk@alcatel-lucent.com) is an authority on routing system design for both connectionless and path oriented routing. He has been an active contributor to several IETF WGs in the routing area including MPLS and CCAMP. During the last couple of years he has focused on the IEEE, specifically on control planes for Ethernet bridges. He is currently the Editor of the Shortest Path Bridging project (IEEE 802.1aq). He is an inventor on approximately 20 communications related patents, and holds B.S. and M.S. degrees in electrical engineering from the University of Waterloo, Ontario, Canada.

Michel Ouellette (michel.ouellette@huawei.com) is a technical advisor in Huawei's IP Network Solutions and Clock Laboratory, where he focuses on mobile backhaul networks, and the development and analysis of packet network architectures/protocols for accurate frequency and phase/time distribution. He participates in ITU-T and IEEE 802.3 standardization. Prior to joining Huawei, he spent 12 years at Nortel focusing on ATM/TDM pseudowires, synchronous Ethernet, clock algorithms for base stations, TCP/IP active queue management, and ATM switching. He has been granted 10 patents and has published several international journal papers. He received his B.A.Sc. and M.A.Sc. from the University of Ottawa in 1995 and 1997, and attended l'Ecole Nationale Superieure des Telecommunication.

Mick Seaman (mickseaman@sbcglobal.net) is an independent consultant working on protocols and operational methods for Ethernet deployment in public and mission-critical networks. He has led the design, implementation, and standardization of routing and bridging switches since the mid-1980s — at DEC, 3Com (VP & CTO Enterprise Systems Business Unit, 1996-99), and Telseon (Founder, EVP & CTO) — and holds more than 20 patents in the field of switching system and protocol design. He chaired the IEEE 802.1 Interworking (1986–2007) and Security (2004–2010) Task Groups, and was responsible for the IEEE 802.1D, 802.1Q, 802.1X, and 802.1AE standards. He has a Master's degree in physics from Cambridge University.

Paul Unbehagen (paul.unbehagen@alcatel-lucent.com) is an active member of the IEEE 802.1 WG. He has also participated in several IETF WGs to include IS-IS, BGP, L2VPN, and IPVPNs, and is currently the author of the IP/SPB IETF draft and has several patents in communication protocols such as IS-IS and BGP. Previously he has worked in numerous diverse networking environments to include the Department of Defense, Bloomberg, MCI, and Nortel. He has had a diverse background working not only in a vendor but also on live networks at a Tier 1 service provider, ASP, the U.S. military, as well as a few startups. He thus has 14 years of deployment, operational, network design, and architectural experience in live networks ranging from enterprise to carrier. He has a B.S in CIS and a B.S. in BA from NC Wesleyan College.

STANDARDS FOR BROADBAND ACCESS AND BEYOND




Mostafa Hashem Sherif

Yoichi Maeda

With more applications coming to rely on high-speed access to the Internet, bringing broadband connectivity to residences and businesses as well as to mobile terminals is growing in importance. Yet, year after year, the data collected by the Organization for Economic Cooperation and Development (OECD) have shown that, among the developed countries, those with a large urban population such as South Korea, Japan, France and the Netherlands are more likely to achieve a higher rate of broadband penetration than those with significant rural communities such as the United States and Canada [1]. Figure 1 plots the data series in Table 1, which represents the broadband penetration per 100 inhabitants vs. the percentage of total landmass used by 50 percent of the population for various OECD countries.

Recently, a study by Oxford University's Saïd Business School and Universidad de Oviedo, Spain, sponsored by Cisco, to evaluate broadband access in the OECD countries confirmed these trends, as summarized in Table 2 [2].

A significant contributing factor to this state is that the cost of deployment of current technological solutions makes them prohibitive in rural and sparsely populated areas. Because the lack of high-speed connectivity can have adverse effects on the overall growth of economic activities, finding less costly ways for broadband connectivity over long distances is an active area of research, development, and standardization.

The first two articles in this section present recent standards in this direction. Vladimir Oksman *et al.* in their article, "The ITU-T's New G.993.5 Standard Proliferates 100 Mb/s," present a functional description of the vectored techniques used in that ITU-T Recommendation, previously known by its code name G.vector. They document its performance, as calculated from theory and as observed in laboratory measurements, and show that it allows the minimum span between repeaters to be increased. They also give guidelines on the performance tradeoffs when the number of vectored subscribers is large.

Improvements in digital subscriber line (DSL) technology are significant because this is the technology of choice for broadband access in many countries. Very high rate DSL (VDSL) technology defined in ITU-T Recommendation G.993.1 first increased the bandwidth available to asymmetric traffic (up to 100 Mb/s downstream and 50 Mb/s upstream). Next, the ITU-T approved Recommendation G.993.2 as a second version of VDSL in 2006. VDLS2 delivers 100 Mb/s symmetrical traffic over copper lines using Ethernet in the first mile per the specifications of IEEE 802.3ah. In a typical installation the fiber runs up to a distribution point in the neighborhood or in the basement of apartment buildings, while the VDSL2 connection takes over to the customer premises [3]. The new Recommendation G.999.5 overcomes a significant limitation of G.993.2 because it increases the connection distance to about 500 m. This is achieved by reducing the far-end crosstalk interference in the DSL access multiplexer (DSLAM) through a transmission technique called vectoring.

The second article, "IEEE 1901 Access System: An Overview of Its Uniqueness and Motivation" by Shmuel Goldfisher and Shinji Tonabe, presents an alternative approach to high-speed access using IEEE 1901 for broadband over power line (BPL). An IEEE 1901 access cell consists of a number of repeater stations, a head-end station to manage the whole cell

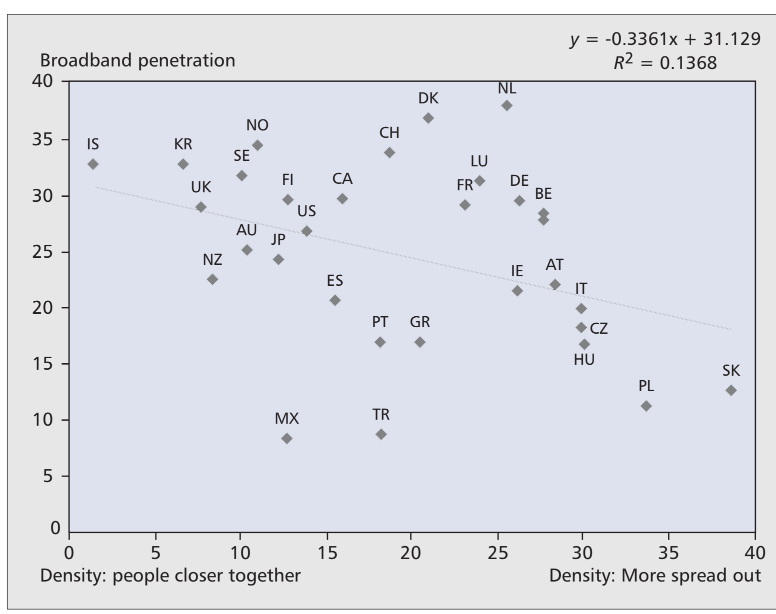


Figure 1. *Broadband penetration per 100 inhabitants and percentage of total land mass used by 50 percent of the population (June 2009). Source: OECD Broadband statistics.*

Country name	ISO 3166 Country Code	Percentage of land mass used by cumulative 50% of the population	Broadband penetration per 100 inhabitants (June 2009)
Australia	AU	10.36	24.9
Austria	AT	28.42	21.8
Belgium	BE	27.63	28.4
Canada	CA	15.91	29.7
Czech Republic	CZ	29.76	18.1
Denmark	DK	20.99	37.0
Finland	FI	12.77	29.7
France	FR	23.08	29.1
Germany	DE	26.29	29.3
Greece	GR	20.49	17.0
Hungary	HU	30.05	16.8
Iceland	IS	1.40	32.8
Ireland	IE	26.19	21.4
Italy	IT	29.99	19.8
Japan	JP	12.31	24.2
Korea	KR	6.68	32.8
Luxembourg	LU	24.08	31.3
Mexico	MX	12.70	8.4
Netherlands	NL	25.52	38.1
New Zealand	NZ	8.31	22.8
Norway	NO	10.97	34.5
Poland	PL	33.69	11.3
Portugal	PT	18.04	17.0
Slovak Republic	SK	38.63	12.6
Spain	ES	15.54	20.8
Sweden	SE	10.08	31.6
Switzerland	CH	18.69	33.8
Turkey	TR	18.25	8.7
United Kingdom	UK	7.72	28.9
United States	US	13.91	26.7

Table 1. *Broadband penetration per 100 inhabitants and percentage of total land mass used by 50 percent of the population (June 2009). Source: OECD Broadband statistics.*

and to connect it to the Internet, and network termination stations bridging in-home networks, both wireless and wired. The article describes the cell topology with a focus on access station addressing, clock synchronization, and bandwidth management and allocation to ensure efficient use of the available channel.

The next two articles are updates of presentations at the 2008 ITU-T Kaleidoscope Conference. The article titled "The Creation of a Ubiquitous Consumer Wireless World through Strategic ITU-T Standardization" by Máirtín O'Droma and Ivan Ganchev proposes a migration path for wireless communications from the current subscriber mode of operation to one with the user at its center. The key concept is to equip the mobile device with a personal identifier instead of the current subscriber identifier and to assign it location-independent IPv6 addresses. In a sense, this evolution mirrors what happened with mobile payment instruments in that they are no longer tied to the customer's bank. The article contains a

Rank	Country	Broadband penetration (% of households)	Broadband quality score
1	South Korea	97%	66
2	Japan	64%	64
3	Hong Kong	99%	33
4	Sweden	69%	57
5	Switzerland	90%	40
6	Netherlands	83%	46
7	Singapore	96%	32
8	Luxembourg	99%	27
9	Denmark	82%	45
10	Norway	84%	38

Table 2. *Highlights of the Saïd Business School/Universidad de Oviedo study on broadband access [2].*

detailed comparison of the attributes of both consumer-centric and subscriber-based modes of operation. It identifies the standards needed to allow third-party operation authentication, authorization and accounting of mobile users as well as call connection establishment.

The final article, "Standards Dynamics through an Innovation Lens: Next Generation Ethernet Networks," is by Tineke Egyedi and one the co-Editors of the standards series (Sherif). They present a framework to analyze the co-evolution of standards and technology, and to anticipate potential change-related problems. New complex technologies depend on the cooperation of many actors, including standards development organizations, particularly with "open innovations." However, for the cooperation to be successful, each party to the collaboration needs to take into account the goals and objectives of the other parties, and the constraints under which they work. The ideas are illustrated with the activities of both the IEEE and ITU-T on the standardization of next-generation Ethernet.

It is hoped that these four articles will stimulate discussions among the readership and encourage further activities on broadband standardization. Finally, the Editors would like to thank the reviewers who graciously accepted to go over the various iterations of the submissions and whose judicious comments greatly improved their quality.

References

[1] OECD Broadband Portal; http://www.oecd.org/document/54/0,3343,en_2649_34225_38690102_1_1_1_1,00.html

[2] Saïd Business School, Oxford Univ. and Univ. de Oviedo, "Global Broadband Quality Study Shows Progress, Highlights Broadband Quality Gap," London, U.K., Oct. 1, 2009; http://www.sbs.ox.ac.uk/newsandevents/Documents/BQS%202009%20final.pdf; also available at http://www.cisco.com/web/MT/news/09/news_021009a.html.

[3] P. E. Eriksson and B. Odenhammar, "VDSL2: Next Important Broadband Technology," *Ericsson Rev.*, no. 1, 2006; also available at http://clanspace.com/r0/download/1456040~fe59ec9cfe508f4d71494c761cae14c7/vdsl2.pdf, pp. 36–47.

Biographies

MOSTAFA HASHEM SHERIF (ms5285@att.com) has been with AT&T in various capacities since 1983. He has a Ph.D. from the University of California, Los Angeles, an M.S. in the management of technology from Stevens Institute of Technology, New Jersey, and is a certified project manager of the Project Management Institute (PMI). Among the books he authored are *Protocols for Secure Electronic Commerce* (2nd ed., CRC Press, 2003), *Paiements électroniques sécurisés* (Presses polytechniques et universitaires romandes, 2006), and *Managing Projects in Telecommunication Services* (Wiley, 2006). He is a co-editor of two books on the management of technology published by Elsevier Science and World Scientific Publications in 2006 and 2008, respectively, and is the editor of the *Handbook of Enterprise Integration* (3rd ed., Auerbach, 2009).

YOICHI MAEDA [M] (yoichi.maeda@ntt-at.co.jp) received B.E. and M.E. degrees in electronic engineering from Shizuoka University, Japan, in 1978 and 1980, respectively. Since joining NTT in 1980, for the last 26 years he has been engaged in research and development on access network transport systems for broadband communications including SDH, ATM, and IP. From 1988 to 1989 he worked for British Telecom Research Laboratories in the United Kingdom as an exchange research engineer. He currently leads the standardization promotion section of NTT Advanced Technology Corporation and is NTT's senior adviser on sstandardization. In October 2008 at the World Telecommunication Standardization Assembly (WTSA-08), he was appointed chair of ITU-T SG15 for the 2009–2012 study period for his second term. He is a Fellow of the IEICE of Japan. He has been a Series Editor of the Standards Series in *IEEE Communications Magazine* since 1999.

Communications | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | qMags

ITU STANDARDS

The ITU-T's New G.vector Standard Proliferates 100 Mb/s DSL

Vladimir Oksman, Heinrich Schenk, and Axel Clausen, Lantiq Inc.
John M. Cioffi, Mehdi Mohseni, and George Ginis, ASSIA Inc.
Carl Nuzman and Jochen Maes Alcatel-Lucent
Miguel Peeters, Broadcom Corporation
Kevin Fisher, Ikanos Communications
Per-Erik Eriksson, Ericsson AB

ABSTRACT

This article explores the recently issued ITU-T G.vector (G.993.5) [1] that allows expanded use of 100 Mb/s DSL. A tutorial description on G.vector's crosstalk noise reduction methods leads to specific projections and measurements of expanded DSL 100 Mb/s reach. A discussion on dynamic maintenance to enhance G.vector's practical application then concludes this article.

INTRODUCTION

Modern life increasingly depends on faster Internet access for applications such as email, voice, information search, and a variety of social interconnections, while IPTV is increasing Internet video use. Growing wireless connectivity will substantially increase access bandwidth demand over the next decade.

Figure 1 shows digital subscriber line (DSL) technology as the current undisputed leader in broadband access. With more than 300 million worldwide subscribers, DSL use is consistently higher than cable modems and passive optical networks. A substantial part of Fig. 1's fiber to the X (FTTX) subscribers are actually very high rate DSL (VDSL2) connections where a fiber runs to an intermediate network point, usually within a neighborhood or in the basement of a multidwelling unit, and the DSL comprises the remaining connection to the customer. G.vector raises VDSL2 connection speeds up to 100 Mb/s at distances beyond 500 m from the fiber termination point with no transmit power increase and no Shannon-Law violation — G.vector simply removes most of DSL's crosstalk noise, thus providing a very high throughput. A similar principle is used in well-known gigabit Ethernet connections.

A copper twisted pair's throughput is unshared, and each individual DSL customer's speed often exceeds those of many other broadband connections. Furthermore, DSL's *fiber to the cabinet* architecture significantly reduces fiber deployment cost by sharing it between hundreds of customers connected to the cabinet via existing copper, enabling a more profitable DSL broadband business case. However, current VDSL2 provisions 100 Mb/s bit rates only over very short distances and, accordingly, requires too many cabinets. G.vector provides the business case's missing ingredient; it significantly increases the 100 Mb/s range, and thus enables reasonable broadband connection cost. We present some test results and illustrate how G.vector supports such a low-capital-cost VDSL2 deployment of 100 Mb/s.

The existing copper connections are ready for use immediately, but can exhibit enormous variability in signal attenuation and noises. Additionally, the line's state depends on the particular installation, and customers' touching of lines or moving of equipment. Today, the operational costs associated with DSL trouble call response, dispatch of technicians (*truck rolls*), and customer service drop/change (*churn*) dominate DSL operational costs, and are a concern for higher-speed DSL network enhancements. As G.vector removes most of the far-end crosstalk (FEXT) noise of other VDSL2 lines sharing the cable, the nonstationary noise and other remaining uncancelled line noises become increasingly important. Thus, link management using line monitoring and optimization techniques, known as dynamic spectrum management (DSM), becomes crucial to retaining G.vector's fundamental gains.

The shorter-line high-speed DSL concept using fiber-fed street cabinets or curb boxes first appeared in standards in 1994 [2], culminating in International Telecommunication Union — Telecommunication Standardization Sector (ITU-T) Recommendations G.993.1 (VDSL1, 2004) and G.993.2 (VDSL2, 2006). VDSL enlarges asymmetric DSL (ADSL) architecture by increasing the number of subcarriers to cover wider bandwidth [3]. Vectoring evolves the multiple-input multiple-output (MIMO) signal processing first suggested by Paulraj [4] for multiple-antenna wireless. Adaptation of this vectoring technology to more stationary channels

Communications Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page qMags

first appears in [5] and was continually refined for DSL in 1999 and 2000 [6]. A reduced complexity near-optimal implementation of linear vectoring was proposed in [7]. Complexity of vectoring is still challenging, especially for large numbers of vectored subscribers.

Vectored DSL proposals first appeared in the American National Standard Institute (ANSI) in the 2001 DSM project [8] as the highest of three levels of crosstalk noise control methods for copper loop management. Vectored DSLs have physically separated customer premises (CP) locations that cannot be coordinated directly by a common customer-side controller. Recommendation G.993.5 describes the necessary interoperable line coordination functions at the DSL access multiplexer (DSLAM) and the individual lines' training protocols within the coordinated vector group.

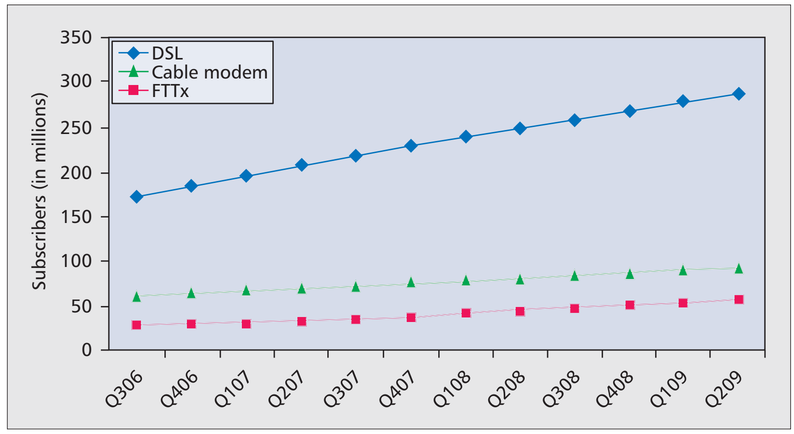


Figure 1. *Broadband access-connection for 2006–2009 (source: Point Topic).*

G.993.5 Standard Vectored Technology

FEXT Cancelation Principles

The dominant VDSL2 noises are near-end crosstalk (NEXT, between transmitters and receivers connected to different pairs of a multipair cable at the same end), far-end crosstalk (FEXT, between transmitters and receivers connected to different pairs of a multipair cable at opposite ends; Fig. 2), and background Gaussian noise [3]. NEXT coupling is usually so strong above 1–2 MHz that VDSL2 systems use nonoverlapping downstream/upstream frequency bands multiplexing up to 30 MHz. Thus, with NEXT largely eliminated, FEXT may dominate the remaining noise. G.vector technology cancels VDSL2's mutual FEXT, thus effecting a performance improvement.

In a typical VDSL deployment, multiple VDSL2 lines connect the DSLAM to the VDSL transceiver unit — remote terminals (VTU-Rs) at physically separated individual residences' CPs. VDSL2 uses digital multitone (DMT) modulation [9] with up to 4096 subcarriers located on frequencies f_i, spaced by $\Delta f = 4.3125$ kHz or $\Delta f = 8.625$ kHz ($f_i = i \times \Delta f$, $i = 0, 1, \ldots, 4095$). Each subcarrier carries a certain number of bits that depends on this subcarrier's signal-to-noise ratio (SNR). FEXT cancellation reduces noise and thus can increase these SNRs. This allows carriage of more bits and therefore increases the line's data rate.

FEXT generated by a particular twisted pair m into a victim twisted pair l can be cancelled at the subcarrier frequency f_i by subtracting from the received signal the value of the signal $Um(f_i)$, transmitted over the pair m, multiplied by the FEXT transfer function from the pair m into the pair l, $H_{l-m}(f_i)$. If the cable binder includes N pairs, and FEXT has to be cancelled on M subcarriers, the subtracted crosstalk signal comprises $(N - 1) \times M$ components, where each component corresponds to a particular binder pair's FEXT into the victim line on the particular subcarrier frequency. Each subcarrier's (f_i) crosstalk signal is obtained by multiplication of the signal vector $\mathbf{u}$ and the FEXT coupling vector $\mathbf{H}$, both of size N; hence the name *vectoring*.

The DSLAM side must perform vectoring for both downstream and upstream, because DSLAM alone contains all the cable's DSL signals (Fig. 2a). A downstream FEXT precoder precedes the modulation, and an upstream canceller follows the demodulation. The vectoring control entity (VCE) supplies updated channel matrices and controls the FEXT cancellation process, such as indicating from which lines to cancel FEXT. Such control allows efficient use of the available processing power to cancel dominant crosstalk.

G.vector defines interoperability only for downstream vectoring, since upstream vector processing needs no interoperability specification. However, G.vector also specifies certain VTU-R control signals and timing that facilitate upstream vectoring: Figure 2a illustrates a potential vectoring implementation.

Figure 2b defines each downstream line's *precoding* for FEXT cancellation. The precoder, for each subcarrier frequency f_i, multiplies the precanceller matrix $\mathbf{C}_p^{(f_i)}$ corresponding row by the signal vector $\mathbf{u}^{(f_i)}$ of N lines. The VTU-R's received vector, $\mathbf{y}^{(f_i)}$, will be: $\mathbf{y} = \mathbf{H} \cdot \mathbf{C}_p \cdot \mathbf{u} + \mathbf{r}$, where $\mathbf{H}$ is the channel transfer matrix, $\mathbf{r}$ is the received non-FEXT noise vector, and the precanceller matrix, $\mathbf{C}_p$, is equal to the inverted normalized channel matrix, $\mathbf{H}_0 = [\mathbf{diag}(\mathbf{H})]^{-1} \cdot \mathbf{H}$, and $\mathbf{C}_p = \mathbf{H}_0^{-1}$. Similarly, for upstream FEXT cancellation, the received signal vector can be described as $\mathbf{y} = \mathbf{C}_c \cdot \mathbf{H} \cdot \mathbf{u} + \mathbf{r}$, and differs from the downstream case since the $\mathbf{H}$ is different.

Vectoring also involves timing requirements. The DSLAM strictly aligns all lines' upstream and downstream DMT symbols to typically within 1 μs. The DSLAM adjusts each vector group VTU-R's upstream timing advance value to meet this small tolerance. Also, all vectored lines transmit downstream sync symbols at the same time, while VTU-Rs transmit all upstream sync symbols also at the same time. This alignment eliminates sync-symbols' FEXT into data symbols, as in Fig. 3. To align all upstream vectored lines' sync symbols, the DSLAM sends the VTU-R a special time marker during the line's initialization, which indicates the time offset between the upstream and downstream sync symbols, and thus allows the new line's VTU-R to align its transmitted sync-symbol position with that of other lines.

Communications | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | qMags

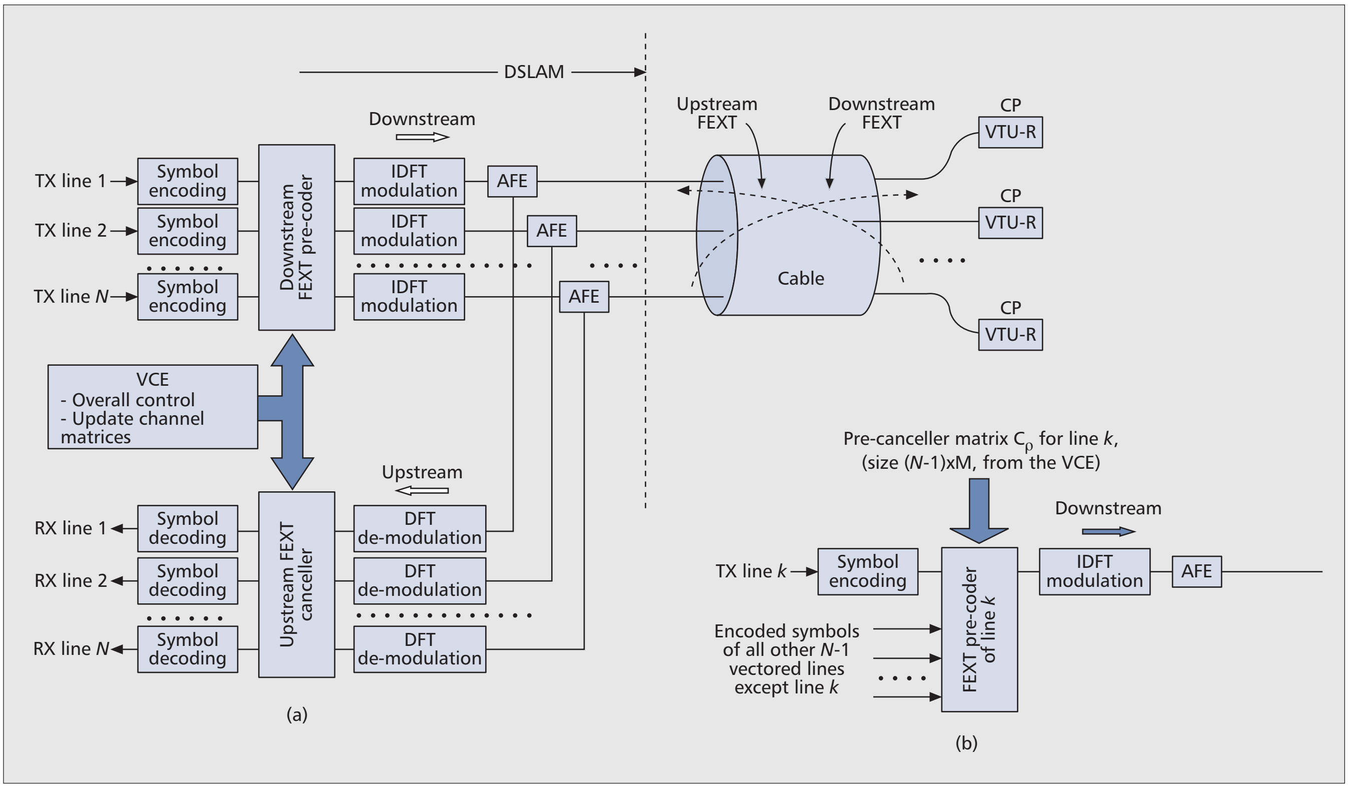


Figure 2. *Functional description of upstream and downstream vectoring.*

FEXT Estimation

FEXT cancellation requires the FEXT coupling coefficients between all pairs of the vectored group at each subcarrier. FEXT coupling estimation between two pairs uses repeating pilot signals to determine the other pair's FEXT component. Since FEXT that is not yet cancelled during joining may cause unacceptably high noise, the pilot signal is transmitted only during sync-symbols. All vectored lines' sync-symbols thus are aligned in time and carry no user data; therefore, even full-power pilot signals sent during sync-symbols do not disturb vectored lines' data transmission (Fig. 3).

VDSL2 transmits a sync symbol every 256 data symbols. For FEXT estimation, a special binary pilot sequence modulates each line's sync symbol on pre-assigned *probe-tone* subcarriers with indices equal to $10n$, $10n + 2$, $10n + 3$, $10n + 4$, $10n + 5$, $10n + 6$, $10n + 8$, and $10n + 9$ ($n = 0, 1, 2, \ldots$). *Flag-tone* subcarriers with indices equal to $10n + 1$ or $10n + 7$ allow communication of standard VDSL2 online reconfiguration (OLR) signals. The reduced number of OLR-carrying subcarriers does not impact OLR robustness because the number of flag tones is still large enough (vectored lines are relatively short and use wide spectrum). The value of FEXT on flag tones is interpolated from adjacent probe tones.

The DSLAM assigns different vector-group lines' binary pilot sequences, which are usually selected as mutually orthogonal. The orthogonality speeds up and simplifies FEXT estimation that correlates victim lines' measured receiver error values with the disturbing lines' orthogonal sequences. One popular class of orthogonal sequences are Walsh-Hadamard sequences, for which the length can be any power of 2.

Apart from crosstalk estimation's use of orthogonal pilot sequences, direct FEXT estimation methods like a least mean square (LMS) algorithm can be used successfully . Selection of a particular pilot sequences is vendor discretionary; each DSLAM vendor can fit the sequence to their preferred FEXT estimation algorithm. G.vector also describes an alternative FEXT estimation method that uses each vector group's reported SNR values, as measured at the VTU-R [1].

Reporting of Error Samples

The VCE needs feedback from the VTU-R on errors that are signal distortions caused by uncompensated downstream FEXT in the pilot sequences of received downstream sync-symbols on subcarriers assigned for downstream FEXT precoding. Samples of errors help the adaptive determination of the FEXT coupling coefficients used in the precanceller matrix. Thus, each VTU-R's measured receiver error samples are reported via the backchannel to the VCE for FEXT estimation and precanceller matrix ($\mathbf{C}_p$) computation. G.vector assures CP-to-DSLAM interoperability by standardization of backchannel parameters and formats of the reported error samples. Backchannel capacity and reported error sample accuracy are critical for the FEXT cancellation algorithm's convergence speed.

Error Samples — A given subcarrier's error sample is the difference $E = Z - D$ between the received complex signal Z and its intended constellation point D, normalized to the scale of a unit-magnitude constellation point. The VCE

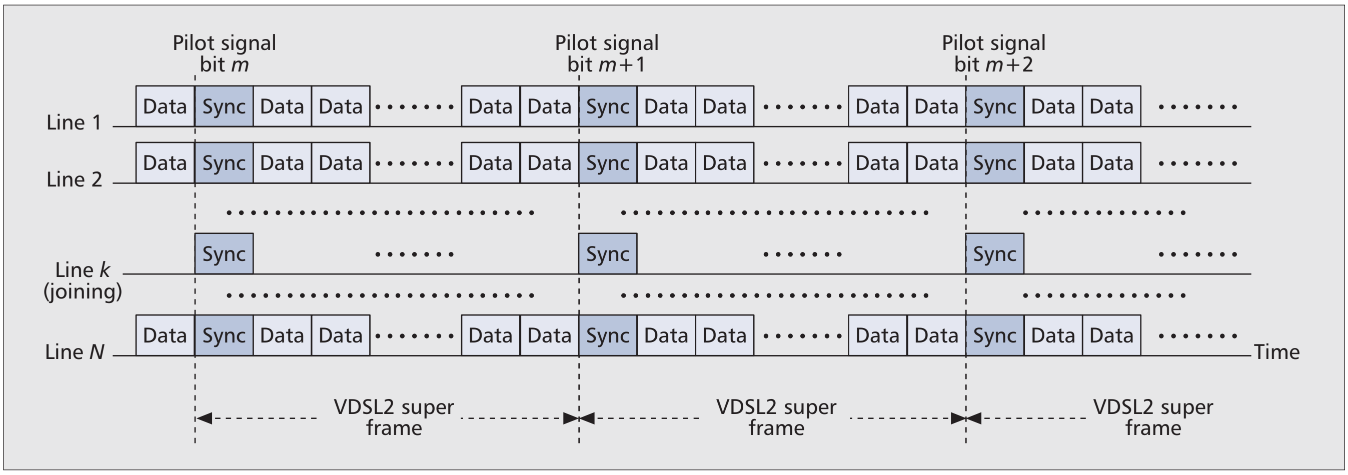


Figure 3. *Symbol alignment, sync-symbol alignment, and pilot bits in a vectored group. (Same in upstream and downstream; line* k *is at the beginning of the joining procedure).*

divides the downstream frequency spectrum into up to eight non-overlapping bands for flexible reporting. A particular band's sync-symbol error samples are reported in a block floating point format that supports different backchannel capacities, FEXT estimation algorithms, and deployment scenarios.

Grouping error samples into blocks saves backchannel bandwidth, since a common 4-bit exponent applies to both real and imaginary parts of all the block's samples; this is effective because neighboring subcarriers' error samples typically have similar magnitudes. A block may contain one error sample, 32 error samples, or all error samples of the band. Further bandwidth may be saved by frequency and time subsampling of the reported error samples (e.g., reporting only even subcarriers' error samples or from only every third sync-symbol). Clipping of error sample components to a specified maximum magnitude to avoid a particular block's exponent is unduly influenced by powerful impairments on certain subcarriers, such as radio frequency ingress (RFI), causing a loss of precision for the remaining subcarriers. The VCE configures the size of the mantissa between 0 (sign only) and 8 bits.

The VCE configures the bands and format of error sample reporting, and can adjust the error feedback accuracy by selecting an appropriate mantissa length and matching the backchannel bandwidth through grouping and subsampling. The VTU-R may also flag a particular report as potentially corrupted (e.g., by impulse noise) and report the error's mean value over the vectored band, thereby assisting the DSLAM assessment of FEXT estimation completeness. With all these means, the DSLAM can effectively address different corner situations, such as excessive SNR variations over the vectored band (due to strong RFI ingress or bridged taps) or very limited bandwidth of the backchannel, and efficiently use the available DSLAM processing power.

Backchannel — The backchannel conveys error samples from the VTU-R to the DSLAM. Three transport mechanisms are specified for flexible backchannel operation: over the special operations channel (SOC) during line initialization, and over the embedded operations channel (EOC) or Ethernet channel during showtime. All three mechanisms use the same error sample formats presented earlier.

Later we describe the O-P-VECTOR 2 stage initialization's use of the SOC-based transport mechanism. Each error report is transmitted as a message, encapsulated in a high-level data link control (HDLC) frame. This transport mechanism is a high-speed VDSL2 SOC that conveys the initialization messages. The DSLAM configures the backchannel data rate and the VTU-R's used error report format via SOC messages.

A DMT symbol's repetition of the same information on several subcarriers achieves SOC protocol robustness. The repetition rate determines the capacity of the SOC, which can range from 16 to 192 b/DMT symbol, providing bit rates from 64 kb/s to 768 kb/s, respectively, for 4000 symbols/s transmission. Transmission with a reduced number of repetitions (relative to VDSL2) is still reliable for vectored VDSL2 due to shorter distances.

VDSL2 uses EOC to convey OLR and operations, administration, and maintenance (OAM) messages. The EOC bit rate is fixed and at a set level below 256 kb/s. For transportation over EOC, error reports are mapped into a standard EOC message and transported with high priority. The format of the error samples is configured through an EOC command sent from the DSLAM.

The Ethernet backchannel has a flexible data rate: the error reports are encapsulated in Ethernet frames and multiplexed with the upstream user data. At the DSLAM, the received Ethernet packets are identified by the CP's medium access control (MAC) address and delivered to the VCE's assigned MAC address. When multiple lines connect the CP to the DSLAM (bonded connection), a CP-assigned ID marks each line. The error samples' reported format is configured as with the EOC-based backchannel. The VCE and CP MAC addresses and the line ID are set during initialization.

By specifying both Ethernet and EOC mechanisms, G.993.5 offers a choice of cancellation algorithms and system architecture. Some cancellation algorithms may require a higher

Data symbols carry the joining CP-line's reported error samples via the SOC backchannel, extended to higher capacity. The DSLAM sets the error-sample format based on the actual SOC throughput and the required error precision.

backchannel peak bit rate than available through the EOC, while some system architectures may exclude reported-error-sample multiplexing among the user data.

Operation of a Vectored Group

Vector group operation comprises three phases: tracking, joining, and leaving. In tracking, no new lines join or leave the group; each line tracks routine FEXT-coupling variations, mainly caused by temperature changes. Tracking's FEXT matrix update is usually very slow, because low-accuracy infrequent error sample reports suffice as derived from a low-speed backchannel.

A vectored group transitions to the joining phase when one or more lines initialize to join the group. First, the joining line transmits only sync-symbols carrying pilot sequences, as in Fig. 3, line *k*. Existing lines then estimate each joining line's FEXT and update their FEXT cancellation matrices without disruption. Furthermore, joining lines accommodate existing lines' FEXT. A joining event requires much higher backchannel throughput than tracking, so the backchannel throughput flexibility accommodates the required change. For that, Ethernet backchannel packets are assigned high priority or the EOC is pre-configures to provide sufficient throughput. Some auxiliary EOC transactions, such as performance monitoring, may be deferred to provide faster joining. Lower backchannel capacity slows the joining event.

Leaving events may be orderly or disorderly. Orderly leaving first terminates transmission in both directions, thus allowing the CP modem to safely disconnect. Disorderly leaving corresponds to sudden CP modem disconnect, power off, or line disconnect, which can potentially increase other vectored DSLs' residual FEXT during several seconds, until the DSLAM and CP modem both detect the disconnect and stop transmitting. Disorderly leaving is currently under study.

Initialization

G.vector line initialization adds new stages to VDSL2 initialization (darker blue boxes in Fig. 4), allowing new lines to join seamlessly. During initialization, prior to steady-state transmission (showtime), VDSL2 modems estimate the channel transfer function (*channel discovery phase*), adapt receiver parameters to the channel (*training phase*), and further compute and exchange bit loading tables between the VTU-O and VTU-R (*channel analysis and exchange phase*). At showtime, small adaptations compensate for natural channel and noise changes.

The handshake phase starts initialization, when the two sides exchange capabilities and agree on a common operational mode. During the VECTOR-1 stage, crosstalk from the joining lines into vectored lines is estimated and compensated; after this stage the initializing line can use standard full-power VDSL2 training signals without disrupting vectored lines. The VECTOR-1 signals consist only of modulated sync symbols transmitted simultaneously with other vectored lines' sync symbols (Fig. 3, line *k*). The VECTOR-1-1 stage is similar to VECTOR-1 and readjusts the crosstalk cancellation from the joining lines into vectored lines after potential changes in FEXT coupling during the channel discovery phase (usually due to deviations in impedances of the modems).

During VECTOR-2 stage joining, lines estimate and compensate crosstalk from vectored lines. After VECTOR-2, all joining lines are ready to compute SNR and FEXT-compensated bit loading. VECTOR-2 signals include sync-symbols modulated by pilot sequences and initialization data symbols at all other symbol positions. These sync symbols allow FEXT estimation from vectored lines into each of the joining lines. Data symbols carry the joining CP-line's reported error samples via the SOC backchannel, extended to higher capacity. The DSLAM sets the error sample format based on the actual SOC throughput and required error precision.

To join several lines (or for a full startup), the VCE synchronizes all joining lines' VECTOR initialization stages, providing a sufficiently long time period of simultaneously transmitted sync-symbols with pilot sequences, and then terminates them at the same time. The VCE thereby acquires the same FEXT estimation data from all vectored lines during VECTOR-1 and VECTOR-1-1, and from all joining lines during VECTOR-2.

Figure 4's two plots illustrate SNR convergence to its steady-state value (FEXT cancelled) for the case of one line joining a group of 31 vectored lines (downstream sync-symbols' subcarrier #684). The upper plot shows the vectored lines' worst SNR (on line 26), and the lower plot shows the joining line's SNR. In both cases, convergence times indicate the required duration of VECTOR-1 and VECTOR-2 transmission, respectively. This duration grows with the number of joining lines and the overall size of the vectored group. The convergence of the upstream SNR looks very similar.

Performance of Vectored System

Theoretical Capacity of Vectored VDSL2

Figure 5 shows downstream bit rates vs. VDSL2 loop length for a North American residential deployment (Profile 17a) operating over 17 MHz with and without vectoring. As each disturber's FEXT magnitude into each subscriber varies based on the particular binder-pair location, cable type, manufacturing tolerances, and deployment practices, the expected vectored VDSL2 gains appear from a statistical perspective using four families of curves. Besides FEXT, simulations also include –140 dBm/Hz additive white Gaussian noise.

Upper and lower performance bounds appear with upward and downward facing triangles. The upward facing triangles show each loop length's FEXT-free bit rate. Downward facing triangles show the bit rate using a standard (conservative) 99 percent worst-case FEXT model. These upper and lower bounds reflect VDSL2's potential capacity range.

The blue crosses show a large non-vectored VDSL2 group's simulated bit rates, assuming a statistically representative loop-length distribution and the NIPP-NAI statistical crosstalk model [10], for 100-pair (4 × 25-pair binders) North

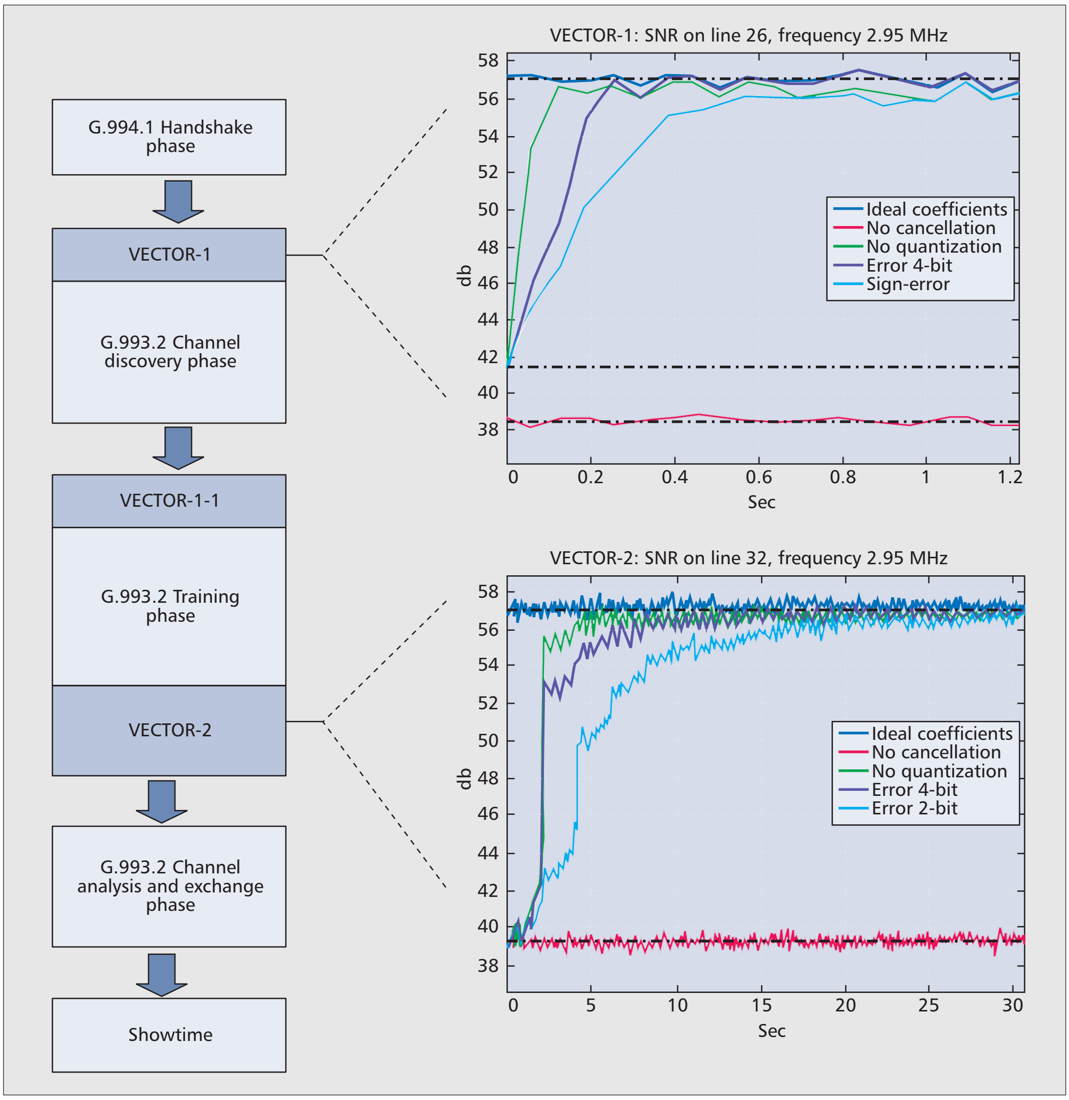


Figure 4. *G.995.3 initialization timeline and simulation results for SNR in vectored and joining line (1 lines joins a group of 31 lines AWG-24 500 m).*

Practical vectored performance gains depend heavily on implementation specifics and service-provider deployment practices. Preservation of vectoring performance gains critically involves all active cable pairs at least to ascertain which pairs actively contribute to any specific lines' crosstalk.

American wire gauge (AWG) 26 cable. Each blue cross represents an individual cable line's data point, where 96 of the 100 pairs carry VDSL2 signals and 4 pairs are unused. Most lines achieve a bit rate higher than the worst-case bit rate; however, there is no easy advance prediction of a particular line's achievable bit rate. Most service providers consequently provision their service based on the worst-case assumption.

The red circles in Fig. 5 show bit rates for the same vectored VDSL2 lines. Since the entire cable employs vectoring, FEXT-free rates are closely achieved. Thus, vectoring can significantly increase FEXT-limited downstream bit rates on loops shorter than 800 m. Similar upstream bit rate improvements are also possible.

Cancellation of all cable crosstalk is usually not necessary, but only the dominant disturbers. Partial cancellation suggests balancing the implementation complexity with the desired deployed bit rate. The presented example cancelled only each line's 63 most significant FEXT sources from 95 (approximately 2/3 of disturbers). Each individual line has its own unique set of significant disturbers based on the line's relative position in the cable, as well as on the cable's construction. Thus, partial cancellation necessarily pre-observes crosstalk across the entire cable to identify each line's unique set of top disturbers.

Practical vectored performance gains depend heavily on implementation specifics and service provider deployment practices. Preservation of vectoring performance gains critically involves all active cable pairs at least to ascertain which pairs actively contribute to any specific lines' crosstalk.

Laboratory Test Results

These vectored VDSL2 17-MHz profile (17a) prototype measurements were taken on 26 AWG and European Telecommunications Standards Institute (ETSI) 0.5 mm cables to validate the following characteristics:

- The bit rate as a function of loop length, with and without crosstalk
- The number of a binder's crosstalking lines that have to be cancelled to achieve significant performance increase
- The impact of adjacent-binder FEXT

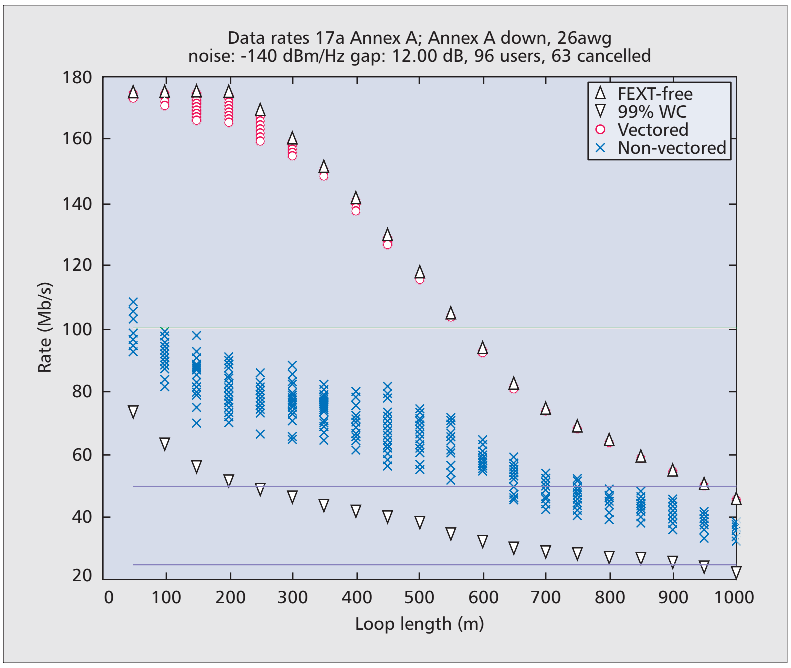


Figure 5. *Downstream bit rates for vectored and regular VDSL2 (Profile 17a, bandplan EU32).*

A 24-port vectored-VDSL2 DSLAM prototype was connected to 50-pair 26-AWG cables of different lengths, comprising two binders with 25 pairs each. Only 24 pairs of 50 were simultaneously accessed. The cable was divided into 500-ft sections, allowing sectional cable extension, while keeping pair consistency at the connection points. Measurements were taken for three test setups:

1. Successive single-line activation to measure each line's FEXT-free performance
2. All lines activated with no vectoring
3. All lines activated with vectoring to record each line's performance with FEXT from other 23 lines cancelled

Figure 6 shows measured vectored results that achieve almost FEXT-free performance. The difference in performance between the 0.5 mm ETSI and 26 AWG cables is caused by different impedance (around 120 Ω for ETSI cable, whereas the 26 AWG impedance is 100 Ω). The rate cap around 100 Mb/s is due to the particular prototype's implementation. Tests for VDSL2 30-MHz profiles are expected in the near future.

Properties for a Vectored VDSL2 System — One vectored VDSL2 property is that its steady-state performance becomes more predictable than legacy VDSL2. One example is the reduced bit rate variation caused by FEXT variation. Another example of the increased predictability is negligible performance variation when lines join or leave. In legacy VDSL2 systems, new joining lines cause unpredictable performance reduction in other lines due to FEXT, and even loss of synchronization. Vectored systems no longer exhibit this instability.

Numbers of Dominant Disturbers and Impact from Adjacent Binders — Measurements also showed how many lines need to be FEXT-cancelled for a significant performance increase. Best vectored performance often requires cancellation of most disturbers in the same binder; cancelling five or six dominant disturbers is not sufficient. The average FEXT from adjacent binders can be expected to be about 10 dB lower than intra-binder crosstalk, although adjacent binders can still contain a number of lines that must be treated as strong crosstalkers. The overall conclusion is that the strongest crosstalk occurs within the same binder with less crosstalk from the adjacent binder.

Management

The management techniques and interfaces for vectored DSL relate to DSM level 3.

Management Challenges for Vectored DSL

First, it is important to manage performance trade-offs between customers' lines. One reason is that the VCE's computational resources can be insufficient to cancel all disturbers. Service providers' management systems can then set higher priority for lines with bandwidth-sensitive services. Also, algorithm settings can favor certain lines at the expense of others (e.g., the upstream vectoring decoding order).

Second, it is important to manage noise sources that become dominant after FEXT among the vectored lines is cancelled. Other noise sources, such as impulse noise, may become dominant and reduce the performance benefits. Sudden external noise changes may lead lines to re-initialize. Management systems can avert such disruptions by appropriate impulse noise protection configuration, or by other means of mitigating abrupt noise changes.

Finally, the management system can improve overall performance in mixed-binder cases where vectored and non-vectored VDSL2 lines coexist in the same cable or binder. If left unmanaged, such non-vectored lines' FEXT may eliminate significant vectoring benefit. The management system can adjust the non-vectored lines' transmitted power based on their service requirements to limit their impact on vectored lines, [11]. ADSL lines have much less impact due to the narrow bandwidth they use.

Management Interface of G.vector

The G.vector management interface has been standardized in ITU-T Recommendation G.997.1, which includes enhancements required for effective G.vector management and defines new management parameters.

A new XLOG test parameter reports crosstalk transfer function on downstream subcarrier basis. XLOG indicates the FEXT coupling strength between any two vector-group lines (for a detailed XLOG definition, see [1]). The XLOG enables some very useful management functions.

Crosstalk diagnosis: XLOG allows the identification of lines creating excessive FEXT. Typically, such lines are characterized by faults (e.g., poor balance) that lead to poor performance and require maintenance action. Additionally,

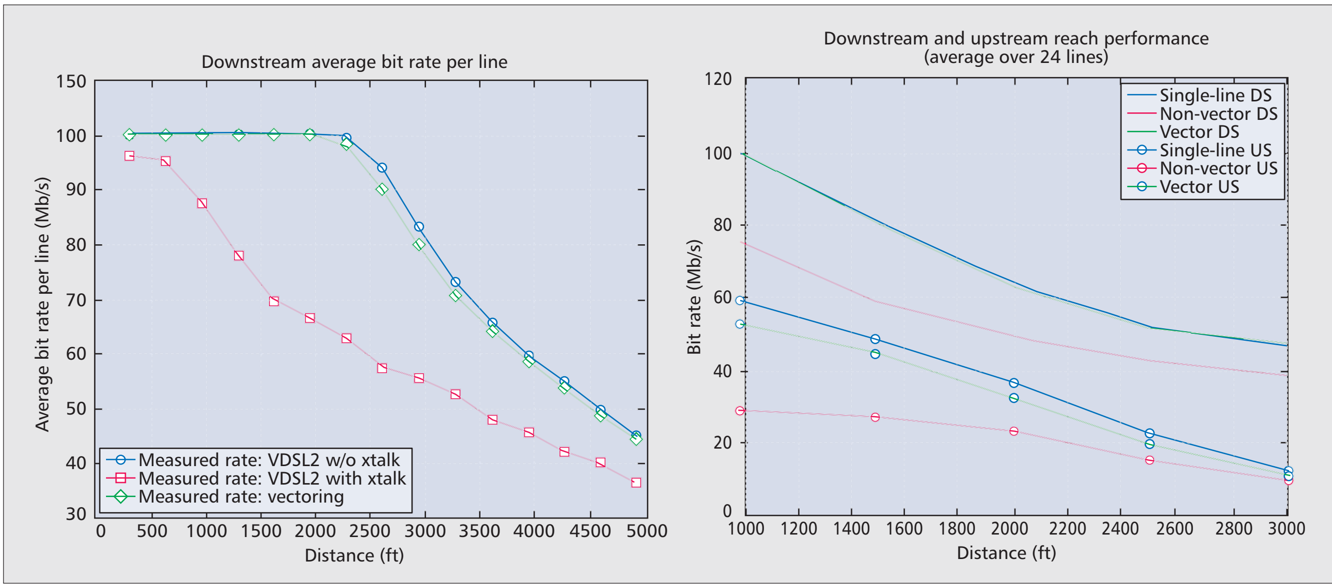


Figure 6. *Measured bit rate for 17a profile: left: downstream, 0.5 mm ETSI; right: downstream and upstream, 26 AWG; the curves for single-line DS and vector DS overlap.*

even though it may be possible to mitigate the FEXT induced by such pairs into vectored lines, high FEXT is likely to also be generated into non-vectored lines, which cannot be cancelled. By identifying such extreme crosstalk *polluters*, the copper network is improved over time.

Performance prediction: XLOG allows expected FEXT cancellation gain prediction for specific pairs. G.vector supports configuration parameters for allocation of vectoring computational resources to prioritized lines. Knowing the expected gains provides essential guidance for priority-level choice or even vector enablement choice.

The G.vector configuration parameters allow:
- Enabling or disabling a particular line's FEXT cancellation
- Selecting FEXT cancellation frequency bands
- Assigning each line's FEXT cancellation priorities
- Assigning each line's target data rate

A particular line's disabled-vectored capability allows service providers allocation of vectoring to maximize benefit (i.e., assisting high-end services to perform maximally). These higher-priority lines can benefit the most when there are limited computational resources.

Control of vectoring frequencies is useful when certain frequencies may suffer from *uncancellable* noises, like radio frequency interference (RFI) or other types of DSL (ADSL or legacy VDSL). The vectored system can then be instructed to disable vectoring on the corresponding subcarriers.

Finally, setting FEXT cancellation priorities is essential for allocating vector computational resources within a vectored group. Since the VCE's computational resources may be limited, the VCE may not be able to cancel all lines' FEXT into every vectored-group line. External guidance on the resource allocation may be based on each line's service requirements. As for legacy VDSL2, a vectored line must always maintain its bit rate between the configured minimum and maximum data rate. In addition, a G.vector system should attempt to allocate enough resources to at least achieve the target data rate. If resources are still available, the VCE should allocate those remaining resources to the lines with line priority set to HIGH. Lines with priority set to LOW might maintain a lower bit rate than the target data rate, as described in Fig. 7. Lines with different priority levels may also use different bit loading algorithms to help the VCE mitigate FEXT.

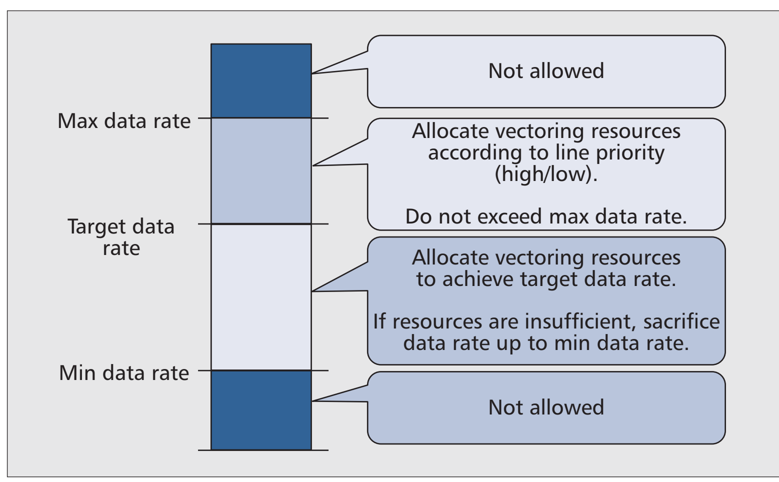


Figure 7. *Setting of line priorities and data rate configuration parameters.*

CONCLUSIONS

G.vector can extend the reach of standard high-speed VDSL2 systems significantly. This reach extension enables service providers to offer higher-speed services to more customers at a lower cost than was previously anticipated in DSL access networks. G.vector or G.993.5 will enable a large interoperable market for equipment and its associated dynamic management across an

G.vector's large DSL benefit is expected to accelerate video, voice, wireless (through backhaul of increasingly smaller cells offering more bandwidth to mobile users) and other high revenue-generating telecommunications services, at a time when such services are of particular interest.

array of vendors, ensuring cost-effective availability of high-speed DSL access networks worldwide. G.vector's large DSL benefit is expected to accelerate video, voice, wireless (through backhaul of increasingly smaller cells offering more bandwidth to mobile users), and other highly revenue-generating telecommunications services at a time when such services are of particular interest.

ACKNOWLEDGMENTS

The authors wish to thank Frank Van der Putten (Alcatel-Lucent), the ITU-T editor of G.993.5, Danny Van Bruyssel (Alcatel-Lucent), Nicholas Sands (Ikanos), Shailendra K. Singh (Ikanos), and Chenguang Lu (Ericsson AB) for simulation results, helpful comments, and discussions, and Tom Starr (AT&T Labs) for support and guidance in this work.

REFERENCES

[1] ITU-T Rec. G.993.5-2010, "Self-FEXT Cancellation (Vectoring) for Use with VDSL2 Transceivers."
[2] J. M. Cioffi and J. A. C. Bingham, "A Proposal for Consideration of a VADSL Standard Project," ANSI contrib. T1E1.4/94-183, Dec. 1994
[3] J. Cioffi *et al.*, "Very High-Speed Digital Subscriber Lines," *IEEE Commun. Mag.*, Aug. 2004.
[4] A. Paulraj and T. Kailath, "Increasing Capacity in Wireless Broadcast Systems Using Distributed Transmission/Directional Reception (DTDR)," U.S. Patent 5,345,599, Sept. 6, 1994.
[5] J. M. Cioffi and G. D. Forney, Jr., "Generalized Decision-Feedback Equalization for Packet Transmission with ISI and Gaussian Noise," Ch. 4, A. Paulraj, V. Roychowdhury, and C. Schaper, Eds., *Communication, Computation, Control, and Signal Processing: A Tribute to Thomas Kailath*, Kluwer, 1997.
[6] G. Ginis and J. M. Cioffi, "Vectored Transmission for Digital Subscriber Line Systems," *IEEE JSAC*, vol. 20, no. 5, June 2002, pp. 1085–1104.
[7] R. Cendrillon *et al.*, "A Near-Optimal Linear Crosstalk Precoder for VDSL," *IEEE Trans. Commun.*, May 2007.
[8] J. Cioffi and K. Song, "Level 3 DSM Results: Vectoring of Multiple DSLs," ANSI T1E1.4 contrib. 2002-059, Vancouver, Canada, Feb. 18, 2002.
[9] T. Starr, J. Cioffi, and P. Silverman, *Understanding DSL*, Prentice Hall, 1999.
[10] ATIS pre-published tech rep. ATIS-PP-0600024, "Multiple-Input Multiple-Output Crosstalk Channel Model," 2009.
[11] M. Mohseni, G. Ginis, and J. M. Cioffi, "Dynamic Spectrum Management for Mixtures of Vectored and Non-vectored DSL Systems," *44th Annual Conf. Info. Sciences and Sys.*, Princeton, NJ, Mar. 2010.

BIOGRAPHIES

VLADIMIR OKSMAN (vladimir.oksman@lantiq.com) received his M.S. and Ph.D. from the Leningrad Radio and Telecommunication State College, USSR, in 1976 and 1987. Currently, he is director of Technical Marketing at Lantiq. He is 34+ years in digital communications, developing the first long-distance digital transmission systems in USSR, early HDLS and ADSL, single-carrier and multicarrier VDSL systems, and home networking technologies. His recent efforts are devoted to vectored DSL and novel networking technologies for energy management. He actively participates in multiple standard development organizations.

HEINRICH SCHENK (heirich.schenk@lantiq.com) received Dipl. Ing. and Dr. Ing. degrees from Friedrich Alexander University, Erlangen, Germany, in 1974 and 1978. From 1979 to 1996 he worked for the Siemens Public Switching Division, Germany, as a senior engineer and head of the Development Division. Since 1997 he has worked in the field of concept and algorithm development in the Semiconductor Department of Siemens and later at Infineon. Presently he is a senior principal at Lantiq in Munich.

AXEL CLAUSEN (axel.clausen@lantiq.com) received his M.S.E. and Ph.D. in electrical engineering from Arizona State University in 1995 and 1999, respectively. He held different positions in system and concept engineering for broadband access, developing algorithms and system architecture for ADSL, VDSL, and home networking products. Currently, he is a principal system engineer in Lantiq, leading the development of system architecture and algorithms for vectored VDSL. He was actively involved in standardization of vectored DSL.

JOHN M. CIOFFI (jcioffi@assia-inc.com) is chairman and CEO of ASSIA Inc. and Hitachi Professor Emeritus, Stanford University. He received his B.S.E.E. in 1978 from the University of Illinois, his Ph.D.E.E. in 1984 from Stanford, and an honorary doctorate from the University of Edinburgh in 2010. He was the founder of Amati Communications Corp., 1991–1999, and the designer of the world's first ADSL modem and basic ADSL/VDSL patents. Affiliations and awards include the Boards of Alto Beam and ClariPhy; IEEE AG Bell (2010), Millennium (2000), Kobayashi (2001), and IEE Tomson (2001) medals; Marconi Fellow (2006); U.S. NAE (2001); FREng (2009); and ANSI Outstanding Achievement (1996).

MEHDI MOHSENI (mmohseni@assia-inc.com) joined ASSIA, Inc., in 2006 as a systems architect, focusing on DSL network optimization and improving management systems. His research interests incorporate multi-user information theory and optimization of wireless and broadband networks using convex optimization techniques. He holds a Bachelor's degree in electrical engineering from Sharif University of Technology, Tehran, Iran, as well as M.S. and Ph.D. degrees in electrical engineering from Stanford University.

GEORGE GINIS (gginis@assia-inc.com) is currently vice president of Expresse Engineering at ASSIA, Inc., overseeing development and deployment of the DSL Expresse management product. Between 2002 and 2005 he was a DSL systems engineer at Texas Instruments. He holds a Diploma in electrical and computer engineering from the National Technical University of Athens, and M.S. and Ph.D. degrees in electrical engineering from Stanford University.

CARL NUZMAN (nuzman@research.bell-labs.com) is a Distinguished Member of Technical Staff in the Enabling Computing Technologies Department, Bell Labs, Alcatel-Lucent. Recent research interests include vectored DSL systems, wireless power control, and optical networking. He holds a Ph.D. in electrical engineering from Princeton University, and B.S. degrees from the University of Maryland College Park.

JOCHEN MAES (jochen.maes@alcatel-lucent.com) received his M.S. and Ph.D. in physics in 2000 and 2004 from Katholieke Universiteit Leuven, Belgium, for his research on charge coupling in nano-structured semiconductor lasers of 1.3 μm and 1.55 μm wavelengths. Since 2006 he has been researching next-generation broadband access systems over copper at Bell Labs Alcatel-Lucent. He acts as a reviewer for several journals and conferences, and has published around 35 papers.

MIGUEL PEETERS (miguel.peeters@broadcom.com) is a graduate in engineering from the Universite Libre de Bruxelles and Ecole Centrale Paris. He joined Broadcom Corporation in 2000. For more than 10 years his activities have been related to the use of multicarrier modulations for DSL. He is a regular contributor to standard bodies like ITU-T, where he is editor of the G.997.1 Recommendation.

KEVIN FISHER (kfisher@ikanos.com) has more than 23 years of experience in the development of firmware, software, and complex silicon devices. His experience includes executive and senior-level engineering positions including CEO of Vector Silicon, VP of IC and DSP engineering at 2Wire, and department manager of advanced recording channels at Quantum. He earned his Ph.D. E.E. and M.S.E.E from Stanford University after a B.S. degree in computer engineering at the University of Illinois. He holds 20 U.S. patents.

PER-ERIK ERIKSSON (per-erik.s.eriksson@ericsson.com) is currently a senior research engineer at Ericsson Research working with next-generation DSL technologies involving VDSL2 vectoring, and technologies for high-frequency transmission over copper. He actively participates in the DSL standardization in ITU-T SG15/4. He holds an M.S. in electrical engineering from the Royal Institute of Technology, Stockholm, Sweden.

IEEE 1901 Access System: An Overview of Its Uniqueness and Motivation

Shmuel Goldfisher, MainNet Communication
Shinji Tanabe, Mitsubishi Electric Corporation

Disclaimer: The points of view expressed here are solely those of the authors, and in no way is it implied here that these points of view also are shared or supported by the IEEE P1901 work group.

ABSTRACT

In 2005 the IEEE P1901 Working Group began standardization activities for broadband over power line networks. The process is now in its final stages, and the latest P1901 draft standard is available for sale to the public. The standard is designed to meet both in-home multimedia and utility application requirements including smart grid. The utility requirements and the resulting features that support those requirements were clustered together and form the basis of what is referred to as the utility access cluster. This article explains the aspects of P1901 power line communication technologies designed to address the access cluster. The differences between access and in-home applications, including addressing methods, clock synchronization, smart repetition, quality of service, power saving, and other access unique mechanisms, are also explained.

INTRODUCTION

In June 2005 20 companies agreed to form the IEEE P1901 Working Group (WG) under the sponsorship of the IEEE Communications Society [1]. The resulting IEEE 1901 standard [2] is applicable to both in-home (IH) multimedia and utilities applications.

Access networks usually cover large areas, may consist of hundreds or thousands of nodes, and are usually centrally controlled. There are two main types of access applications offered by utilities: broadband applications and utility applications. Typical broadband applications include providing Internet data access and voice over Internet Protocol (VoIP). Typical utility applications include controlling energy use (smart grid) and building/factory controls.

Utility applications have been getting a lot of attention recently. For example, the IEEE established a dedicated portal for handling this hot topic [3]. There are two primary approaches for implementing utility applications over the power line network:

- A narrowband power line communication (PLC) low-speed approach, for which some solutions are already available
- A broadband PLC (BPL) approach such as the one developed by the IEEE P1901 Working Group.

This article begins by explaining the preference for the broadband PLC approach over a more conventional narrowband PLC approach. Next, we cover the access network's features and requirements (particularly utility systems) and compare them with those needed for IH systems. We then address some of the design differences in detail such as special addressing methods, clock synchronization, repetition, quality of service (QoS), power saving, and centralized time-division multiple access (TDMA) systems for large topology multihopping access systems.

WHY BROADBAND PLC IS NEEDED FOR LARGE-SCALE ACCESS SYSTEMS

The main differences between narrowband (low-speed) and broadband (high-speed) PLC are their bandwidths and the carrier frequencies they use. Figure 1 shows a typical inverter noise spectrum that may be found on the power lines. Narrowband PLC typically uses carrier frequencies below the U.S. amplitude modulated (AM) band (<500 kHz). At these frequencies the high noise floor reduces range and available bandwidth. Broadband PLC requires a much better signal-to-noise ratio (SNR), and typically uses carrier frequencies between 2 and 30 MHz where the noise tends to be less.

Higher-frequency PLC not only improves the data rates but also provides better transmission distances and coupling across circuit breakers and between lines.

Utility applications such as smart grid need 100 percent accurate data communications. The system has to be *commercial-grade* and very robust. Broadband PLC is able to choose the best of many carriers supported by a wider (2–30 MHz) band, a variety of modulation methods that adapt to the channel SNR, and a selective automatic repeat request (ARQ) mechanism in order to achieve a robust link. In addition, when there are a large number of nodes, broadband PLC may be needed to provide even small bandwidth allocations to each node.

For access applications like VoIP, utility applications, and Internet surfing, several levels of QoS may be required when these applications contend for bandwidth on the same medium. For example, some packet streams are isochronous video or audio streams, and others may be asynchronous real-time data. These data may require much higher bandwidth than utility applications in order to share the same medium.

It is expected that more demanding applications and commercial opportunities will arise in the future. Because utilities traditionally invest in technologies for the long term and have to predict their needs over that term, it is logical to select a broadband PLC solution with more potential rather than narrowband PLC.

The Main Differences between Access and In-Home Systems

In-home networks are relatively simple and small. The distance between stations is generally short, and the topology is typically a star or a tree. A repeating function is rarely required, and if it is needed, most of the time a single repeater between endpoints is enough. All of the devices in the network usually belong to and are managed by one owner.

It is a greater challenge to define a solution for access networks since they have no definitive structure. They can also contain tens, hundreds, or even thousands of stations managed by the same logical network manager. Different power line networks may topologically be a tree network, a mesh network, a ring network, or a hybrid structure comprising all three of them. As a result, the optimal Access solution has to be as flexible as possible in order to support all relevant topologies. Automation and optimization mechanisms that make use of the flexibility need to be incorporated into products in order to adapt each station to overcome topology-specific issues and simplify the installation.

The dynamic characteristic of access networks is another challenge. The number of customers, number of end stations, and location of devices (e.g., electric cars) can vary over time. The network itself may also change over time due to electric switches being opened or closed. In addition, the impedance of the line and external interference also affect the usable bandwidth. These challenges require scalable and dynamic solutions that are far beyond the requirements of an IH network.

As a result, the access protocol stack should be designed to support generic topologies, a high number of neighbors and edge station addresses, and fast repetition, and should adapt to dynamic changes.

IEEE 1901 Access System Topology

The access system defined by the IEEE P1901 Working Group has a cell structure. A cell is a group of stations managed and authorized by a single station, usually the head end. The cell is built from stations managed directly by a logical *cell manager* and additional network elements, which are managed indirectly using stations as proxies.

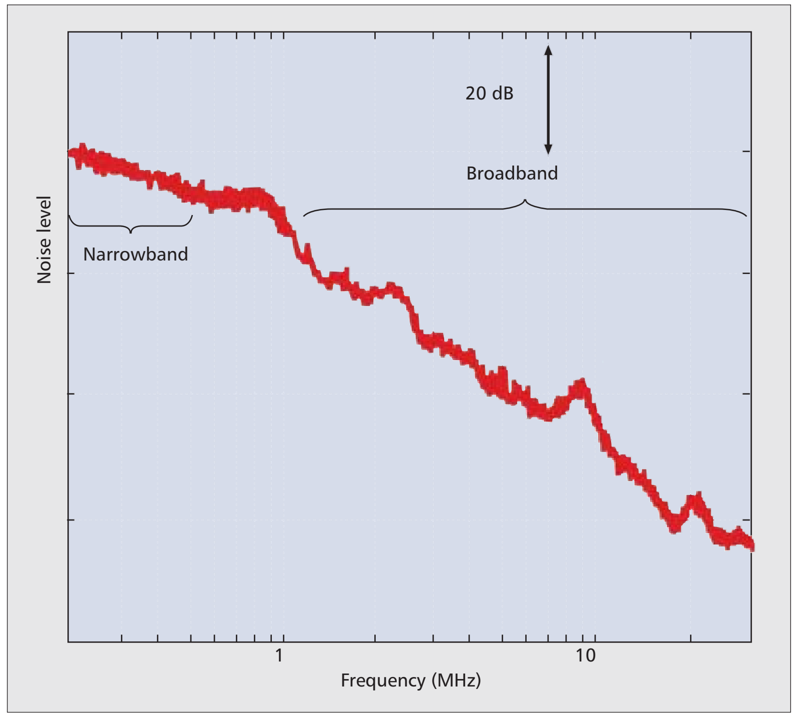


Figure 1. *Inverter noise spectrum.*

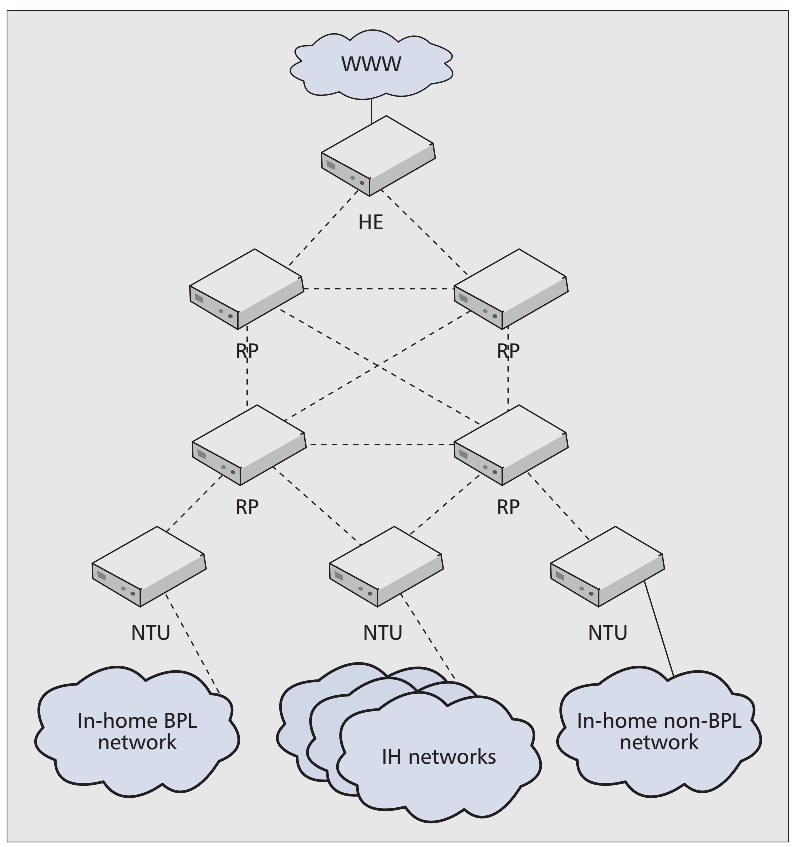


Figure 2. *Elements of a 1901 access cell.*

Figure 2 describes the elements of a 1901 access cell, which includes the head end station (HE), a number of repeating stations (RPs), and network termination stations (NTUs). The HE manages the cell and connects the whole access

IEEE Communications | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | qMags

In an access network where repeaters are likely to be used, end-to-end communication is defined as the entire communication path between edge stations in the Access Cell, including the repeaters, which are being used between them.

network to the backbone. RPs are stations that can also repeat from one station to another. NTUs are RPs that can also bridge between the access cell network and external customer networks such as IH networks or non-BPL technologies networks, such as IEEE 802.11 (Wi-Fi) and IEEE 802.3 (Ethernet).

Dotted lines in this figure represent the power line links between stations. Solid lines represent other network media.

BPL and non-BPL IH networks shown in this figure may be smart grid devices or networks served using the access cell as their backbone.

Specific Functions for Access Systems: Access Station Addressing

Access systems need to be able to address and control more than 1000 nodes. Every station inside the access network requires a unique address. The HEs are given short network identifications (SNIDs). SNID values support a maximum of 63 neighbor networks (63 SNID hexadecimal values, 0x01 to 0x3F, are the network addresses, and the value 0x00 is reserved to mark unassociated stations). Each SNID uniquely defines and logically separates one access cell from its neighbors. Each new station searches for a cell with which to associate and uses the SNID to differentiate between neighbor cells. A station that is already connected will use the SNID to search other neighbor cells if it can improve its connectivity by hopping to a neighbor cell with better performance. If the station cannot hear the HE directly, it will use the RP's repeating ability and associate using another station as a proxy toward the HE.

After the association process ends, the HE allocates a 12-bit address, the terminal entity identification (TEI), to the new station. The TEI and SNID combination will be the address to be used by the station for synchronizing with its neighborhood and transferring information, for bridging, and for repeating purposes.

TEI values support more than 4000 stations in a single access network. This large address range enables different topologies and scenarios for large-scale access installation with low numbers of backbone connection points, and for small-scale access installation with large numbers of backbone connection points. Due to the reality of typical installations, there is no practical limit in terms of addressing for access installation scenarios.

The TEI and SNID combination is what makes each station unique and yet allows it to communicate with its neighbors. Because this address is used for the point-to-point communication, it is also the basic identification tool for higher-layer algorithms such as channel estimation, channel accessing, bridging, routing, and management.

Using the TEI/SNID addressing instead of the unique medium access control (MAC) address of each station is more efficient in terms of address space. The combined TEI and SNID are 18 bits long, while an Ethernet MAC address is 48 bits long. Because each packet includes at least source and destination addresses, TEI and SNID addressing will require only 30 bits (source TEI, destination TEI, and SNID) compared to 48-bit MAC addressing, which requires 96 bits (source Ethernet MAC and destination Ethernet MAC). The result is a saving of 66 b/packet.

Neighbor network detection is another aspect of addressing. Access neighbor networks should fairly coexist, communicate, and synchronize by informing each other about their transmission times. In addition, a neighbor network could act as a failover option. The failover option increases the network robustness in case a neighbor network loses connectivity or changes topology to a point where the link does not meet the application's requirements and needs a new path. The first step toward coexistence is the detection of other networks. Using SNIDs, each station will be able to detect its neighborhood activity and sort each transmission to each cell.

1901 Access Cell: End-to-End Communication

In an access network where repeaters are likely to be used, end-to-end communication is defined as the entire communication path between edge stations in the access cell (e.g., the HE or NTU, which are located at the edges of the access cell and may have an external network port), including the repeaters being used between them.

In a typical access network, packets are repeated through stations that are members of the network. However, these stations are not usually the final destination of packets, nor are they usually the initiators of packets. Also, the edge stations in the access cell act as a bridge between external (BPL or non-BPL) network devices to the access network. For example, a PC connected to an IH station may send a packet to the Internet (*www* in Fig. 2). The NTU will bridge this packet to the power line, and the HE will bridge it back to the backbone network toward the Internet.

This concept of access bridged networks has an analogy to a switch with an ingress Ethernet port (e.g., the HE station) and an egress Ethernet port (e.g., the NTU station). From a bridging point of view, this analogy is accurate. Each edge station bridges an incoming packet toward the relevant edge station. The route over which the packet travels toward this edge station is managed by a separate network layer — the routing layer.

In order to allow quick transfer of the packets from one end to the other via repeaters, the access system frame format is built from two headers:

- An external header, which determines the end-to-end communication path and contains the edge devices' addresses
- An internal header created by every repeater in the path according to its path toward the edge device station

The external header handles the end-to-end bridging communication, while the internal header deals with point-to-point communication. The combination of both allows the end-to-end communication from the HE to the NTU using one or more RPs in the path.

Clock Synchronization

Several applications within the access network require accurate clock synchronization between neighbor stations. An example of such a requirement is PHY clock synchronization, which is essential in order to communicate between two stations using higher-order modulations. Another example is the time allocation management reference clock described below.

In established IEEE 1901 IH networks, the clock synchronization process is quite simple. The station that can hear the most other stations is elected and becomes the network manager. This manager synchronizes the IH network to its clock using beacon messages. If one or more stations in the network cannot hear the master, the master assigns one or more proxy masters to forward the beacon clock and information toward these stations.

The access network synchronization is also based on a centralized clock. All access cell stations synchronize on the clock of the HE. The algorithm for this synchronization is more complex than the IH and uses a multihop mechanism. Due to the long distances between the access network's master and the access stations, the beacon period of access networks should be much longer than the IH beacon single-hop synchronization period.

The multihop mechanism used by the HE periodically transmits a beacon message with the network time base (NTB). The NTB is a 32-bit clock maintained by the HE, which indicates the time when the beacon was sent. Another field used in this process is the beacon level (BL), which is always set to zero by the HE. Every station that hears this beacon message synchronizes its clock based on this master clock.

The interval between HE beacons transmission is a function of the maximum interval that allows the network to maintain an accurate clock between neighbor stations.

In order to keep the whole access cell synchronized to the HE clock, each station that hears the HE beacon synchronizes its clock with the HE clock. When the beacon is transmitted by the station, it includes the station's transmitting time based on the station's clock and its inherited NTB. The beacon also contains the BL, which is set to 1 in this case. Only stations that are not synchronized to the HE (BL = 0) synchronize on these beacons.

This process continues for every BL until all the stations in the network have shared and synchronized the same clock value between all stations in the access cell according to their BL.

Given the fact that beacon messages are transmitted using robust but inefficient modulation (in terms of potential line bandwidth usage), the number of BLs is minimized, and is equal or less than the actual number of hops that would be used in an optimal and efficient data transmission path.

Intracell Smart Repetition

The forwarding mechanism described in the IEEE 1901 document has a distributed nature. Each station maintains its own forwarding table and independently makes its forwarding decisions. However, because all stations in an access network need to communicate with their HE, a connection path from the HE toward all access stations has to be established.

To make it easy to learn the connection path, each station listens to beacon messages carrying the neighbors' connection level information. Each station that receives a beacon message gathers the connectivity level information from all its neighbors. After taking into consideration the connection level difference between it and each of its neighbors, and the connection level every neighbor has between it and the HE, the station chooses the optimal neighbor to use as a repeater toward the HE.

Because each station constantly optimizes its route toward and from the HE, the result is a dynamic and optimized repeating tree network structure. Every change in link degradation or link improvement may add, remove, or reroute a hop from the previous route. This supports dynamic physical topology changes such as that caused by closing a power circuit or opening a different power circuit (very common in mesh power line networks).

It is important to emphasize that the optimal forwarding tree is usually different than the BL tree described in the clock synchronization section. The beacon-based tree is a function of the simple ability of the stations to hear beacons. There are a relatively small number of BLs. The forwarding tree, however, is optimized based on finding the maximum bandwidth path between edge stations, which usually favors shorter distances and better SNR.

Figures 3, 4a, and 4b summarize the topic of BLs and repetition level. In these figures the connection points of the BPL stations are marked with two adjacent circles. This symbol emphasizes the fact that BPL stations are connected to an existing power line network.

Figure 3 demonstrates the physical topology of an access network. In this scenario the HE has two adjacent stations (RP 100 and RP 101) that have good connection levels (marked with solid blue curves). Stations RP 103 and RP 202 are examples of stations that have bad connections (marked with dotted red curves). These stations have links to the HE, which is good enough for synchronization but not sufficient for high-rate bandwidth usage. Good and bad connection levels between other stations in this cell are marked in the same fashion (Solid blue for good links and dotted red curves for bad links).

Figure 4a and 4b show two network trees built according to the connection levels in Fig. 3. The beacon tree in Fig. 4a is created by the beacon time synchronization mechanism, which relies only on the ability to connect and does not take into account performance — it simply requires a minimal ability of low modulation reception. The repetition tree in Fig. 4b is created by the optimal forwarding mechanism, which

Several applications within the access network require accurate clock synchronization between neighbor stations. An example of such a requirement is PHY clock synchronization, which is essential in order to communicate between two stations using higher-order modulations.

considers the connection level and the bandwidth between the stations in order to achieve optimal performance.

Each station in the trees is marked with its hierarchical level (starting from the HE, which is Level 0). Notice the level changes between beacon tree connections and repetition tree connections.

The difference between the Fig. 4a and 4b trees is that beacon trees require only connectivity between different levels, while repetition trees require finding the optimal path in terms of bandwidth and service from the HE toward each station and vice versa. For example, the connectivity between RP 103 and NTU 303 is enough for beacon passing (synchronization) but (in this example) not good enough for higher-bandwidth transmission. Better performance was achieved using RP 203 as a repeater in the middle.

Power Saving

Most access systems are outdoors and consist of more than 1000 nodes. If a smart meter consumed 1 W-h, the total would be a large load on the system. In order to save power and minimize noise emissions, the access *sleep mode* feature is introduced.

The main challenge of supporting sleep mode in access networks (compared to IH networks) is

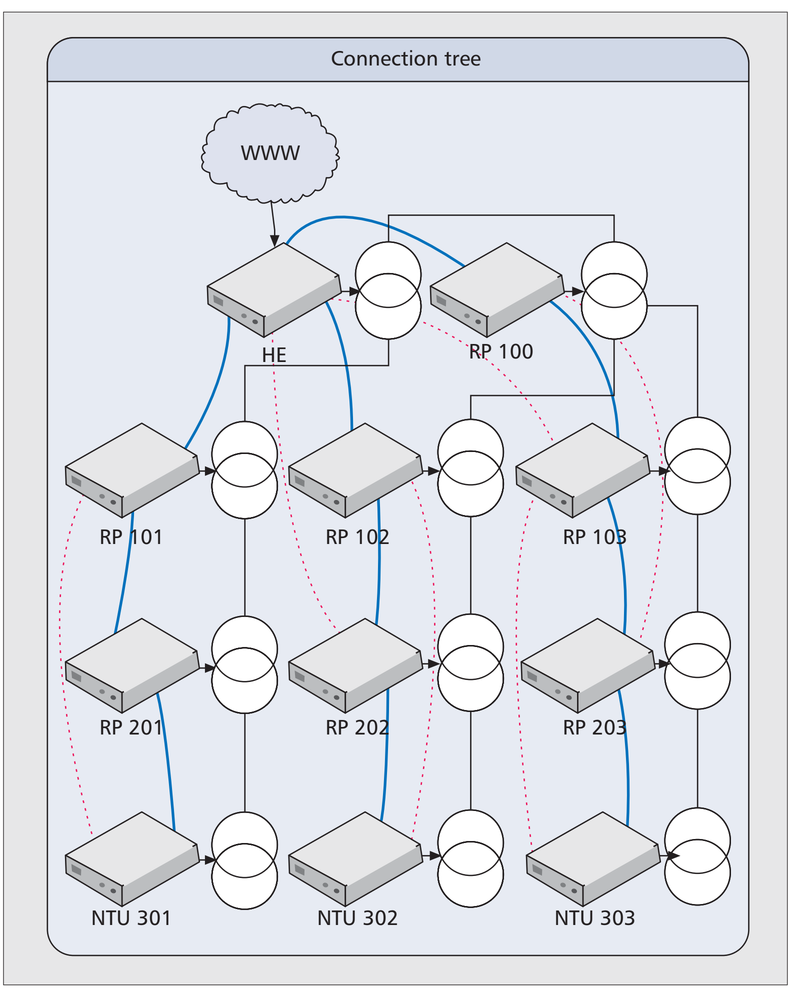


Figure 3. *1901 access cell physical installation topology example.*

the distance between the stations. The complexity and signal delay make it hard to schedule the sleeping periods and their duration.

In order to support sleep mode and at the same time ensure availability to the network infrastructure, the access system sleep mode works in a hierarchical manner. It starts from the HE and moves toward the edge stations through the RPs in a method similar to the beacon clock synchronization mechanism. Whenever a station intends to go to sleep, it publishes its sleep duration information using beacon messages. The stations located under it in the tree hierarchy will be able to synchronize their sleeping period to their parent's sleeping time. This will ensure that the stations below the parent will be awake in time to get updates and synchronization messages.

In addition, the system is able to define an *exclude list* of stations that are either essential to the proper work of the access cell or serve applications that are sensitive to latency. The HE will create the exclude list and publish it throughout the access cell (using beacon messages). Stations may also use the exclude list to keep their parents awake if they require full-time service.

Bandwidth Allocation

CSMA and TDMA

The default MAC protocol of access networks is carrier sense multiple access with collision avoidance (CSMA/CA) with prioritization. This MAC protocol is most suited for the broadband access topology that may contain hidden nodes. Moreover, CSMA/CA actually gives each station in the network the independent ability to compete with its neighbor stations for the right to use the medium without considering a master station, making it a good fit with the dynamic and multihop topology of the access network. Moreover, the dynamic nature of the access network and the unstable topology make it hard to set a local manager and define good master-slave relationship between the HE, the repeaters, and the edge stations throughout the network.

Nevertheless, CSMA/CA lacks a very important feature: it is not deterministic. This medium access algorithm is statistically based and does not guarantee bandwidth or minimum latency. Thus, it may not satisfy services and applications that require more deterministic QoS. Real-time applications such as telephony services (e.g., voice over IP) and videoconferencing are good examples of applications that require such determinism.

It is important that utility applications such as smart grid systems get access to the network in a deterministic way. TDMA can ensure bandwidth allocation as long as it is designed to function in the access cell multihop environment.

In order to support utility applications, a bandwidth allocation mechanism is required. The IEEE 1901 document introduced a concept of *TDMA over CSMA* in which the overall time resource is sliced into time regions. Each region defines the intervals for CSMA, for guaranteed TDMA allocations, and stay-out regions where TDMA allocations are granted to other stations

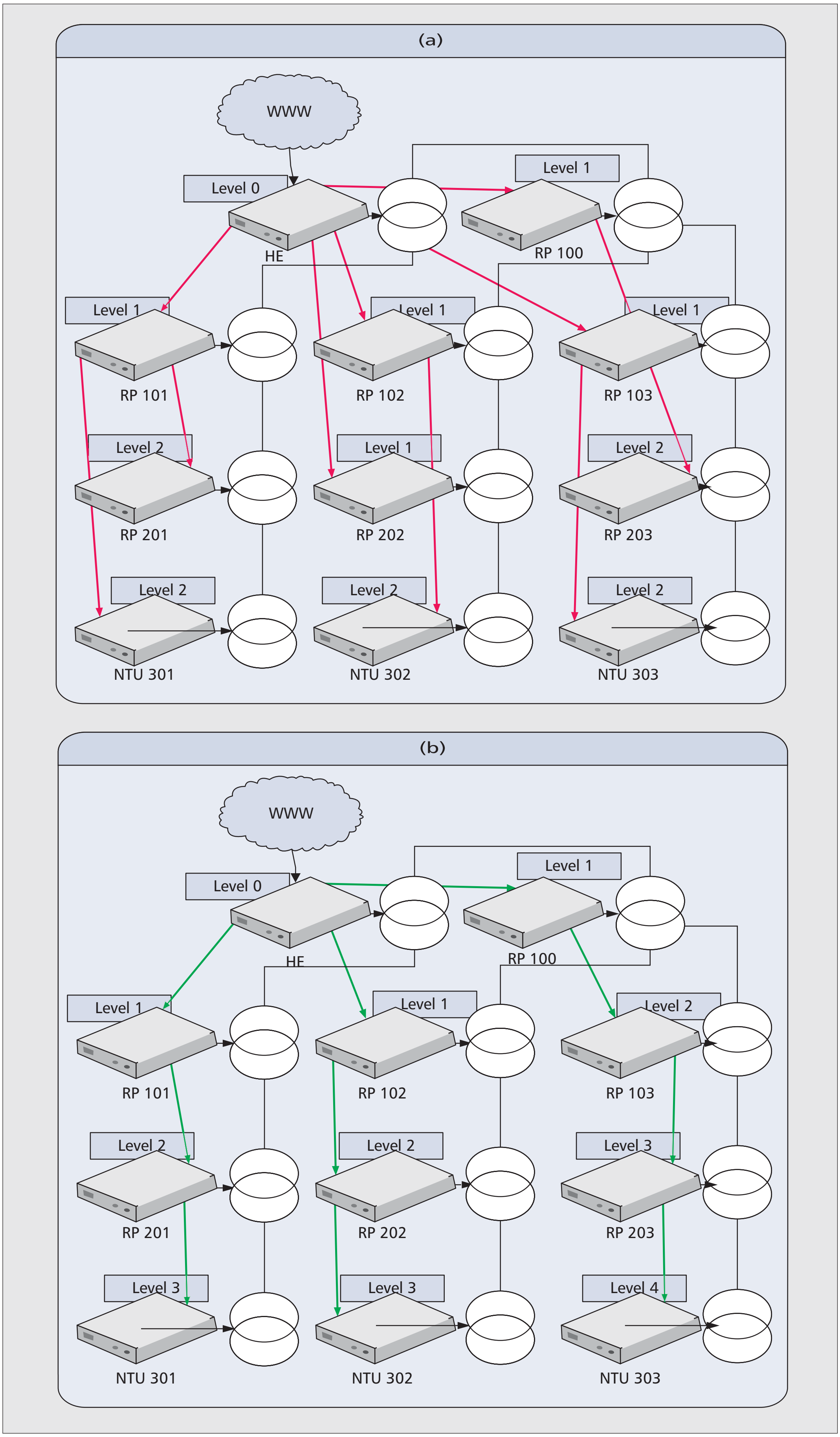


Figure 4. *a) Beacon tree example; b) repetition tree example.*

It is important that utility applications such as Smart Grid systems will get access to the network in a deterministic way. A TDMA can ensure bandwidth allocation as long as it is designed to function in the access cell multi-hop environment.

Communications | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | qMags

The IEEE 1901 document is a result of many years of work to include the presented access cluster functions that made it possible to apply it to large multi-hopping systems consisting of more than 1000 nodes.

in the station's neighborhood. Stations can contend for access in the CSMA region, and when allocated they get a persistent slot in the TDMA region. In the stay-out region the station will not transmit in order to avoid interfering with its neighbors' communication.

The HE is in charge of authorizing the opening of a TDMA channel within its cell. Each edge station that wants to use the TDMA service will request a channel based on the level of service required by the application. The HE will decide whether to authorize this request and start the allocation procedures according to the properties of the initiator of the TDMA channel and the available resources in the network.

The HE may use two methods to allocate the TDMA slots. It can schedule a slot per repetition hop on the way to and from a station as a central manager. Alternatively, it can initiate a remote and distributed procedure where the RPs allocate the slots, and each RP is responsible for its allocations according to its available time and bandwidth resources. The first method is referred to as *centralized TDMA* (since the HE manages everything itself). The second method is referred to as *distributed TDMA* (since each station manages its own allocations independently after getting authorization from the HE).

The HE may use these two methods in parallel in order to have control of specific TDMA channels and simultaneously have other TDMA channels automatically adapted. This is also a method to achieve a level of prioritization and different service levels between TDMA channels.

Until a TDMA channel is established, traffic is sent using the CSMA default access method.

Centralized TDMA

When the HE uses the centralized TDMA scheme, the HE becomes the sole manager of the channel between it and the relevant edge station. It is responsible for allocating a TDMA slot for each repeater in the route toward and from the edge station. The HE senda a message toward each relevant repeater and assigns the slot according to its knowledge of vacant bandwidth in the cell. Using this centralized method can ensure good synchronization between the time allocations of slots through the entire path. The HE can allocate the slots in order to optimize their timing to minimize latency for the whole channel. The HE can also take into consideration several channels being aggregated into one slot according to the topology of the channels.

Dynamic TDMA Polling

Another flavor of the centralized TDMA method is TDMA polling. In this method the master, usually the HE, will fix the slot schedule at the time of initiation. Because the channel condition and network resources change dynamically, the schedule table will be modified simultaneously after considering the polling results and channel conditions.

Master polling can be useful for managing numerous automatic meter readers (AMRs) by avoiding collisions due to CSMA. The signaling is designed to allocate time slots in a semi-persistent manner. Indeed, one allocation is signaled in beacons, and beacons are expected to have a periodicity around one or two seconds. It means that the same allocation is repeated during these one or two seconds.

This type of semi-persistent allocations is typically useful for data flows such as voice or video, with a rather constant bit rate over a long period. It is not applicable for AMR polling, were the typical flow profile is a single short data packet (a few hundred kilobits) every several minutes. In fact, if the TDMA schedule scheme defined in the previous sections is not carefully used for AMR devices, much bandwidth may be wasted.

Distributed TDMA

If the HE uses the distributed TDMA method, it controls only the first hop toward the destination edge station. Each repeater in the route is in charge of allocating a time slot according to its local time allocation map built from the information it gathers from its neighbor's beacons and from its already assigned slots. The time slots are being allocated in serial fashion per each hop until the channel is built from the HE toward the edge station and vice versa.

The main advantage of distributed TDMA is that due to its local management nature, each station can set the time slot according to its local constraints and is able to perform fast recovery in cases of allocation collision between neighbor stations, hidden nodes, and neighbor cells. The distributed method is also much easier to manage in cases where the route path has changed, causing the need for fast teardown and recreation of whole or parts of the TDMA channel path.

Conclusions

This article has covered the main challenges in the broadband power line access network environment, and has given a glimpse of concepts introduced by the IEEE P1901 Working Group. A short description was given for access's most unique mechanisms such as the network concept of an access cell, smart repetition, data forwarding, QoS requirements, and power saving sleep mode.

The access network related mechanisms share some common functions with the IH network such as the physical layer and basic channel access. The main difference between the access and IH environments is the large number and variety of dynamic network topologies the access network needs to support.

The most challenging aspect of access oriented systems is the fact that they are designed for a multihop environment, while a typical IH network only has one or two hops.

The IEEE 1901 document is the result of many years of work to include the presented access cluster functions that made it possible to apply it to large multihopping systems consisting of more than 1000 nodes. The QoS mechanism based on master-slave topology is ideal for optimizing the multihop access environment.

ACKNOWLEDGMENTS

The authors wish to express their gratitude to Jim Allen (Arkados, Inc.) and the reviewers for their valuable contribution and feedback.

REFERENCES

[1] S. Galli and O. Logvinov, "Recent Developments in the Standardization of Power Line Communications within the IEEE," *IEEE Commun. Mag.*, vol. 46, no. 7, July 2008, pp. 64–71.
[2] IEEE P1901, "Draft Standard for Broadband over Power Line Networks: Medium Access Control and Physical Layer Specifications"; http://grouper.ieee.org/groups/1901/index.html
[3] IEEE Smart Grid Portal; http://smartgrid.ieee.org/

BIOGRAPHIES

SHMUEL GOLDFISHER (shmulikg@mainnet-plc.com) received his B.Sc. degree in computer science and mathematics from Bar-Ilan University of Ramat-Gan, Israel. He started his career as a software engineer in 3Com in the ATM networking field. In 2000 he joined MainNet Communication Ltd. and has been involved in design and development of the MainNet main product line in the Access over Broadband Power Line Communication segment. Since 2005 he is a software manager involved in MainNet's system architecture. As a member of the HomePlug Broadband PowerLine Communication Specification Working Group, he was one of the contributors of the HomePlug access specification. Since 2007 he has been an active participant in and contributor to the IEEE P1901 Working Group. His interest is in access communication in general and power line communication specifically. He holds several patents in power line communication and smart grid technologies.

SHINJI TANABE [SM] (Tanabe.Shinji@dh.MitsubishiElectric.co.jp) is a chief engineer in the Advanced Technology R&D Center of Mitsubishi Electric Corporation. He is a member of IEEE SA. He received his B.Sc. in physics from Tohoku University, Sendai, Japan, in 1980 and his Ph.D. in electromagnetic field analysis and high-speed communications from Tohoku University in 2003. He joined Mitsubishi Electric in 1980 and developed magnetic heads for hard disk drives and VCRs, and electromagnetic numerical simulation methods for EMC analysis. He was a visiting researcher at Tohoku University from 1982 to 1983 and a visiting researcher at the University of Minnesota from 1989 to 1990. He has been involved with IEC and IEEE standardization activities since 1995 as a member of IEC TC106, and IEEE P1528, P1675, P1775, and P1901. He published 12 papers and eight conference papers in the field of magnetic recording, magnetic materials, electromagnetic field simulations, EMC, and high-speed communications, and one book about electromagnetics.

IEEE Communications Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page qMags

The Creation of a Ubiquitous Consumer Wireless World through Strategic ITU-T Standardization

Máirtín O'Droma and Ivan Ganchev, University of Limerick

This article is based on a paper presented in the ITU-T Kaleidoscope Conference, Geneva, May 12–13, 2008.

ABSTRACT

The theme of this article is a proposed global wireless communications environment — ubiquitous consumer wireless world — enabled through the strategic creation of International Telecommunication Union standards. Specifics include new consumer identity module cards, new network-independent and location-independent personal IPv6 addresses, new protocol and interface infrastructures underpinning trusted third-party authentication, authorization, and accounting services, and new consumer-oriented incoming call connection services.

INTRODUCTION

The ubiquitous consumer wireless world (UCWW) is a proposed global environment that brings a different approach to the wireless communications business [1]. Through it, users become consumers instead of subscribers. Consumers will have full mobility among participating access networks, switching between them, for instance, to access more suitable price/performance offerings for specific services.

Besides a range of new benefits for the user, UCWW has the potential to stimulate the creation of a number of new interesting business opportunities and to create a more liberal, more open, and fairer wireless marketplace for existing and new access network providers (ANPs). The primary ANP business success indicator will shift from subscriber numbers to the volume of consumer transactions. This will increase the range of competitive price/performance and price/quality of service (QoS) offerings, specialist and niche service offerings, and so forth, all of which will drive forward innovation in the wireless communications services market.

In this article we seek to present the whole picture of this new consumer-centric wireless communications environment from a standardization perspective and to make the case, in a tutorial style, for undertaking a global standardization program in support of it. In this we have constrained ourselves to describing the standardization needs of the essential new core infrastructural elements, identifying exactly in what respect they should be standardized and in harmony with what guiding principles and objectives (final two sections). For this we first consider the standardization context and related research work (next section), and also present the relevant defining aspects of the subscriber-based and consumer-centric techno-business models (following section), together with indications of transition trends toward the consumer-centric model.

STANDARDIZATION CONTEXT

Three points will help to clarify the standardization issue. First, while requiring some distinct technological infrastructural modifications [1], UCWW will nonetheless benefit fully from almost all the existing global technological development and standardization efforts in wireless communications, such as Next Generation Mobile Networks (NGMN) Alliance proposals, Third Generation Partnership Project's (3GPP) Long-Term Evolution and System Architecture Evolution (LTE-SAE), and the International Telecommunication Union — Telecommunication Standardization Sector's (ITU-T's) ongoing work on next-generation networks (NGN).

Second, the underlying techno-business philosophy in wireless communications today is founded in the broad sense on the commercial personality of the user being that of a subscriber. Standards naturally reflect this. It may be called the subscriber-based business model (SBM). Its characteristic attribute, as in the legacy fixed network, is that of the user having a *fixed* point of attachment to the network and having a (long-term) contract with an ANP — *home ANP* — mapped onto a phone number (or a unique globally significant identity). The new consumer-centric business model (CBM) — founded on the commercial personality of the user being that of a consumer — is an alternative to this. Global standards need augmenting to reflect this model, in fact to create the necessary conditions for its realization.

Lastly, the ITU-T recently has made provision for *third-party authentication and authorization* services [2]. This is a good platform on which to build the full UCWW third-party authentication, authorization, and accounting

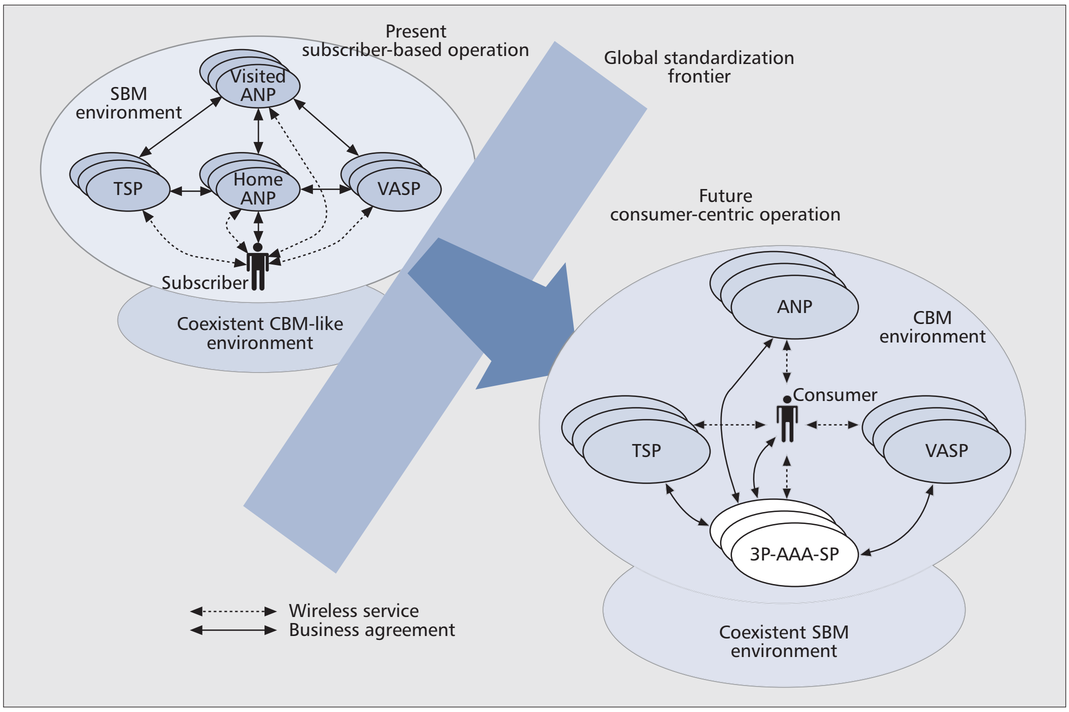


Figure 1. *Subscriber-based (SBM) and consumer-centric (CBM) models.*

That there are techno-business-model transition trends in wireless networking towards the CBM already present today is not surprising. Examples include: cellular operators' and ticket-based WiFi hotspots, per-call WiFi hotspot payment products such as metakall, and opening carriers' networks to third parties

(3P-AAA) service, which also incorporates charging and billing (C&B) aspects (addressed later in the article).

Published ideas nearest to UCWW, other than those reviewed in [1], include the Wireless World Research Forum's (WWRF's) work of placing the user at the center of the business model [3]. Also, among the business models considered for cellular operators in [4], the one envisaging the operator acting only as an open bit pipe provider is close to the CBM idea. Another CBM-like model is the open heterogeneous mobile network architecture in [5] realizable on Telecommunications and Internet Services and Protocols for Advanced Networks (TISPAN)-NGN and supported by third-party organizations. Heterogeneous there, and in this article, refers to users gaining wireless access over various disparate access networks, with handovers between such networks being possible. Within the m-commerce services domain, Groupe Spéciale Mobile Association's (GSMA's) recent *mobile-money* and *pay-buy-mobile* [6] initiatives manifest functionality and security attributes akin to certain requirements in UCWW. While these schemes are clearly subscriber-based, the role of the new *trusted services manager* entity echoes aspects of the functional role of 3P-AAA service providers (3P-AAA-SPs) in UCWW.

Wireless Environment in Transition

That there are techno-business-model transition trends in wireless networking toward the CBM already present today is not surprising. Examples include cellular operators' and ticket-based WiFi hotspots, per-call WiFi hotspot payment products such as *metakall* (www.metakall.com), opening carriers' networks to third parties, mobile money, and credit systems for multihop wireless networking. Only with global standardization support, however, will these evolutionary trends toward an open, globally pervasive CBM environment be facilitated.

Figure 1 portrays schematic representations of both techno-business models (SBM and CBM), illustrating the main wireless service paths and business agreement relationships. Transition to the UCWW, where the new CBM environment may coexist side by side with the SBM environment, is shown as a passage through a global standardization frontier.

In the SBM the subscriber primarily gets wireless services through his/her *home* (cellular) ANP or from visited ANPs if roaming agreements are in place. WiFi hotspots, femtocells, and the like, when they offer services using the AAA infrastructure of a user's home ANP, come under the visited ANP umbrella. Teleservice providers (TSPs) and value-added service providers (VASPs)[1] can also offer their own services through access networks under respective bilateral business agreements. If, however, other means are used to obtain payment for services, decoupled from the home ANPs (e.g., direct service purchase), these could be described as *CBM-like* and their presence be part of the trend toward a CBM environment coexistent with the SBM (bottom left of Fig. 1).

Certain downsides are becoming ever more apparent in the SBM, mainly linked to the *lock-in* constraint of the (long-term) subscriber contract with a home ANP. Examples include roaming charges (which are often perceived as not cost-based), domination of the ANP marketplace by a few large providers, poor market openness for new or niche ANPs due to pro-

[1] *VASPs provide services other than services of ANPs and TSPs (e.g., content provision or information services) for which additional charges are incurred.*

In CBM, all users, whether they be roaming visitors or not, are perceived as local to the access network from which they seek services. The technical basis for the roaming business thus no longer exists and the distinction between home and visited ANPs disappears.

hibitive startup costs, and limited *number portability and mobility* among ANPs. In regard to the latter, subscribers who desire to change ANP, rather than porting their number — a slow formal process where it is enabled — tend to the easier solution of buying a new (U)SIM card with another phone number in the other network. *Spinning* is another approach whereby subscribers using multiple (U)SIM card phones can choose to operate on any ANP at any time. However, this is still quite far removed from the full number portability and mobility within the CBM, where consumers will be allowed always to use the best ANP for each particular service instance. As regulators will readily attest, the SBM downsides are not easily addressed while staying within the SBM environment. The legislative efforts in the European Union to remove roaming borders across its Member States is an example.

Transition to UCWW and enabling the growth of CBM opens opportunities to address these issues. The two principles underpinning this new CBM environment are [1]:

- The decoupling and separation of the administration and management of users' AAA activity from the supply of a wireless access (transport) service and its devolution to new non-ANP trusted 3P-AAA-SPs entities
- The full consumer ownership and portability of their globally significant address [1]

These principles are inextricably linked.

The kind of standardization needed to open the possibility for these new business entities to create their 3P-AAA infrastructures is set out in the next section. The ANP-independence of new 3P-AAA-SPs is important for fairness in the ANP marketplace (e.g., in relation to access to consumer wireless access usage information) and as such may require regulation.

This 3P-AAA infrastructure innovation engenders a strong distinction between CBM and SBM. Users will be able to access wireless services anytime-anywhere-anyhow through any CBM-compliant access network on a transaction-by-transaction basis. This is a significant step up in the level of user-driven *always best connected and best served* (ABC&S) empowerment. In CBM all users, whether they be roaming visitors or not, are perceived as local to the access network from which they seek services. The technical basis for the roaming business thus no longer exists, and the distinction between *home* and *visited* ANPs disappears.

In regard to realizing the second principle, the standardization goal is the creation of a personal, globally significant, network- and device-independent address capable of full number portability. We recommend a new separate class of *personal IPv6 addresses* satisfying these attributes, standardizing this and its container — a smart universal consumer identity module (CIM) card to replace the present (U)SIM card. With CIM, the user may obtain and securely pay for services from any CBM-enabled access network or teleservice provider.

Security around this CIM card will require careful standardization. For instance, it needs to be securely locked to enable the user to be identified and authenticated for execution of 3P-AAA procedures, provision and purchase of communications services, teleservices, and so on. The ITU-T's X.509 digital certificate, with its potential for large-scale deployment and capacity for binding several IP addresses to one certificate, may be used. Typically, users' digital certificates will be embedded into the CIM cards, with secure ways to add new personal IP addresses and so on. Certificates issuers could be 3P-AAA service providers, among others. Responsibility for secure access and management operations could be jointly shared between users and the 3P-AAA-SPs, although other forms are possible.

With the existence of 3P-AAA services and the full portability and mobility of CIMs, an environment is realized whereby the development and implementation of user-driven ABC&S services through new terminal and/or TSP end-system applications become feasible, and this with various degrees of seamlessness and especially network transparency. An example is the hot access-network change (HAC) between heterogeneous networks controlled and managed by mobile terminals (MTs) described below (final section). This is an altogether different paradigm from network-driven and -managed heterogeneous networking, for example, as envisaged for handover mechanisms in next-generation systems working within 3GPP's Evolved Packet Core (EPC). The UCWW environment also includes the potential for user-driven integrated heterogeneous networking, whereby end systems realize different parts of a single service in a coordinated way over different wireless access networks, and transparently to those networks. An example of this would be a user-TSP negotiation over a cellular network for a download service, with the downloading actually being done, independently and transparently, over a broadcast or hotspot (WiFi) network.

Evolution of CBM will be gradual, take time to mature, and be perceived as quite supplemental to the SBM environment. As Fig. 1 indicates, both it and SBM will coexist. The SBM's established reliability, QoS, global depth of penetration, and universal user acceptance will be effective in maintaining its subscriber base. However, dual-mode user terminals (e.g., with both (U)SIM and CIM cards for SBM and CBM operation) are likely. The same duality may apply also to ANPs, especially during the transition period from SBM to CBM. In general, in fact, within UCWW the thrust is for users to have unfettered access to access networks and vice versa.

The main SBM and CBM attributes, where there are distinctions, are compared in Table 1.

3P-AAA Infrastructure and Signaling Protocols

The 3P-AAA interface infrastructure's and signaling protocols' standardization program to enable global provision of 3P-AAA services is outlined here. Standards proposed for agreement should respect the need for these services to be scalable, hierarchical, and cognizant. Such

Item	Attribute	CBM	SBM
User	Commercial essence	Consumer	Subscriber
	Presence in a network	Always as local	As a local or visitor (roamer)
	Network choice, mobility, and portability	Full, unfettered	Limited
	Identification	Through (single) CIM card	Through (multiple) U/SIM card(s)
	Connectivity	Anytime; anywhere; any network	Anytime; anywhere; home network (or visitor network with roaming agreement)
Address/ Identity	Ownership	User	Joint: user and home-ANP
	Type	Network-independent and location-independent	Network-dependent and/or location-dependent
	Portability	Full, immediate, and inherent	Limited (slow formal process)
AAA	Service provider	Extra-network third-party entity (3P-AAA-SP)	Home ANP
	AAA client	On terminal	In network
	Protocol (Diameter)	New 3P-AAA application	Unchanged
	Signaling	Through new open application-layer interfaces Increased (3P-AAA-SP ↔ user) Decreased (network ↔ network)	Unchanged
ICC	Service provider	Extra-network autonomous entity	Home-network
	Service management	Highly customized Greater user management of ICC policies	Constrained user management
	Signaling	Through new open application-layer interfaces Increased (ICC-SP ↔ user) Decreased (network ↔ network)	Unchanged
ANP	ANP market	Profitability primarily linked to the number of consumer transactions Openness of entry More competition	Profitability primarily linked to the number of subscribers Difficult entrance Less competition
	Home and visitor network distinction	No	Yes
	Inter-ANP business agreements necessary for roaming	No	Yes (Multiple bilateral home-ANP ↔ visited-ANP agreements)
	Technical foundation for the application of roaming charges	No longer exists	Unchanged
Generic items	AAA, ICC, and general teleservice decoupling from wireless access service	Yes	Unchanged
	ABC&S (main driver)	User	Network
	Hot access network change (HAC) between autonomous heterogeneous networks	Yes Network-transparent, user-controlled	Yes User-transparent, network-controlled
	Advanced intelligent access & service control evolution	In terminals & TSP entities primarily	In network core primarily
	Wireless environment	Consumer-centric	Network-centric

Table 1. *CBM and SBM attribute values compared.*

attributes may be seen in global 3P-AAA service solutions where service providers deploy their AAA servers regionally, in multiple hierarchical layers, reflecting the dimensions and characteristics of their customer bases. Service to customers may also be improved (e.g., better response times achieved) by porting a copy of a customer's data to the AAA server in the region closest to where a customer is at the moment.

We envisage the creation of an infrastructure of three new application-layer interfaces (and suitable signaling protocols) as indicated in Fig. 2:

- User ↔ ANP and user ↔ TSP
- User ↔ 3P-AAA-SP
- ANP ↔ 3P-AAA-SP and TSP ↔ 3P-AAA-SP

Corresponding to foreseeable major market sectors, three specialized 3P-AAA-SP classes are posited [7]: class A for ANPs, class B for TSPs and VASPs, and class C for consumers. It is not intended with this division to exclude other forms or even that a single 3P-AAA-SP would cater for all three markets. There will be similarities but also distinctions in the services provided within each class. For ANPs, for instance, services will include accounts and related AAA policies, C&B policies, pricing and rating functions, charging detail records generation, account balances, and the like. For consumers (besides accounts), these may include various types of credit top-up services, billing system configuration functionality, functionalities to enable customer retention discount and promotional schemes, user-definable account format and layout, and so on. The advantage of such class differentiation is that each 3P-AAA-SP can focus on a subset of the overall required functions so that these can be made considerably more sophisticated than they might otherwise be. Also, with establishing such classes any consequent effects on the new interface protocols may be more easily handled.

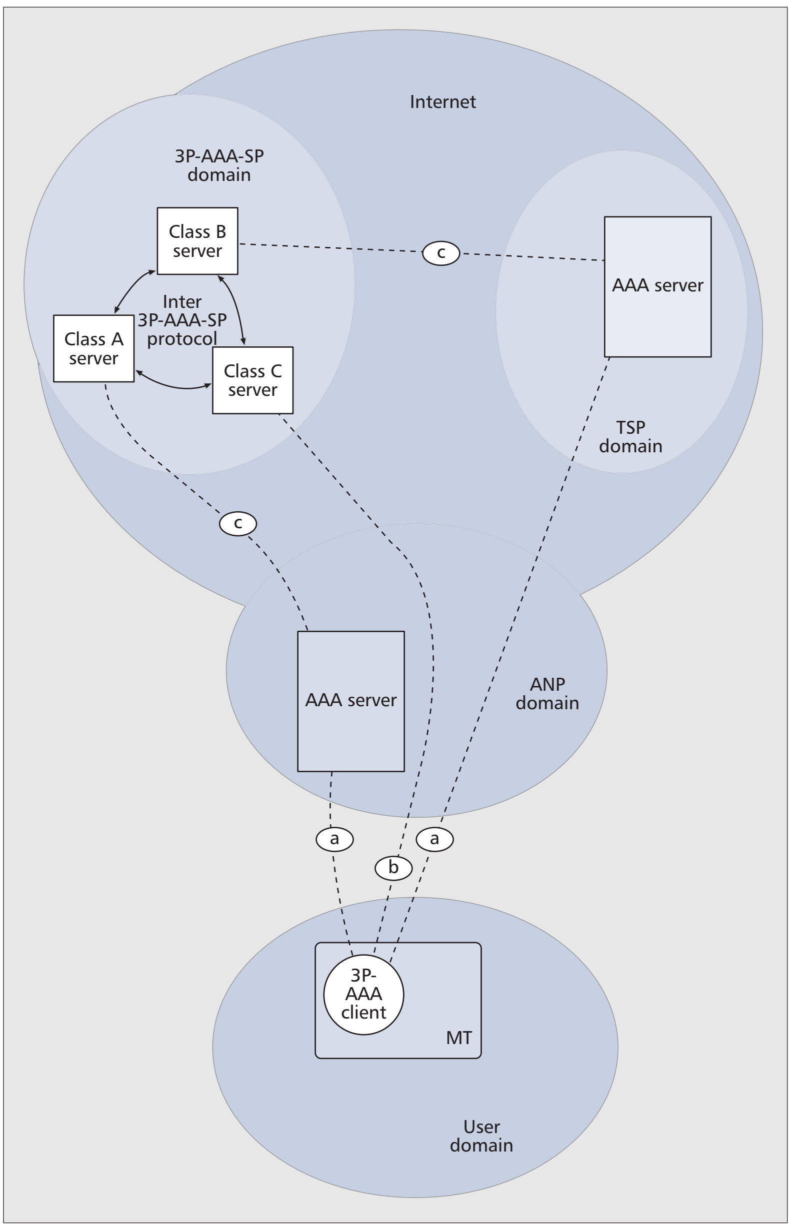


Figure 2. *3P-AAA functional model schematic (with three new application-layer interfaces indicated).*

On these new interfaces (Fig. 2), the existing Internet Engineering Task Force (IETF) Diameter protocol [8] has suitable attributes for carrying 3P-AAA signaling. Being conceived for the SBM environment, however, adjustments for 3P-AAA will be required. Exploiting the Diameter design, whereby it accepts new autonomous applications that run on its core, is a potential solution and has the advantage of not constraining the Diameter core from evolving independently.

The multi-application processor CIM card inserted in the MT will have one or more 3P-AAA application clients installed. This devolution of application programs and processing power to the CIM is a significant change from (U)SIM operations. Simultaneous accounts with different 3P-AAA-SPs will be possible. Cards, for instance, could come with generic 3P-AAA clients pre-installed and several free spaces on the card chip leasable to specific 3P-AAA-SPs for installing (and upgrading) their own specific 3P-AAA client, designated for that consumer.

Through standardized protocols the 3P-AAA client on the MT interacts with AAA servers of the ANPs and TSPs for mutual authentication and the exchange of security credentials. A part of these standardized protocols will be the signaling to establish authentication securely prior to any service purchase. An authentication scheme based on ITU-T's Recommendation X.509 can support strong secure mutual authentication and trusted relationship establishment with a minimum number of protocol exchanges. Its three-way exchange option has the attraction of not requiring the communication parties to have synchronized clocks.

Illustrated in Fig. 3 is an example related to the difference in approaches in ongoing standardization work that UCWW thinking, based on this 3P-AAA and CIM infrastructure, would bring. The figure represents the ITU-T's graphical illustration of an NGN authentication architecture for interworking among heterogeneous wireless networks operating within the SBM [9]. Its objective is to enable a subscriber of one (home) network (e.g., a 3G cellular network) to gain wireless access services as a roamer from another heterogeneous network (e.g., a WLAN

Communications Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

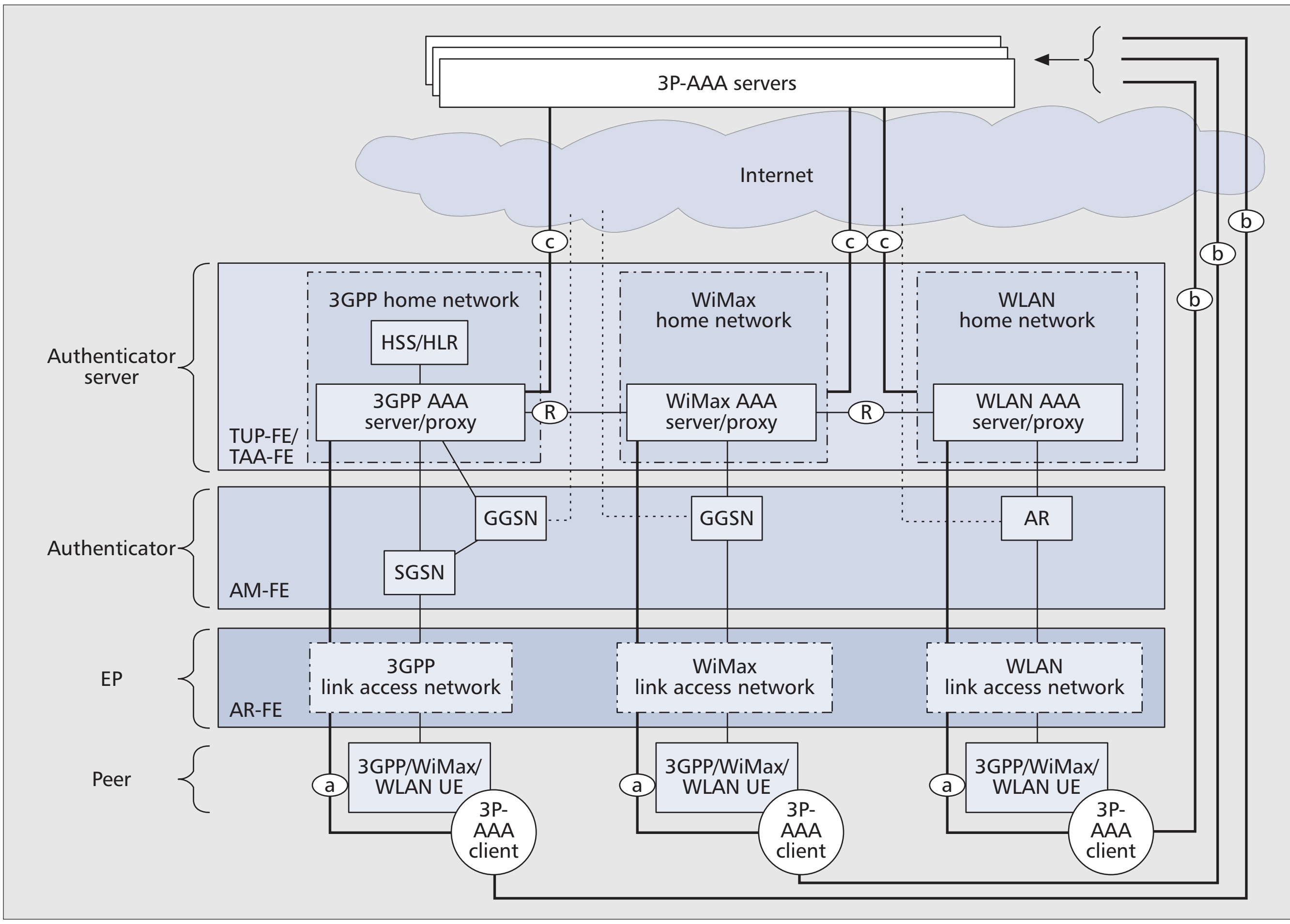


Figure 3. *ITU-T's SBM authentication architecture for interworking in NGN [9] and proposed modification (new entities in white) for UCWW.*

or WiMAX network) with payment executed through the home network. The approach also may form a basis for subscriber-transparent network-driven handover while roaming among collaborating heterogeneous networks. It uses a four-layer architecture, with the user equipment (UE) at the bottom. The network access control function entities are the authenticator (access management functional entity, AM-FE) acting as an AAA client and the home-ANP's AAA server (transport user profile functional entity/transport authentication and authorization functional entity, TUP-FE/TAA-FE) positioned at the third and fourth layers, respectively. The second layer (access relay functional entity, AR-FE, in the access network) acts as an enforcement point, filtering packets and allowing through only packets exchanged for initial authentication and subsequently authenticated data packets. The key communication to allow roaming on a heterogeneous network is on the *roaming interface* (marked *R*; our notation) between the AAA servers in the different network domains.

Our modifications of this architecture for a CBM environment are overlaid on this ITU-T illustration. In CBM the *home network* attribute no longer applies to access networks. Also, the roaming interfaces *R* are not required. Instead we have the presence of 3P-AAA server entities with which the network-specific AAA servers communicate over standardized 3P-AAA interfaces of type *c*. Similar type interfaces are needed between 3P-AAA servers and the service-specific AAA servers (service authentication and authorization functional entity, SAA-FE). In addition, the AAA client is shifted from the network to the terminal, and its communication with the ANP's AAA server and 3P-AAA server is performed over the proposed standardized 3P-AAA interfaces of type *a* and *b*, respectively. The balance of the heterogeneous networking decision making power is now much more in the hands of mobile users and away from networks.

The signaling between 3P-AAA servers and local AAA servers of ANPs (and TSPs) could lead to significant network traffic. This would be the case if the C&B implementation includes a requirement for continuous processing of the frequently clocked records coming from individual charging functional entities in order to calculate the correct service charge. This is a typical implementation for C&B of phone calls, and other non-flat-rate services, in the SBM environment today. To address this, we propose the concept of a *C&B agent*. Downloaded in advance from 3P-AAA-SP and deployed in the metering domain of an ANP (or a TSP), this agent would perform all associated C&B functions there. For this, the agent would come with a budget for the

Communications | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | qMags

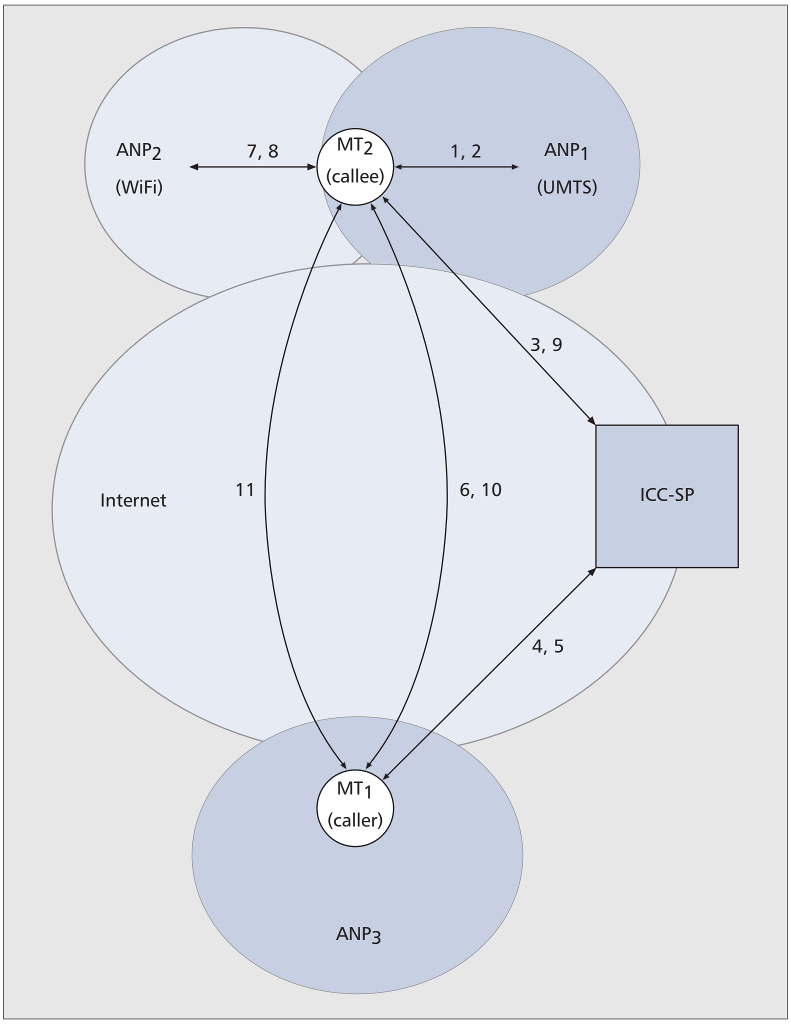


Figure 4. *ICC session set-up (actions 1 to 6) and consumer-driven ICC session switching (actions 7 to 11).*

service to be supplied to the consumer and any other relevant consumer account details. This agent would have the functionality to manage the budget, for example, expending it in response to the service provider's (ANP or TSP) metering triggers, sending budget replenish requests to the 3P-AAA server when a budget depletion threshold is crossed, and at the end of the service session informing the 3P-AAA server of the total charge [7].

CONSUMER-ORIENTED INCOMING CALL CONNECTION SERVICE

As a consumer in CBM does not have a fixed point of attachment (which a subscriber has in SBM), incoming call requests would be directed to autonomous dedicated ICC service providers (ICC-SP). Standardization would facilitate an ICC service operation based on a contact address (CA) scheme — a globally routable, temporary, forwarding IP address. It can be associated with an NAI-like contact address identifier (CAI) — a public identifier, in the form of `callee@ICC-SP_domain`. Consumers may use the services of more than one ICC-SP and thus have more than one CAI. An example of the consumer-oriented ICC service operation is shown in Fig. 4. It illustrates an initial ICC session setup and the subsequent consumer-driven live call session switching, and is described in more detail below.

To enable ICC, a callee using a mobile terminal, MT_2, who initially is without a globally routable active presence through which s/he can receive incoming calls, will first associate with an access network — in this example with ANP_1's Universal Mobile Ttelecommunications System (UMTS) network (*action 1*) — and seek such presence by acquiring a contact address, CA_1, from ANP_1 (*action 2*) and sending it to his/her ICC-SP (*action 3*). There it is bound to the callee's CAI and is available to callers in accordance with the callee's policies, e.g., that CA_1 be given out only to certain categories of callers. The callee now has the desired globally routable active presence for receiving incoming calls.

For an ICC, a caller, here MT_1, in using a callee's CAI, will automatically send a *CA request* for the callee's present CA to the relevant ICC-SP (*action 4*) and will receive it (*action 5*). The ICC session setup can now proceed between MT_1 and MT_2 using the callee's CA_1 (*action 6*). Effectively within ANP_1's UMTS network, translation between CA_1 and the consumer's personal address is performed by a network address translation (NAT) mechanism.

Live call session switching could occur when the callee enters the area of another network — here a WiFi network — provided by ANP_2 which offers the same service at an attractively lower cost. To benefit immediately from this, the callee's terminal needs to arrange a live-call session switched to ANP_2's network. It does this through a consumer-driven, network-transparent, hot access network change (HAC) as follows. The callee first associates with the new ANP_2 for ICC service support and obtains a new contact address, CA_2, (*actions 7 and 8*). MT_2 also notifies its ICC-SP (*action 9*) and perhaps sets a specific policy for CA_2. Then to execute HAC, the callee informs the caller about the new CA_2, directing the caller to change the session IP address to CA_2 (*action 10*). A technology solution with potential to support this form of end-to-end seamless HAC could be the IETF's mobile Stream Control Transmission Protocol (mSCTP)[2] [10]. By means of this, the active ICC session will be switched for delivery over the ANP_2's network (*action 11*). This call switching action is network-transparent; for example, to ANP_1 it would be perceived simply as a normal termination of the current ICC session. HAC could be further supported and speeded up by installing a smart IEEE 802.21 Media Independent Handover client on the callee's terminal. This acts to provide layer 2 (L2) support for terminal mobility across heterogeneous networks, such as fast discovery of neighboring access networks, quick appraisal of new L2 network points of attachment, setup of multiple L2 links for make-before-break handling, quick teardown of unused L2 links, and QoS parameter mapping among different access technologies.

[2] *mSCTP is primarily targeted at multistreaming and multihoming transport, where it facilitates end-to-end controlled mobility by allowing endpoints to dynamically add, change, and delete the primary IP address in an active association using address configuration (ASCONF) chunks. The design goal is the provision of soft handover (at the transport layer and without help of routers or additional network agents) for terminals moving to a different access network during an active teleservice session.*

In summary, to enable global operation of this reinvented ICC service the signaling protocols for communication with the ICC-SP, the CA assignment and operation, as well as the address resolution schemes will need standardization.

CONCLUSIONS

Standardization in support of 3P-AAA, CIM, and consumer-oriented ICC will make global transition to UCWW feasible. With this, or as part of this process, an implementation of a pilot global system would provide direct experience of CBM system design, validation, and deployment. This may stimulate stakeholders such as the new 3P-AAA-SP and ICC-SP business entities, terminal manufacturers, and new ANPs, attracted by the new consumer business potential, to invest in its realization and evolution. Considering the benefits and attributes already mentioned, in many ways it may be claimed that the UCWW harmonizes well with the ITU's NGN objectives [11].

REFERENCES

[1] M. O'Droma and I. Ganchev, "Toward a Ubiquitous Consumer Wireless World," *IEEE Wireless Commun.*, vol. 14, no. 1, Feb. 2007. pp. 52–63.

[2] ITU-T Rec. Y.2702, "Authentication and Authorization Requirements for NGN Release 1," 2008.

[3] WWRF WG2, "Service Architectures for the Wireless World — White Paper on Business Models for the Wireless World v1.0," 2003.

[4] C. Andersson *et al.*, *Mobile Media and Applications — from Concept to Cash*, Wiley, 2006.

[5] Y. Murata *et al.*, "The Architecture and a Business Model for the Open Heterogeneous Mobile Network," *IEEE Commun. Mag.*, vol. 47, no. 5, 2009. pp. 95–101.

[6] Telstra, "GSMA Pay-Buy-Mobile Business Opportunity Analysis," White Paper, Nov. 2007; http://www.telstraenterprise.com/SiteCollectionDocuments/Whitepapers/GSMA%20PBM%20White%20Paper.pdf

[7] J. Jakab, "New Charging and Billing Models and Mechanisms for the UCWW," TRC internal rep., Univ. of Limerick, Ireland, 2009.

[8] P. Calhoun *et al.*, "Diameter Base Protocol," IETF RFC 3588, 2003.

[9] ITU-T Draft Rec. Q.3202.1 (Q.nacf.auth1), "Authentication Protocols Based on EAP-AKA for Interworking among 3GPP, WiMax, and WLAN in NGN," 2008.

[10] S. J. Koh and Q. Xie, "Mobile SCTP (mSCTP) for IP Handover Support," IETF draft, Oct. 2005.

[11] ITU-T Rec. Y.2001, "General Overview of NGN," 2004.

BIOGRAPHIES

MAIRTIN O'DROMA [SM'91] (Mairtin.ODroma@ul.ie) received his B.E. and Ph.D. degrees from the National University of Ireland in 1973 and 1978, respectively. He is a senior academic and director of the Telecommunications Research Centre at the University of Limerick, Ireland. He is an IEEE Subject Matter Expert, an ITU-T Invited Expert, a Fellow of the IET, and Chair of Ireland's Royal Irish Academy's Communications and Radio Science Committee. His previous activities include founding partner and steering committee member, TARGET — Top Amplifier Research Groups in a European Team — the EU FP6 Network of Excellence IST-507893, 2004–2008, and section head of the RF power linearization and amplifier modeling research strand; founding partner and steering committee member, ANWIRE — Academic Network for Wireless Internet Research in Europe — the EU FP5 Thematic Network of Excellence IST-38835, 2002–2004; Ireland's delegate to two European Science Foundation Cooperation in the field of Science and Technology Research Actions: "Modelling and Simulation Tools for Research in Emerging Multi-Service Telecommunications" and "Traffic and QoS Management in Wireless Multimedia Networks," (COST 285 and 290). Currently he is delegate (Ireland) to the COST Action IC0906, "Wireless Networking for Moving Objects" (WiNeMO). His research interests include wireless NGN infrastructural innovations and new paradigms; complex wireless telecommunication systems simulation and behavioral modeling, linearization and efficiency techniques in multimode, multicarrier broadband nonlinear microwave, and mm-wave transmit power amplifiers; and smart antennas and MIMO channels. He has served on the Technical Program Committees of numerous IEEE and other international conferences and workshops.

IVAN GANCHEV [SM'08] (Ivan.Ganchev@ul.ie) received his Dipl.Eng. and Ph.D. degrees from the Saint Petersburg State University of Telecommunications in 1989 and 1994, respectively. He is an associate professor at the University of Plovdiv and an ITU-T Invited Expert. Currently he is lecturing at the University of Limerick, where he also acts as a deputy director of the Telecommunications Research Centre. Previously he served as a member of the Academic Network for Wireless Internet Research in Europe (ANWIRE) and two European Science Foundation COST research actions: "Modelling and Simulation Tools for Research in Emerging Multi-Service Telecommunications" and "Traffic and QoS Management in Wireless Multimedia Networks" (COST 285 and 290). Currently he is a member of the COST Action IC0906, "Wireless Networking for Moving Objects" (WiNeMO). His research interests include simulation and modeling of complex telecommunication systems, new communications paradigms for wireless next-generation networks, always best connected and best served, 3P-AAA, wireless billboard channels (WBC), and Internet tomography. He has served on the Technical Program Committees of a number of prestigious international conferences, including IEEE GLOBECOM (2006), IEEE VTC (2007–2010), and IEEE ISWCS (2006–2010).

Standardization in support of 3P-AAA, CIM and consumer-oriented ICC will make global transition to UCWW feasible. With this, or as part of this process, an implementation of a pilot global system would provide direct experience of CBM system design, validation, and deployment.

Standards Dynamics through an Innovation Lens: Next-Generation Ethernet Networks

Tineke Mirjam Egyedi, Delft University of Technology
Mostafa Hashem Sherif, AT&T

ABSTRACT

The inherent need for stable standards is difficult to reconcile with the aim of developing state-of-the-art standards and combine standardization with innovation. Standards change is inevitable, leads to increased transaction costs, and calls prior interoperability into question. In this article we analyze the problem of standards change as a feature of innovation. The focus is on the dynamics associated with committee standards when specifications and technologies co-evolve. We consider disruptions in the technology and/or the value chain simultaneously to classify innovations into four types. By identifying the type of innovation at hand, innovation-specific issues of standards change can be singled out. We illustrate this with the case of standards for next-generation Ethernet networks.

INTRODUCTION

The value of committee standards (i.e., specifications developed by consensus and intended for repeated use) depends to a large extent on how stable they are. Standards make life easier because we can refer to them and thus reduce what economists term the *informational transaction costs*. In contrast, changing standards increases them. In addition, switching costs could hinder standards upgrade or evolution.

In this article we address standards change as an inherent but potentially problematic source of innovation, an area not yet addressed in a systematic way in the innovation literature. Given that technology innovation requires changes to standards, when and where are issues of change likely to occur, and how can the various stakeholders deal with change? We use the case of the next-generation Ethernet to illustrate our points. In the next two sections we introduce the standards dynamics and innovation framework. We then apply it to developments in next-generation Ethernet standards. We attempt to systematically identify moments and situations in which the tension between standards stability and innovation is likely to emerge, and define tools decision makers could use to handle these tensions. In the conclusion we summarize the main findings and propose future areas of research.

STANDARDS DYNAMICS: DRIVERS FOR CHANGE

The term *standards dynamics* refers to changes to and interactions among standards after they have been set [1]. In the following we focus on two types of changes: maintenance and succession. Standards maintenance includes, for example, the development of corrigenda, the adoption of standards from other standards bodies, and standards withdrawal. Standards succession refers to the substitution of one standard or a generation of standards by another. Both changes are driven by the evolving nature of technology (i.e., *external* causes of change) and sources that are standardization-specific (i.e., *internal* causes of change).

Regarding external causes, information and communication technologies (ICT) and their standards are subject to the same market and technological forces within the same regulatory context. These forces include the competition among standards and the interactions among complementary standards in complex systems. Internal causes of standards change are specific to the context in which standards are developed and applied, such as:

- A flaw in the concept or content of the standard (e.g., if a committee's scope is too wide or a standard has become too comprehensive)
- A deficiency in the standards process (e.g., lack of consensus or absence of an important stakeholder)
- Lesser quality of the standard specification (e.g., ambiguous terminology, errors, or omissions)
- An upgrade to increase reliability or error recovery, or to extend the standard's domain of use

Although standards committees typically hope to develop stable standards and avoid unnecessary changes, the desire to lay an early claim on a

This is an updated version of the paper presented at the First ITU-T Kaleidoscope Conference on May 12–13, 2008.

 0163-6804/10/$25.00 © 2010 IEEE

market could favor a quick-and-dirty process leading to a low-quality standard. Thus, in the case of internal causes, a new revision or replacement is an unintended by-product of flaws in the *management* of the standards process when standards committees are involved. For example, a standard that is not based on a solid foundation of testing and implementation experience is more likely to require revision later on.

Although technology evolution sources of change are the focus of this article, the internal causes are also important because they are under the control of standards committees.

Standardizing Innovations

We focus below on how the evolving nature of a technological innovation affects the standardization process and the resulting standard. To do this we highlight some items from the literature on technological innovation, in particular, the technology life cycle and its relation to innovation and standardization.

Technology Life Cycle

The literature on the management of technologies considers five main stages in the life of a technology: emergence, improvement, maturity, substitution, and obsolescence [2]. An emerging technology stimulates the consolidation of new functional areas and the accumulation of new types of knowledge through research and field experience. As the properties of this emerging technology become better understood, new designs ameliorate its performance and increase the efficiency of the production processes. If the technology moves to the mainstream, its market share expands until its performance saturates. At this point, any substantial performance improvement will require a switch to a new technology (i.e., a technology transition) (Fig. 1). Traditionally, standardization takes place in the first three stages of the technology cycle — emergence, improvement, and maturity — although the movement for sustainable development is expected to expand the focus to all the phases of the life cycle.

Timing of Standardization

Standardization as it relates to these three stages is anticipatory, enabling, or responsive, respectively.

Anticipatory Standards — Anticipatory standards are forward-looking answers to expected interoperability problems. Their specification runs parallel to the production of prototypes, pilot services, and field trials. They provide a way of sharing ideas by systematically reworking the results of investigations and experimental data into useful engineering knowledge. Collaboration with competitors on the same set of problems is crucial when the market is as yet unknown and risks are high. Examples of anticipatory standards are the X.25 packet interface, the Secure Sockets Layer (SSL) for secure end-to-end transactions, and the Universal Mobile Telecommunications System (UMTS).

Because of the lack of detailed experience with the field performance of the technology as well as definitive market conditions, anticipatory standards are highly susceptible to the problem of *scope creep*. Also, incorrect market assumptions may lead to a dead end, such as in the case of Group 4 facsimile (facsimile on integrated services digital network [ISDN]) or the Wireless Access Protocol (WAP). Ideally, therefore, anticipatory standards should offer a minimum set of features to allow interconnectivity, stimulate the network externalities, and manage risks.

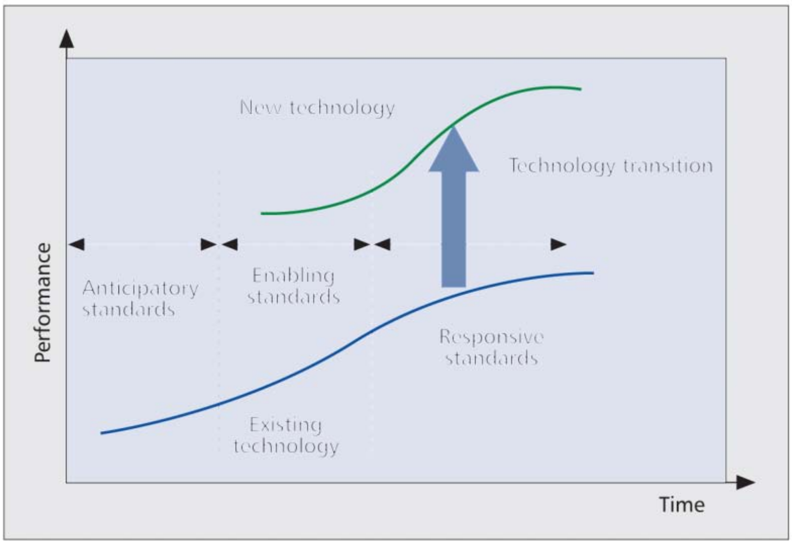


Figure 1. *Timing of standardization along the technology S-curve.*

Enabling Standards — The specification of enabling standards proceeds in parallel with market growth and enhancements to the technology and products. The principal goals are to reduce production costs, and improve agreed-on designs that extend their robustness and scale. Competitive market pressure is exerted on standards organizations to speed up the finalization of enabling standards [3]. An example of an enabling standard is International Telecommunication Union (ITU) Recommendation V.90 for modems at 56 kb/s. Even though proprietary designs were already available, to avoid market fragmentation as well as increase the overall market size, chip manufacturers agreed to collaborate with the rest of the value chain (modem manufacturers, computer manufacturers, and Internet service providers) and come up with an implementation that would work independent of the chip set used.

Responsive Standards — Responsive standards arise late in the technology life cycle to improve efficiencies or reduce market uncertainties for auxiliary products or services. Examples of responsive standards abound in the area of quality assurance. Another contemporary example is transport layer security (TLS), which was defined following the widespread use of SSL.

Classification of Innovations

There is an endless debate on the impact of standardization on innovation. Some studies purport to show that standardization spurs innovation, while equally valid data support the opposite claim. In general, leaps in technological

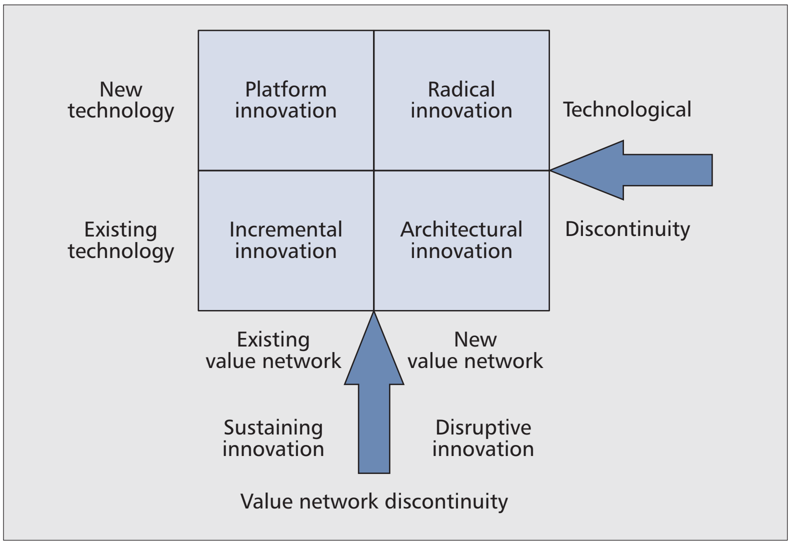


Figure 2. *Classification of innovations in terms of the value network and technological competencies.*

performance require technological discontinuities. C. M. Christensen [4] introduced the concept of discontinuities in a *value network* as well, that is, the set of attributes used to rank products, services, or technologies and determine their cost structure. Factors that can cause a discontinuity include new legislation, emerging standards, evolution of customer profiles, and others. Such a discontinuity opens opportunities to new entrants. Innovations that change the rank order are called disruptive, while those that preserve it are called sustaining.

Depending on the degree of changes they introduce in the existing technology and value network, innovations can be grouped into four categories, as shown in Fig. 2: incremental, architectural, platform, and radical innovations [5].

Incremental Innovations — Incremental innovations build on well-known technological capabilities to enhance an existing technology through improved performance, enhanced security, better quality, and reduced cost within the established value network. The purpose of the innovation is to enhance a company's competitive position by lowering costs through economies of scale and improving productivity through automation.

It is estimated that half of the economic benefit of a new technology comes from process improvements after the technology has been commercially established. This is why incremental innovations are typically process innovations. They tend to reinforce the existing industrial order because they are more readily integrated within the firm's strategy from both the technological and financial viewpoints. This contrasts with other types of innovation that alter the order and offer opportunities to new entrants.

Kuhn's [6] description of the way *normal science* operates gives us insight into the limits of incremental innovations: they should be consistent with and/or optimize the established design. This limits the overall ability to cope with changing markets or technical requirements.

Architectural Innovations — Architectural innovations (sometimes called system innovations) provide new functional capabilities by redefining the arrangements of existing technology to satisfy unmet needs (simplicity, cost, reliability, efficiency, convenience, etc.). Architectural innovations result from a market pull: new uses of an existing technology. For example, Bluetooth is a marriage of local area networks (LANs) and wireless communications in the unlicensed part of the radio spectrum at 2.45 GHz reserved for industrial, scientific, and medical (ISM) use. This technology for communicating with portable computers and personal digital assistants requires expertise in radio chip integration in addition to radio transmission, antenna design, and protocol engineering.

Architectural innovations not only modify supply chains; they tend to reorganize market segments, ultimately forming one or more new value networks. For example, in the case of Bluetooth, an industry-wide coalition promoted Bluetooth as a technology and as a standard.

Platform Innovations — Platform innovations correspond to a quantum leap in performance without major changes to the existing value chain. Programs to upgrade the existing infrastructure require large capital investments and are complex. New platforms alter the competitive position of larger firms by changing the technical criteria for competition. They weaken smaller firms because these have less access to financing as well as technical and financial talent. Because technology push is the main characteristic of platform innovations, technological considerations dictate business strategies.

For example, while the digital coding of signals is a radical innovation, the introduction of digital transmission in the 1970s was a platform innovation. Similarly, frame relay and the asynchronous transfer mode (ATM) are platform innovations in the core technology of public networks, both based on packet switching along the *connection-oriented* paradigm.

Radical Innovations — According to Kuhn [6], a "scientific revolution is a non-cumulative developmental episode in which an older paradigm is replaced in whole or in part by an incompatible new one." Radical innovations likewise provide a totally new set of functional capabilities discontinuous with both existing technological capabilities and value networks. Accordingly, these innovations face four types of uncertainties: technical uncertainty, resource uncertainty, organizational uncertainty, and market uncertainty. Technical uncertainty arises from two factors:
- Many of the technical characteristics of the innovation are not well understood.
- An even better technology may become available and displace the technology under development.

In telecommunications this is how optical transmission displaced the emerging waveguide technology. Resource uncertainties relate to the unknowns regarding the cost of development and implementation, as well as the collaborative network of technical, managerial, and marketing experts. Organizational uncertainties are due to

Communications | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | qMags

the tension from simultaneous discontinuities in the technology and the value network. Market acceptance is another unknown, because the more radical the technical innovation, the less likely that existing customer needs will be able to steer its development: market research methodologies typically focus on existing applications.

In the following, this framework is applied to analyze current efforts for standardizing the next generation Ethernet networks.

Standardization of the Next Generation Ethernet

Ethernet technology was originally used in LANs of enterprises and standardized in the IEEE 802.1 and 802.3 Working Groups. The bit rates were successfully increased from the original 2.4 Mb/s to 10 Mb/s to 100 Mb/s, 1 Gb/s, and finally 10 Gb/s; in parallel the physical media migrated from coaxial cables to twisted pairs and fiber optic cables. Virtual LANs (VLANs) were introduced in 802.1Q to partition terminal stations among communities within the enterprise. The majority of these extensions were incremental innovations that increased the access speed from desktop and laptop machines to the LAN.

10 Gb/s Ethernet, however, was of a different nature. Its standardization over twisted-pair cable (known as 10GBASE-T) required a sophisticated PHY layer and took more than three years to develop. Current 10 Gb/s Ethernet applications are mostly used in wide area networks (WANs) and carrier networks. Within enterprises, they are used to interconnect servers in data centers. Their use for connecting workstations in a LAN environment, however, has been limited by the restricted capacity of disk drives to process high data rates [7]. Accordingly, we can view 10 Gb/s Ethernet as a platform innovation suitable for two distinct value chains: carrier networks and high-speed server interconnections.

Initially, the extension of Ethernet to WANs was envisioned as a simple architectural innovation that would reuse existing bridge technologies to interconnect LANs with the WAN at the medium access control (MAC) layer. The lack of centralized network management, however, called for some form of coordination to allocate VLAN addresses. IEEE 802.1ad therefore defined a new service VLAN tag to be added by provider bridges at the ingress of the provider network and then removed at the egress to separate the VLAN address spaces of each entity. Because the MAC address space remained shared, customer addresses were also used in the provider network, and changes in one network would still be reflected in other networks (through flooding of frames with unknown destination, the construction of spanning trees, etc.) with obvious security and congestion implications. Therefore, a distinct MAC address space for the backbone network was defined in 802.1ah to be added to Ethernet frames at the ingress of the backbone and removed at the egress (so-called MAC-in-MAC operation). In addition, multiple Ethernet switched paths were defined in the core network for traffic engineering and load balancing (802.1Qay). All these changes were incremental steps.

For carrier applications, however, more basic technical changes were needed, such as a centralized approach to the management of MAC addresses and connection-oriented transport rather than the original connectionless mode of operation. Moreover, other substantive changes indicative of a platform innovation were needed. For example, from 2002 onwards the ITU — Telecommunication Standardization Sector (ITU-T) defined a management structure for Ethernet bridges in Recommendation G.8010. Also, the introduction of a fault management mechanism for fault isolation, verification, and notification, as well as protection switching and methods for clock synchronization had to be added [8]. To address these specialized areas, additional standardization organizations like the Metro Ethernet Forum (MEF) and the Internet Engineering Task Force (IETF) became involved. Thus, fault management procedures were defined in IEEE 802.1ag and ITU-T Recommendations G.8031, G.8032, and Y.1731. The case of fault management shows that tight coordination between standards bodies is essential to avoid incompatibilities; otherwise, standards may have to be modified prematurely. Clearly, the views on and approaches to achieving the 10 Gb/s bit rate started to diverge across application areas and value networks. As we argue, a more significant divergence is occurring at 100 Gb/s.

At present, the number of services requiring packetized 10 Gb/s pipes is increasing (e.g., video distribution). It is anticipated that within a few years there will be a need for a comprehensive network solution that aggregates traffic at rates of $N \times 10$ Gb/s. The current consensus is to standardize two such rates: $N = 4$ and $N = 10$ [9, 10]. There are already proprietary solutions operating at these rates. Therefore, the proposed standardization would not be anticipatory but enabling [3], provided they start before the expected market growth. This would reduce the time for standards development by building on experience gained from prestandard proprietary implementations. The corresponding innovation, however, would be a platform innovation since the aim is to radically improve technology performance.

To enable Ethernet transmission at 100 Gb/s, new technologies are needed that increase the reach of 100 Gb/s over fibers for WAN applications, and a higher-speed copper Ethernet interface must be agreed on. New multiplexing and line coding schemes are required for transmission over copper links, and new optical interfaces need to be designed, including 100 Gb/s transceivers, opto-electronic modulators, higher-performance network interface controllers (NICs), higher-speed input/output (I/O) buses, adapters, and so on.

Given space limitations, we illustrate our point by focusing on the 100 Gb/s standardization activities taking place in the IEEE 802.3 Higher Speed Study Group (HSSG) and ITU-T Study Group (SG) 15. IEEE 802.3 is positioning its activities as an incremental innovation intended to extend the longevity of existing 10 Gb/s and 40 Gb/s designs through technology reuse; keeping the same frame format, MAC layer, physical coding sublayer (PCS), and physical media attachment (PMA), backward compatibility would be ensured as well.

Market acceptance is another unknown, because the more radical the technical innovation, the less likely that existing customer needs will be able to steer its development: market research methodologies typically focus on existing applications.

Communications | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | qMags

Aspects	Standardization objective	
	Platform innovation	Incremental innovation
Value network	Carrier networks (e.g., long distances; high reliability)	Server interconnects (e.g., short distances)
Technologies to be standardized	New container for OTU/ODU, new modulation schemes, optical interfaces, network interface controllers, etc.	Extending existing technologies
Characteristics of standards dynamics	Enabling standard (time pressure); succession	Enabling standard (time pressure); revision; wide scope as cause for competing standards
Impact of standards change: compatibility	More effort needed to achieve backward compatibility	More straightforward backward compatibility

Table 1. *Two alternative approaches to standardization of next generation Ethernet networks at 100 Gb/s.*

Conversely, even when taking into account ITU-T's earlier work on 40 Gb/s, the scope of activities in ITU-T SG 15 corresponds unequivocally to a platform innovation because of the definition of a new transport container for optical transport units (OTUs)/optical data units (ODUs) — denoted OTU4/ODU4 — for 100 Gb/s Ethernet, definition of the characteristics of the interface (e.g., the forward error control scheme), and so on.

The way innovations for 100 Gb/s Ethernet are positioned as a platform or an incremental innovation could affect the long-term stability of their standards. In particular, the incongruence between the short-term advantages of incremental innovations and the technical challenges posed by platform innovations may lead to difficult compromises that would sooner require a standard's revision. Of course, some conceptual or technological breakthrough that resolves the above difficulties cannot be excluded. However, if we extrapolate existing conditions — and given the time pressure enabling standardization always faces — a reframing of the standards work should be explored. For example, activities in the IEEE 802.3 HSSG could be repositioned as a platform innovation by decoupling the 40 Gb/s activities from the 100 Gb/s efforts to restrict the incremental aspects to the first rate. The two alternative approaches to next-generation Ethernet are summarized in Table 1.

INNOVATION AND CHANGES IN STANDARDS

In this section, we identify moments and situations in which the tension between standards stability and innovation is likely to emerge. We close with a brief reflection on the Ethernet standards case.

TECHNOLOGY LIFE CYCLE AND TIMING OF STANDARDIZATION

First of all, let us not forget to mention the seemingly obvious: where a technology is mature and standardization more or less occurs ex post (responsive standardization), standardization resembles a selection process. Knowledge about and experience with the technology is readily available. The standard's implementation is already ensured. As a result, the standards committee's scope is likely to be focused and realistic. Those with a stake are sitting at the negotiation table. Technical ambiguity is less likely to become part of the specification. In short, a more stable standard is likely to result.

Standards are less likely to be stable where an emergent technology is standardized (anticipatory standardization). In this case there is little certainty about how the technology will develop and be applied due to the lack of experience. In standards committees scope creep is more likely to occur. Because the future market and market positions are as yet unclear, economic interests may favor ambiguities in a standard. In short, anticipatory standards are much more likely to be revised at a later time than enabling and responsive standards.

Moreover, the timing of standardization also determines the amount of pressure exerted on a committee process, and therefore on the quality of the process (e.g., amount of mistakes made, or need to accept a bad compromise). From this perspective, the pressures of the various stakeholders on enabling standards (where the coevolution of technology, markets, and standards takes place) would be the highest. The impact of standardization timing on potential changes in the specifications is shown in Table 2.

INNOVATION TYPE AND STANDARDS CHANGE

The combined effects of changes in the technology and the value chain on the expected changes to the underlying standards differ. For example, where incremental innovation occurs and a change of standard is necessary, standard revision will most likely include backward compatibility; whereas the disruptions that accompany radical innovation will require a new standard and thus preclude backward compatibility.

More specifically, the disruptions that define innovation categories also affect the actors participating in standardization and the area of standardization. To take the example of the 40/100 Gb/s Ethernet standardization activities, because the IEEE effort is defined as an incremental one, the same actors will be sitting around the table as in earlier rounds of Ethernet standardization. Market shares between providers are likely to remain stable. However,

Communications Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page qMags

Timing of standardization standards dynamics	Anticipatory standards (emergence)	Enabling standards (improvement)	Responsive standards (maturity)
Amount of expected changes	+++	++	+
Required speed of standardization	++	+++	+
External causes of change — firm's perspective	Immaturity of technology	Technology and market uncertainty	Cost reduction; performance optimization
External causes — application area perspective	Imprecise customer and supplier requirements	Market growth	Changes in market and regulations
Note: The number of + represents the qualitative significance of the factors.			

Table 2. *Factors affecting standards dynamics.*

when a platform innovation is the goal, we hypothesize that even when roughly the same actors participate in standardization (because the value chain is about the same), the stakes will be higher and the negotiation process harder than for an incremental innovation due to the technological disruption.

Conclusions

Standards co-evolve with technology innovation. Our goal was to identify when and where changes are likely to occur and in what way standardizers can anticipate possible change-related problems. We used the example of the next-generation Ethernet because its use as a carrier technology has encouraged cooperation between the IEEE and ITU-T. For this cooperation to bear full fruit, however, the difference in perspectives between the two groups needs to be considered at the standards management level.

The standards dynamics and innovation framework presented in this article address phases in the technology life cycle, timing of standardization, and innovation classes to identify sources of standards dynamics. Equally significant and at times difficult to do is to assess the kind of innovation to be standardized (i.e., whether the disruption affects the technology or the value network or both). Identifying the type of innovation at hand helps to prevent scope creep, anticipate potential committee tensions (e.g., height and diversity of stakes), and assess the implications of the type of standards dynamics (maintenance or succession, backward compatible or not). Table 2 summarizes the effects of various factors on standards dynamics. One possible line of future research would be to apply the same framework to the IETF and MEF as well as other standards groups interested in carrier Ethernet applications. Another challenge for research is to measure and quantify the entries of Table 2.

We recommend that standards committees use our framework to help stakeholders create a common ground for defining a feasible standardization scope before embarking on the work itself, and for assessing the impact of business pressures, differences in the value networks of the participants, and the achievability of backward compatibility.

Similarly, standard bodies whose work areas overlap can use the framework to discuss the nature and degree of their cooperation. The example of the next-generation Ethernet standards could help them gain insight into the reasons for choosing different standardization strategies, and improve their communication and collaboration.

References

[1] T. M. Egyedi and K. Blind, *The Dynamics of Standards*, Edward Elgar, 2008.
[2] T. Khalil, *Management of Technology: The Key Competitiveness and Wealth Creation*, McGraw-Hill, 2000.
[3] M. H. Sherif, "When is Standardization Slow?" *Int'l. J. IT Standards and Standardization Research*, vol. 1, no. 1, 2003, pp. 19–32.
[4] C. M. Christensen, *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*, Harvard Business School Press, 1997.
[5] W. J. Abernathy and V. B. Clark, "Mapping the Winds of Creative Destruction," *Research Policy*, vol. 14, no. 1, 1985, pp. 2–22.
[6] T. S. Kuhn, *The Structure of Scientific Revolutions*, 2nd ed., Univ. of Chicago Press, 1970.
[7] A. F. Brenner, P. K. Pepeljugoski, and R. J. Recio, "A Roadmap to 100G Ethernet at the Enterprise Data Center," *IEEE Apps. and Practice*, vol. 45, no. 11, suppl. 3, 2007, pp. 10–17.
[8] K. Fouli and M. Maier, "The Road to Carrier-Grade Ethernet," *IEEE Commun. Mag.*, vol. 47, no. 3, Mar. 2009, pp. S30–S38.
[9] M. Cvijetic and P. Magill, "Delivering on the 100 GbE Promise," *IEEE Apps. and Practice*, vol. 45, no. 11, suppl. 3, 2007, pp. 4–5.
[10] J. McDonough, "Moving Standards to 100 GbE and Beyond," *IEEE Apps. & Practice*, vol. 45, no. 11, suppl. 3, 2007, pp. 6–9.

Biographies

Tineke Mirjam Egyedi (T.M.Egyedi@tudelft.nl) is a senior researcher in standardization at Delft University of Technology. She leads a Next Generation Infrastructure project about standards and infrastructure flexibility, and another project on the rise of Inverse Infrastructures. She has published widely and participated in many European projects. Her latest book is *The Dynamics of Standards* (Edward Elgar, 2008), co-edited with K. Blind. She is President of the European Academy for Standardization and chairs the International Committee for Education about Standardization.

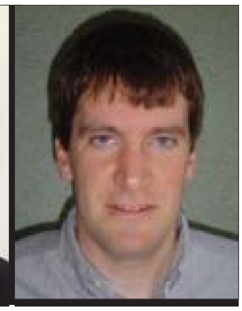
Mostafa Hashem Sherif has been with AT&T in various capacities since 1983. He is author of *Protocols for Secure Electronic Commerce*, 2nd ed. (CRC Press, 2003), *Paiements Électroniques Sécurisés* (Presses Polytechniques et Universitaires Romandes, 2006), and *Managing Projects in Telecommunication Services* (Wiley, 2006). He is the editor of the *Handbook of Enterprise Integration* (Auerbach, 2009). He is Standards Editor for *IEEE Communications Magazine*, an Associate Editor of the *International Journal of IT Standards & Standardization Research*, and a member of the editorial board of the *International Journal of Marketing*.

Communications Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page qMags

SERIES EDITORIAL

TOPICS IN INTEGRATED CIRCUITS FOR COMMUNICATIONS





Charles Chien *Zhiwei Xu* *Stephen Molloy*

In the past 10 years, we have witnessed a rapid paradigm shift in computing from that of mainframe to client-server. This paradigm shift has enabled the realization of cloud computing where data and computing resources are no longer located at the client side but are readily accessible in a remote data center via the Internet. To support cloud computing for a large population of Internet users, data centers must support enormous aggregate bandwidth on the order of hundreds of terabits per second. The required bandwidth will continue to rise due to increased broadband Internet users as attested by the more than 80 million in the United States alone as of 2010. By 2013, cloud computing is projected to increase by 300 percent.

The steep rise in aggregate bandwidth demands necessitates high-speed communications in the backbone and distribution networks that host the ever increasing Internet connections among clients and servers. Such communication technology includes synchronous optical network (SONET) for traditional carrier grade metro area networks, Ethernet for local and wide area networks, and passive optical network (PON) for fiber to the premises, all of which can currently support at least 10 Gb/s and 100 Gb/s in the future. The high speed requirement extends from external connections outside of the data center to within the data center itself. Each data center must host terabit-per-second routers, servers, and network storage devices that must be interconnected in complex backplanes at tens of gigabits per second. Since more than 50 percent of the cost in a data center results from the energy spent on cooling, the design of high-speed interconnect circuit technology must address the challenge of minimizing the power consumed per gigabit per second transferred while maintaining high reliability and link quality (e.g., less than 10^{-15} bit error rate).

While bandwidth has scaled up significantly, computation speed has also steadily increased over the past 10 years, from 1 GHz in 2000 to 3 GHz in 2010. The total computation power, however, has been boosted more than three times due to the introduction of a wider 64b bus and multicore CPU architectures. Today, it is customary to find processors with four cores in a server-grade CPU. Interestingly, the realization of a terahertz multiprocessor server system has been hampered by speed bottlenecks in the communication bus connecting the multiple processors and memory subsystems. The problem is even more severe when the multiple processors are located in separate chips. Similar to interconnect technology for backplanes of servers and routers, in this case circuit design must address the challenges in maintaining equally reliable link quality at high power efficiency but at lower complexity suitable for chip-to-chip communications.

In this issue of the Topics in Circuits for Communications Series, we have selected three articles that mark recent progress in the communications semiconductor industry for highly integrated transceiver systems on chip (SoCs) that enables future trends in high-speed interconnect technology for server backplanes and chip-to-chip communications for terahertz computing.

In the first article, "Design Considerations for High-Data-Rate Wireline Serial Interconnect Systems," the author addresses the challenges to realize power-efficient electrical interconnects for interchip communications over a range of ~80 cm in a server or router backplane, and ~10 cm on a printed circuit board line card that plugs into a backplane. Such interconnect technology enables high-speed interfaces such as PCI Express at 8 Gb/s, 10Base-KR at 10 Gb/s, SATA3 at 6 Gb/s, USB3 at 4.8 Gb/s, and IEEE 802.3ba at 40–100 Gb/s. The article gives an overview of architecture and circuit techniques that address the challenges of providing a low-jitter high-speed reference clock and compensating for line dispersion, crosstalk, skin effects, and impedance mismatch to achieve low bit error rate with low power dissipation. The author then discusses the design of a 2 × 2 full duplex 11 Gb/s interconnect chip in complementary metal oxide semiconductor (CMOS) technology.

A potential limitation in electrical interconnects lies in the increased crosstalk and higher power consumption as the data rate scales up. The second article, "Optical Technology for Energy Efficient I/O in High Performance Computing," reviews recent developments in optical interconnect circuit technology and architecture that may overcome such limitations at an acceptable price point as the required technology matures. In the near term, the article presents a hybrid MCM package to integrate high-performance optical components in existing microprocessor packaging technology to reduce cost at the expense of lower power efficiency (1 mW/Gb/s). In the longer term, the article describes a monolithic photonic CMOS process that can deliver data rate beyond 40 Gb/s at less than 0.3 mW/Gb/s.

The third article, "Wireless Proximity Interfaces with Pulse-Based Inductive-Coupling Technique," presents an interesting alternative to chip-to-chip interconnect based on near-field magnetic coupling among arrays of on-chip micro-coils. The micro-coils replace the traditional I/O pads and do not require electrical contact to make a connection with another chip. Such interconnect technology may potentially reduce the cost and area of high density interconnects. As an illustration, the article describes the effectiveness of wireless interconnect technology in the 3D integration of an 8-core CPU chip and a memory chip while achieving an aggregate data rate of 19.2 Gb/s with 0.15 mm^2/Gb/s area overhead and 1mW/Gb/s power efficiency.

We would like to take this opportunity to thank all the authors and reviewers for their contributions to this series. Future issues of this series will continue to cover circuit technologies that are enabling new and emerging communication systems. If the reader is interested in submitting a paper to this series, please send your paper title and an abstract to any of the Series Editors for consideration.

BIOGRAPHIES

CHARLES CHIEN is the president and CTO of CreoNex Systems, which focuses on technology development for next-generation communication systems. Previously he has held various key roles at Conexant Systems, SST Communications, and Rockwell. In his career he has architected several key products including a CMOS/SiGe chipset for multimedia over coax, IEEE 802.11abg WLAN RF CMOS transceiver and GaAs PA/RF switches, a wireless audio CMOS chipset for home theatre in a box, CDMA2000 cellular RF CMOS transceivers, and low-power wireless networked sensors. He was also an assistant adjunct professor at the University of California at Los Angeles (UCLA) from 1998 to 2009. His interests focus mainly on the design of SoC solutions for wireless multimedia and networking applications. He has published in various journals and conferences, and has authored a book entitled *Digital Radio Systems on a Chip*. He received his B.S.E.E. from the University of California at Berkeley, and his M.S. and Ph.D. from UCLA. He was a member of the technical program committee of ISSCC from 1998 to 2006.

ZHIWEI XU received B.S. and M.S. degrees from Fudan University, Shanghai, China, and his Ph.D. from UCLA, all in electrical engineering. He held industry positions with G-Plus Inc., SST Communications, Conexant Systems, and NXP Inc., where he did development of wireless LAN and SoC solutions for proprietary wireless multimedia systems, CMOS cellular transceivers, multimedia over cable systems, and TV tuners. He is currently with HRL Laboratories, working on software defined radios, high-speed ADC, and analog VLSI. He has published in various journals and conferences, one contribution to the *Encyclopedia of Wireless and Mobile Communications*, and has five granted patents.

STEPHEN MOLLOY received M.S. and Ph.D. degrees in electrical engineering from UCLA in 1993 and 1997, respectively, where his research focused on low-power circuits and architectures for video signal processing. This work led to the award of the Showman Prize from UCLA in 1997, and resulted in over a dozen conference and journal publications. He received his B.S. degree in electrical engineering from Rensselaer Polytechnic Institute in 1991. He served as Associate Editor of *IEEE Journal of Solid-State Circuits* from 2001 to 2004 and was a member of the technical program committee of the IEEE International Solid-State Circuits Conference from 1998 until 2005. He is currently vice president of engineering at Qualcomm, leading architecture development.

Design Considerations for High-Data-Rate Chip Interconnect Systems

Troy Beukema, IBM Research

ABSTRACT

Over the past decade, data rates for electrical interconnects in interchip communications systems have experienced a dramatic increase from <1 Gb/s to 10 Gb/s and beyond to keep up with ever increasing demands for more I/O bandwidth from modern high-capacity storage, networking, and data processing systems. This article presents an overview of the high-data-rate chip interconnect design space, including a short description of the channel, line equalization architecture, and design considerations for key I/O core subsystems realized in nanoscale CMOS technology.

INTRODUCTION

At the heart of modern information technology systems are powerful large-scale integrated circuits (ICs) that rely on high-data-rate electrical interconnects to rapidly bridge information flow among data processing, memory, network, and mass storage devices. To avoid bottlenecking the information processing capacity of these systems, industry standards are evolving to define new electrical interconnect protocols capable of transmitting data at line rates of 10 Gb/s [1] and higher, with advanced standards already considering wireline data transfer in the range of 25 Gb/s [2] to support future 100 Gb/s data transport systems [3]. These ultra-high data rates require design of line transceivers that incorporate low-jitter clock and data recovery with advanced equalization functions to combat loss of signal integrity arising from channel bandwidth limitations, reflections from impedance discontinuities, and crosstalk in package escape and board connectors. This article presents an overview of such high-data-rate interchip transceiver systems, covering both the high-level system architecture and key design trade-offs in performance-critical input/output (I/O) core subcircuits realized in deep-submicron complementary metal oxide semiconductor (CMOS) technology.

SYSTEM OVERVIEW

High-data-rate interchip communication systems can be categorized into two general application areas: long-range and short-range interconnects. Long-range (LR) systems operate over a backplane through one or more connectors with a total line run of up to ~80 cm (Fig. 1a). Short-range (SR) systems are used to realize direct chip-to-chip (C2C) interconnects on a common PCB without going through connectors, with a typical line run of ~10 cm (Fig. 1b). Many variations on these two major classes of interconnect systems are also used, such as chip-to-memory (C2M) channels, which may require one connector on a PCB, and short chip-to-optical module channels, which form the electrical interface between an IC and an optical transceiver for networks or optically attached disks.

A large number of off-chip I/O channels may be used to achieve a high aggregate data transfer capacity in an IC. As an example, a high-end switch application-specific IC (ASIC) may integrate 100 full-duplex LR channels, and a C2C memory interconnect bus may be 32 lanes wide, with several such buses integrated onto a single chip. The large number of I/Os in these systems impose a severe power constraint on the design of the line transceivers due to thermal design power (TDP) limits in the IC. A rough power guideline for SR I/O (C2C and C2M) systems realized in modern (65–32 nm) CMOS technologies is approximately 10–15 mW/Gb/s, while LR (backplane) I/Os typically target 20–30 mW/Gb/s. Although advances in CMOS technology can bring inherent power savings through reduction in transistor size, careful optimization of the line transceiver system architecture and circuit design is necessary to fully leverage the capability of the technology and achieve needed system power targets.

Further design considerations for the I/O core include chip area and performance. Chip area (roughly 0.5 mm^2/LR I/O in 65 nm process technology) must be minimized to lower cost and improve I/O escape density, permitting more I/O channels in the limited space available on the die. Key system performance metrics include transmit amplitude, transmit jitter mask, and receiver sensitivity and jitter tolerance. Jitter tolerance (JTOL) measures the ability of a receiver to operate at or below a threshold bit error rate (e.g., 1E-12 or 1E-15) with additional timing jitter added to the transmit signal. While both chip area and power dissipation constraints might be relaxed in a given application, the low bit error performance constraint is less forgiving. Lack of tolerance to line bit errors arises from minimal bit error correction capability in most high-speed

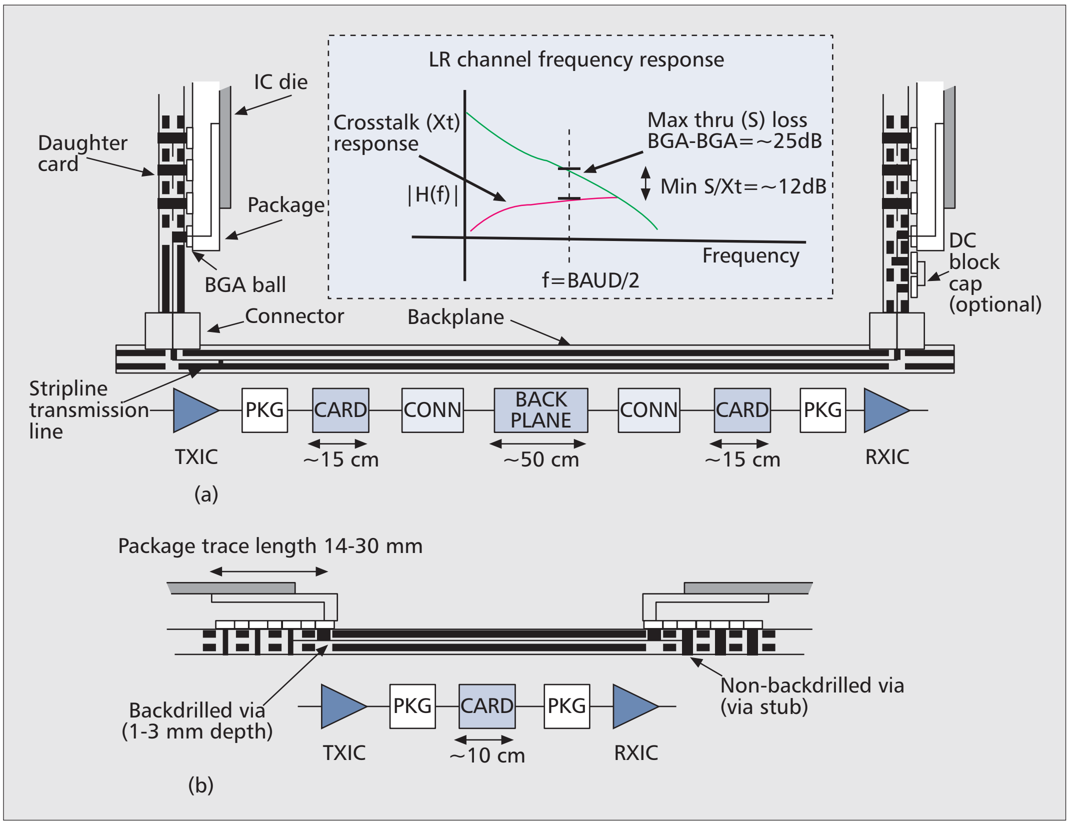


Figure 1. *Chip interconnect channels: a) long-range (LR) backplane interconnect channel; b) short-range (SR) chip-to-chip interconnect channel.*

The channel response is dominated by skin-effect loss in the copper transmission line traces, dielectric loss in the package and board material, and time dispersion arising from channel component impedance mismatches which generate reflections.

interconnect protocols. Low to no bit error correction protocols are often used to minimize both system complexity and decode latency, with memory access interconnect systems in particular requiring low data transmission latency to maximize system performance. As a result, the line transceiver must be designed to minimize hard decode errors and the data throughput loss associated with error detect/retransmission protocols.

ELECTRICAL INTERCONNECT CHANNEL

The characteristics of the electrical interconnect channel, including both the line dispersion and crosstalk, place an upper bound on achievable data rate [4] and directly drive the system design requirements for the chip I/O cores. In the high-data-rate chip interconnect space, the channel typically comprises a transmitter, a receiver, two package escapes, PCB via drops to impedance-controlled stripline transmission lines, and, in some cases, DC blocks and connectors, as shown in Fig. 1. The point-to-point (one transmitter to one receiver) channel design avoids signal integrity degradation from stub reflections that arise from multiple-terminated parallel buses, enabling the line data rates to scale to 10–20 Gb/s and beyond. Both single-ended and differential line signaling may be used, although differential line signaling is prevalent at multigigabit data rates due to inherent benefits including higher crosstalk tolerance and better common-mode noise rejection.

The channel response is dominated by skin-effect loss in the copper transmission line traces, dielectric loss in the package and board material, and time dispersion arising from channel component impedance mismatches that generate reflections. Although dielectric loss can be mitigated through the use of advanced board materials in ultra-high-data-rate (here defined roughly as 20 Gb/s+) systems, the skin-effect loss, which increases in proportion to the square root of frequency, is difficult to avoid as this loss is fundamental to an electrical interconnect. A plot of the impulse response of an 80 cm transmission line illustrating the time dispersion relative to 25 Gb/s signaling (40 ps bit period) is shown in Fig. 2a.

Reflections arising from impedance mismatches in the elements that constitute the channel are inherent in any practical physical system. A simplified channel model of a package escape reflection is illustrated in Fig. 2b as a 30 mm transmission line terminated with 300 fF capacitances. The impulse responses of this channel model for both lossless and lossy transmission lines are shown in Fig. 2a and 2b. The dominant (first double-bounce) reflection has a time delay of approximately 400 ps, while higher order reflections decay exponentially in amplitude and spread out in time. The 400 ps time delay is significant compared to bit period durations that drop to 40 ps at 25 Gb/s. This delay, coupled with the dispersion arising from conductor and dielectric bandwidth loss, directly influences the design of line equalizers in the I/O core.

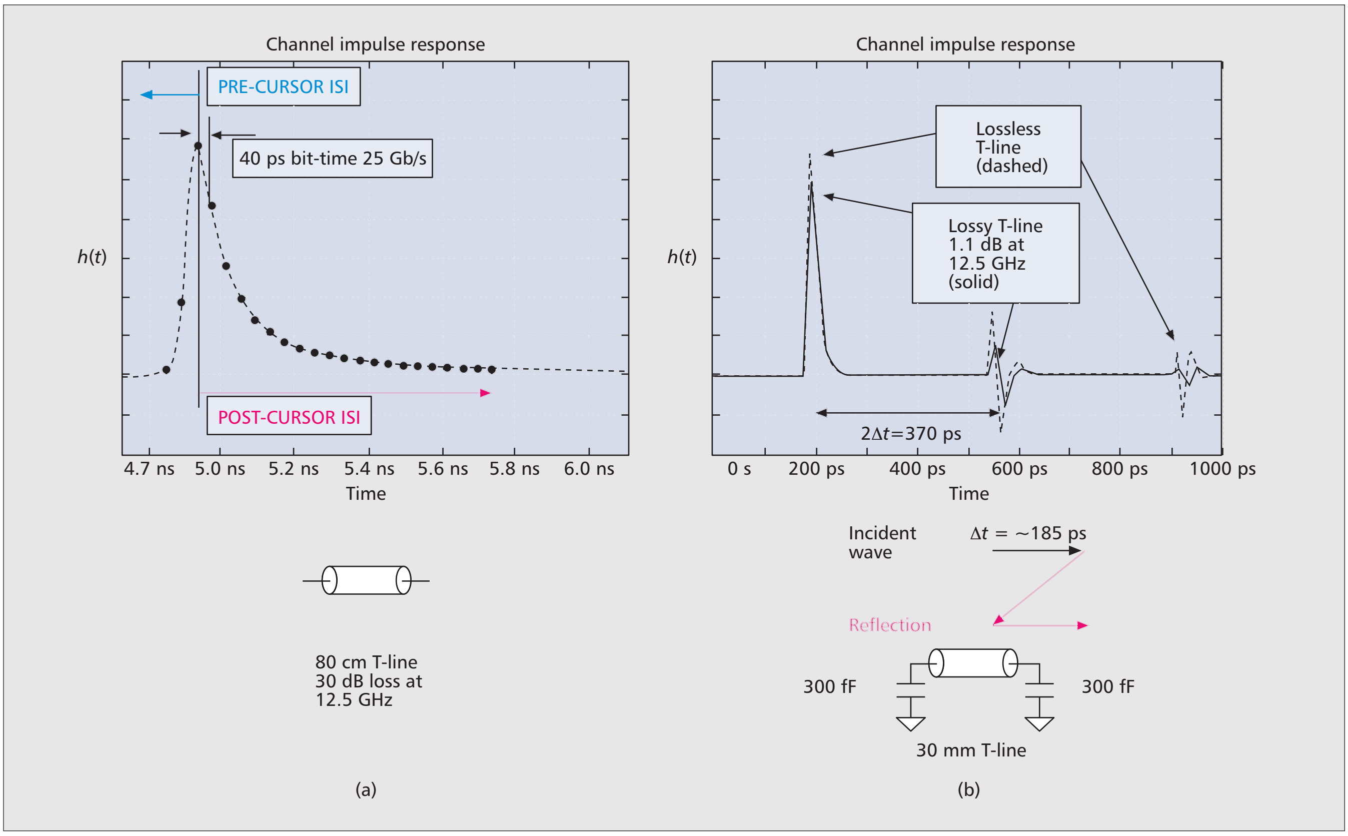


Figure 2. *Channel dispersion and reflections: (a) dispersion from channel loss; b) reflections from impedance discontinuities.*

Stub resonances, which can occur in vias, result in excess line attenuation beyond that of the inherent skin effect and dielectric loss of the PCB. In practice, via stubs must be eliminated (using techniques such as via back-drilling, illustrated in Fig. 1b) in high-data-rate interconnect channels to minimize loss and reflections arising from the via [5]. Another potential source of impedance discontinuity in the line is an external DC block capacitor, which may be used to isolate the common mode voltages of the transmitter and receiver. The external DC block may be eliminated if either the receiver integrates an AC coupling system on chip, or DC coupling is used in the system design. An integrated DC block in the receiver or DC coupling is desirable at extremely high data rates to eliminate the loss and reflection distortion arising from the extra PCB vias and discrete DC block capacitor.

Since groups of high-data-rate channels are used in almost all interchip I/O applications to form wide data buses, crosstalk among the channels must also be carefully considered in system designs. Crosstalk can become a dominant factor in limiting the maximum achievable data rate, since it rises at high frequencies where the data channel has higher attenuation (Fig. 1a inset). Large crosstalk can arise from coupling in the dense via field connecting the chip package to the board and in board connectors. Although crosstalk cancellation could be considered to improve system performance, the crosstalk cancellation process requires advanced knowledge of the data transmitted in the aggressor channels before the received data can be decoded. This requirement cannot be satisfied in a general backplane interconnect design, where crosstalk might arise in a connector from a completely unknown source. The extra power and area needed for crosstalk cancellation circuitry are also undesirable in a high-density chip I/O application. Without crosstalk cancellation in the I/O, the system crosstalk must be managed by providing sufficient isolation in the physical channel design, including all coupling paths arising in on-die circuit interactions, in the package escape region, and in the signal transmission lines and board connectors.

Line Equalization

High-data-rate chip I/O systems use line equalizers to combat the intersymbol interference (ISI) arising from dispersion in the channel. A standard equalization architecture employs a baud-spaced transversal feed forward equalizer (FFE) at the transmitter and a continuous-time equalizer (CTE) followed by a decision-feedback equalizer (DFE) in the receiver as shown in Fig. 3. The receiver CTE forms a high-pass frequency response to invert channel loss using a linear circuit with adjustable poles/zeroes. The CTE is followed by a DFE to cancel ISI in the channel from previously detected data bits. The DFE differs from FFE and CTE in that it is a nonlinear equalizer, meaning that the equalization signal is derived from a hard-limited (decision-feedback) version of the received signal. This feature enables it to compensate for high-frequency channel loss without amplifying high-frequency

Communications Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

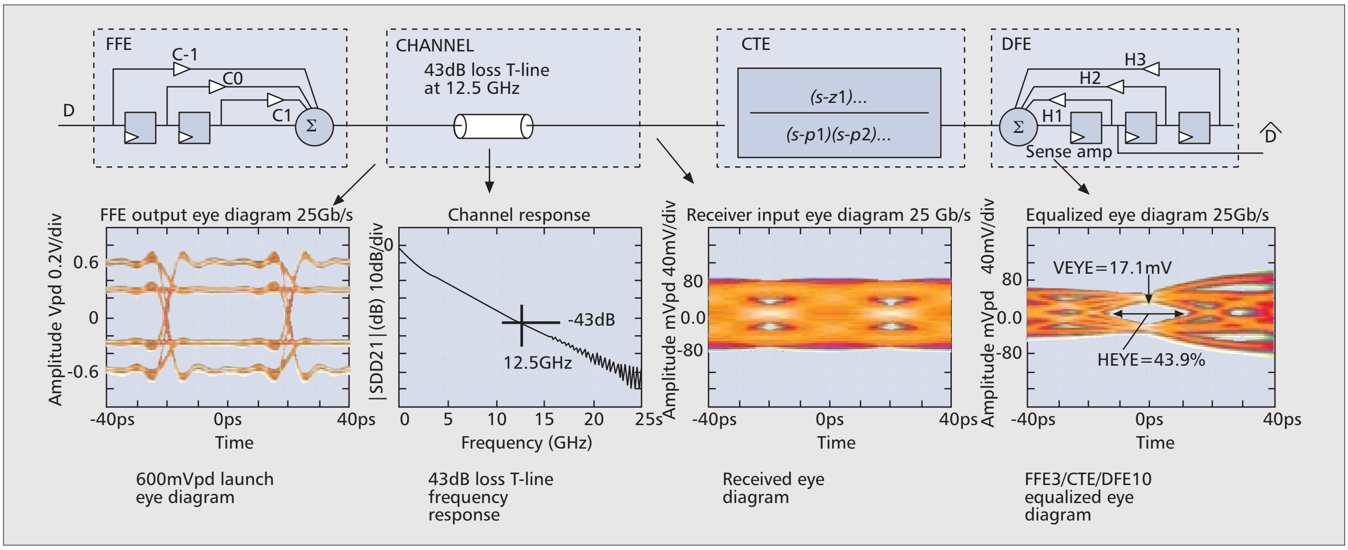


Figure 3. *Line equalization.*

thermal or crosstalk noise. The DFE can therefore improve the signal-to-noise ratio (SNR) of the received signal over that achievable by linear equalizers (FFE/CTE). This helps the transceiver system operate at low bit error rate (BER) in the presence of combined high frequency loss, crosstalk, and thermal noise.

The combination of transmit FFE and receive CTE/DFE can effectively equalize end-to-end channel losses up to 30 dB or even higher at the BAUD/2 frequency [6]. A simulated receiver eye diagram of an FFE3/CTE/DFE10 equalization chain applied to a 43 dB loss (at 12.5 GHz) transmission line at 25 Gb/s data rate is shown in Fig. 3. Although such a high level of loss most likely could not be tolerated in a practical system with added distortion from line reflections and crosstalk, this result illustrates that the FFE/CTE/DFE architecture is well matched to the task of equalizing high-loss transmission lines.

I/O Core System Design

The electrical interconnect I/O core is normally realized as a serializer-deserializer (SERDES) device in an advanced CMOS technology. The SERDES translates a parallel stream of data within an IC to a serial data stream at the interconnect line rate with a rate multiplier typically in the range of 8 to 32. In modern high-data-rate systems, the serial line rate varies from ~5 Gb/s to over 20 Gb/s. Non-return-to-zero (NRZ) line encoding is used in most chip interconnect standards. This signaling method has advantages of low-complexity encode/decode, low peak-to-average power ratio, and backward compatibility with legacy interconnect standards. Examples of modern industry interconnect standards include PCI Express Gen 3 (8 Gb/s), IEEE 802.3ap 10BASE-KR (10 Gb/s), Fibre Channel 16GFC (14 Gb/s), SATA3 (6 Gb/s), USB3 (4.8 Gb/s), Infiniband, and upcoming IEEE 802.3ba (40–100 Gb/s), which is defining the 40/100 GbE protocols. The 100 GbE protocol will most likely employ 10 × 10 Gb/s chip I/O in the near term with a goal to evolve to 4 × 25 Gb/s in the long term [2].

Key functional blocks of the SERDES include low-jitter clock generation using a phase-locked loop (PLL), clock and data recovery (CDR), and line equalization. A block diagram of a typical SERDES architecture is shown in Fig. 4. The system design includes a serializer, line equalizer, and line driver at the transmitter, a PLL for clock generation in the transmitter/receiver, and a line equalizer/sense amplifier/deserializer combined with a CDR unit at the receiver. A shared PLL with local phase mixers in each receiver can be used in system designs to leverage the low-jitter performance of one power- and chip-area-intensive inductor-capacitor (LC) tank PLL [6–8]. The die microphotograph in Fig. 4e illustrates a 2 × 2 full-duplex 11 Gb/s I/O core test chip with one shared LC PLL realized in 65 nm CMOS technology [7].

A complication in the design of the SERDES in modern systems is the need to satisfy a wide data rate range in one physical I/O. This wide data rate range permits a chip using the transceiver (e.g., a switch ASIC) to communicate to a multiplicity of different rate line cards simultaneously (back-supporting the lower data rates in earlier versions of interconnect standards), or be deployed in a range of different products with widely varying interconnect protocols/data rates. This requirement places a burden on the transceiver data path and clock generation circuits, which must be both broadband and as power efficient as possible at the same time, ideally scaling power linearly as the data rate changes. The following sections will present an overview of some design approaches for the key building block circuits used in chip I/O transceiver subsystems, with a short introduction on CMOS process technology used to implement the designs.

CMOS Process Technology

The high performance and integration of modern data processing and switch fabric ICs is achieved through process technology advances in

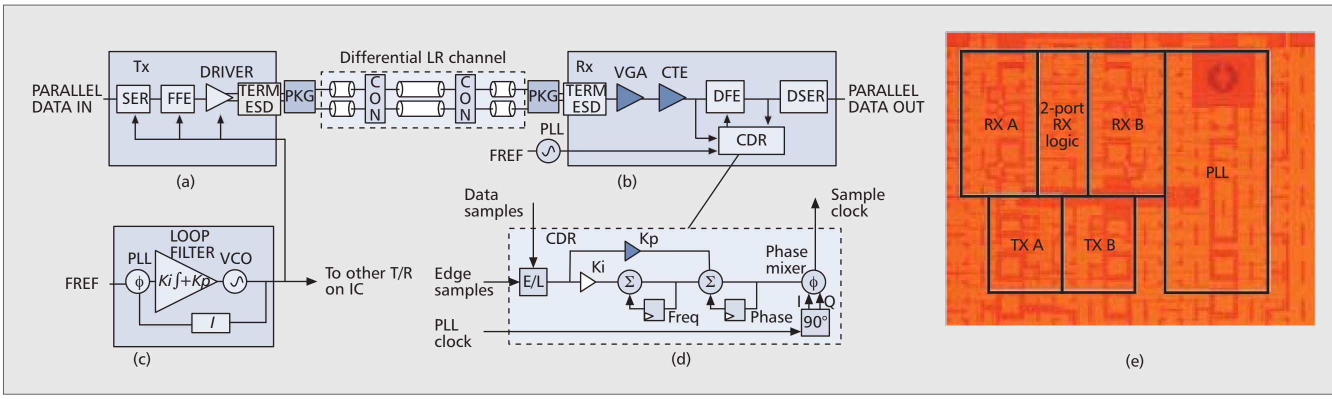


Figure 4. *High data rate SERDES system design: a) transmitter; b) receiver; c) PLL; d) clock-and-data recovery; e) die micrograph of 4-port 65 nm 11 Gb/s test chip.*

CMOS, primarily through lowering the device feature size, or minimum metal oxide semiconductor field effect transistor (MOSFET) gate length, while simultaneously increasing the maximum operating frequency (ft) of the devices. From 2000 to 2010 CMOS device feature size has dropped from ~130 nm to 32 nm, while ft has increased from ~100 GHz to ~300 GHz [9]. As feature size scales down, operating voltage is also lowered, from approximately 1.2–1.5 V in 130 nm to ~0.9–1 V in 32 nm. Reduction in operating voltage is dictated by fundamental device limits (in turn arising from reduced gate oxide thickness needed to maintain FET transconductance gain as feature size and operating voltage scale down) or the desire to lower system power. Although reduction in feature size directly benefits digital circuits by permitting higher levels of integration, other device characteristics such as FET output impedance (Gds) tend to degrade as the process technology node shrinks and operating voltage drops. This increases the difficulty of designing circuits with sufficient gain, linearity, and signal swing to achieve needed dynamic range in analog data paths. Further technology trade-offs are discussed in the following sections, which present an overview of I/O core subsystem designs in modern nanoscale CMOS process technology.

Transmitter System

Key system specifications for the transmitter include launch amplitude, deterministic jitter (including bandwidth loss, duty cycle, and sinusoidal jitter), random jitter (arising primarily from noise in the PLL), line impedance matching, electro-static discharge (ESD) robustness, and equalization capability. Two example line driver/equalization architectures that may be used are a current-mode logic (CML) based design [6, 7], shown in Fig. 5a, or a switched-CMOS- based design [10], as shown in Fig. 5b. Advantages of the CML driver include high-data-rate operation, power supply noise rejection from the balanced differential pair, and constant current draw for low switching noise. These advantages come at a cost of lower efficiency and higher system power arising from the linear bias current in the CML stage, which dissipates power across the current switch devices and tail current device that is not delivered to the load. The CMOS driver avoids the linear bias current by using hard-limiting switching devices to connect the line to either pull-up or pull-down terminations. This design approach adds power efficiency and receiver termination flexibility at the cost of increased power supply switching noise and reduced power supply noise rejection as described in [10]. Because power efficiency is critical in the I/O, the switched-CMOS approach is a preferable architecture as long as the CMOS switch devices are fast enough to support the required data rate.

Transmitter Equalization — Equalization can be added to the line driver through use of a baud-spaced feed-forward equalizer (FFE) built using a tapped delay line, as shown in Fig. 5a. The equalizer commonly employs one pre-cursor tap and either one or two post-cursor taps with adjustable sign and magnitude control. By adjusting the sign and magnitude of the delay-line taps, the frequency response of the transmitted signal is effectively predistorted at the transmit launch to compensate for ISI arising from dispersion in the channel. In most cases the transmit FFE will provide a low-frequency de-emphasis function. With FFE de-emphasis, high-frequency data patterns such as 1010... will be sent down the line with no attenuation, while low frequency data runs such as 1111... are effectively launched with a lower drive level. This provides a frequency peaking response to help compensate loss at high frequency in the channel. The transmitter FFE taps may be either programmed to fixed values based on known channel characteristics or adapted on a per-channel basis using an ISI minimization criterion [11].

Electro-Static Discharge Protection — To protect the CMOS devices that connect to the external line from potential ESD damage, a diode and/or silicon-controlled rectifier (SCR) can be used to shunt ESD transients to ground [10]. The ESD protection device(s) add unwanted capacitance to the I/O port. This capacitance degrades both the signal bandwidth and impedance matching to the channel. To mitigate this problem, passive compensation networks constructed using T-coils [10] can be added to the line terminations on both transmitter and

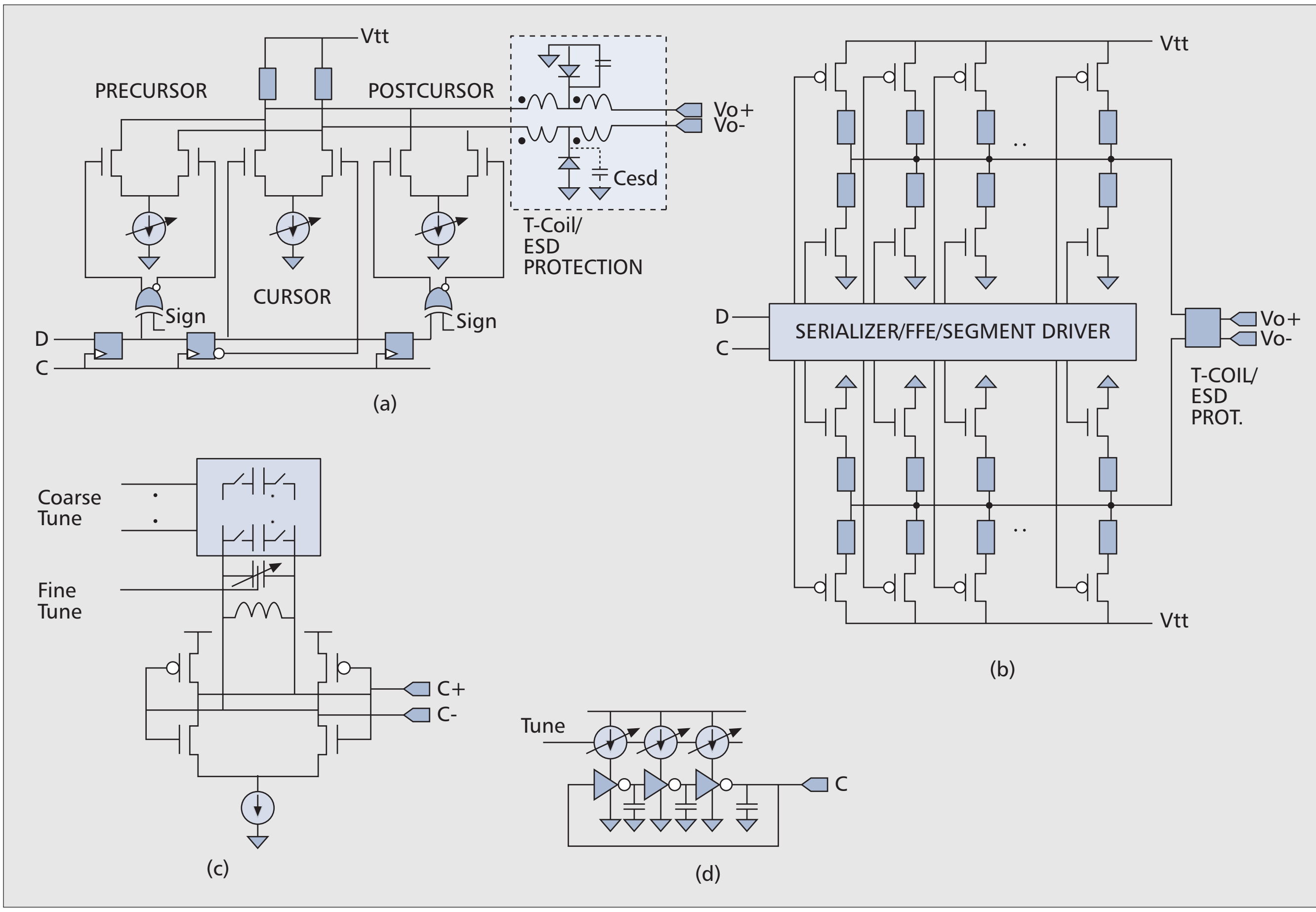


Figure 5. *Transmitter and VCO circuits: a) full-rate CML FFE/line driver; b) CMOS source-series terminated FFE/line driver; c) band-tuned LC VCO; d) ring VCO.*

receiver. A diagram of a T-coil/ESD protection network at the transmitter output is shown in Fig 5a. The compensation network improves line impedance matching at a cost of increased chip area and a small series ohmic loss due to the inductors in the T-coil, which can normally be absorbed by the line termination.

Phase-Locked Loop

A key component of the serial transceiver is a low-jitter PLL used to generate the data launch clock at the transmitter [12] and detection latch clock at the receiver. The critical system performance metric of the PLL is its output clock random jitter, or variation in clock edge time with respect to that of a noiseless clock. Short-range interconnect systems commonly use clock forwarding to increase tolerance to jitter on the PLL clock. In a clock-forwarded system, a given data bit launched with a jitter-advanced or -delayed transmit clock will automatically be sampled earlier or later by the receiver since it makes use of the same jittered clock that launched the data bit to sample it at the receiver. Non-clock-forwarded (asynchronous) systems used in LR interconnects do not have this advantage, however, since in this case the data launch and detect clocks are generated with independent PLLs with uncorrelated jitter. Some LR interconnect systems may lock both transmit and receive PLLs to the same reference clock to achieve jitter correlation within the tracking bandwidth of the PLL, but the LR I/O cannot in general rely on any transmit-to-receive jitter correlation. This results in the need for both low-jitter PLLs in the transmitter and receiver, and a CDR function in the receiver capable of tracking the relative jitter between transmit and receive clocks.

A block diagram of a basic PLL is shown in Fig. 4c. The key components of the system include a phase detector, a proportional-integral loop filter, a low-jitter voltage-controlled oscillator (VCO), and a frequency divider. The frequency divide ratio is determined as the ratio of the output frequency to the PLL reference frequency, which is typically on the order of 100 MHz for asynchronous I/O applications. This results in divide ratios of 100 to 200 to achieve 10–20 GHz output frequencies. The PLL output jitter is related to the closed-loop phase noise at the output of the VCO, which in turn is influenced by the phase noise of the fundamental oscillator, the phase noise of the reference clock, the PLL loop dynamics, and the noise in the circuits used to construct the PLL. A rough rule of thumb for maximum PLL loop bandwidth is 1/10 the reference frequency, setting the PLL bandwidth to <10 MHz for a reference clock at 100 MHz. Within the loop-bandwidth of the PLL, the phase noise of the reference clock is multiplied in direct proportion to the loop divide

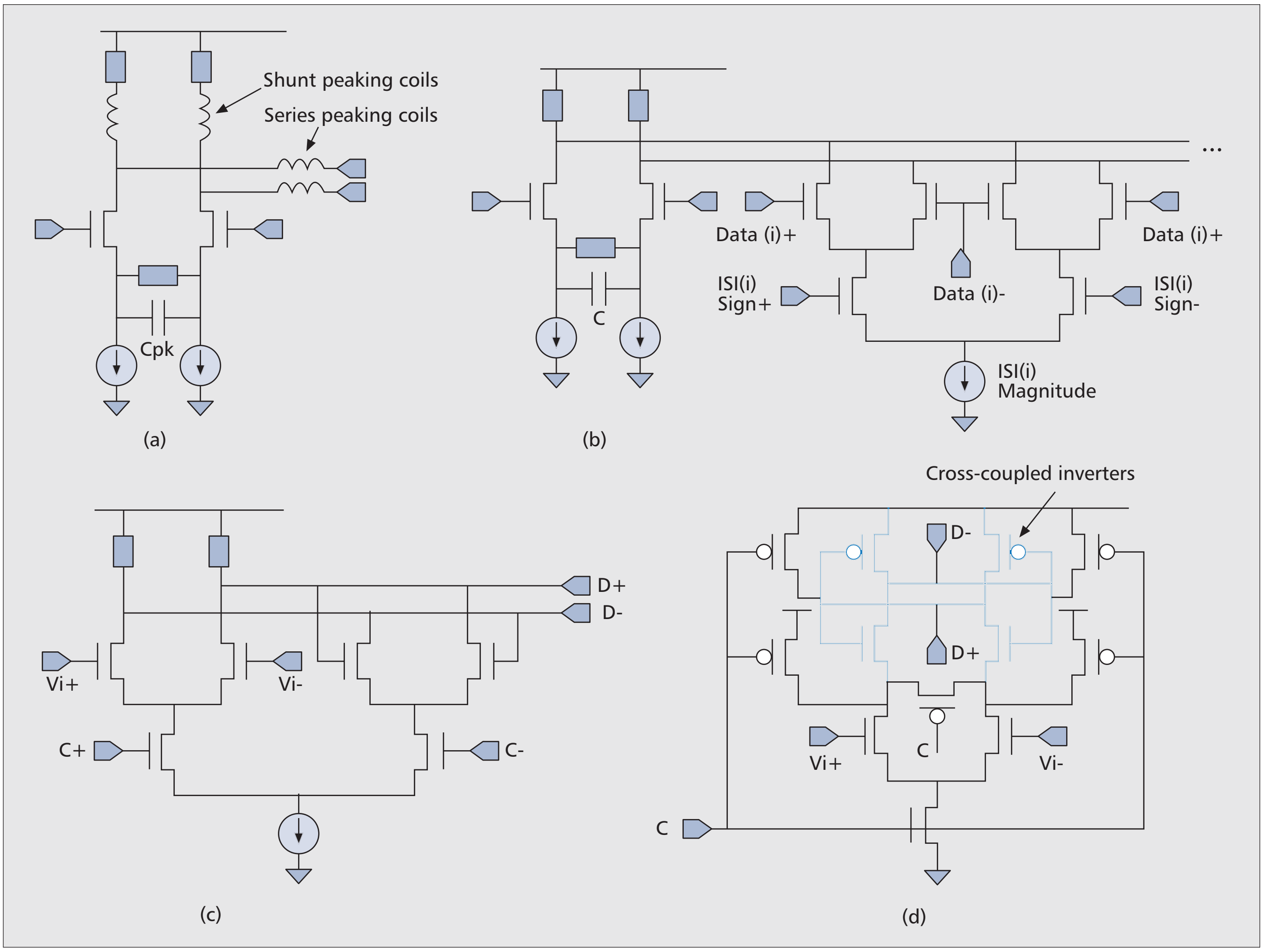


Figure 6. *Receiver circuits: a) continuous-time equalizer; b) DFE summer; c) CML sense amp; d) DCVS sense amp.*

ratio. This places stringent low-phase-noise requirements on the reference oscillator for loop multipliers of 100 to 200, which increase the reference phase noise by 40 to 46 dB at the output of the VCO within the tracking bandwidth of the PLL.

Above the loop bandwidth of the PLL, the output jitter is dominated by the phase noise of the open-loop VCO. To achieve a low-jitter VCO, needed in asynchronous LR I/O, LC-tank-based oscillators [12, 13 chapter 7] are used due to the improved Q provided by the second-order resonant circuit (Fig. 5c) compared to a lower-Q CMOS ring oscillator (Fig. 5d). The CMOS ring VCO [14] is useful in more jitter-tolerant forwarded-clock C2C systems due to its smaller chip area compared to the LC VCO. To lower phase noise in the LC tank oscillator, the VCO tuning gain (in Hertz per volt) is reduced to as small a value as possible by employing the use of several overlapping coarse pretune bands across the frequency coverage range of the PLL [6, 7, 12]. The pretune bands are switched in using discrete capacitance steps in the LC tank circuit as shown in Fig. 5c, while a continuously variable capacitance varactor is used to achieve loop phase lock. This approach enables a wide frequency coverage range for the PLL while minimizing output phase noise. The PLL must also be resistant to noise sources from power supply or other adjacent circuits on the IC die, which can result in unwanted spurious tones (i.e., sinusoidal jitter) and increase the time jitter of the PLL output.

Receiver System

The line receiver is responsible for making correct binary decisions on a data signal that may have a completely closed eye diagram at the receiver input (Fig. 3) due to a combination of channel bandwidth limitation, reflection ISI, crosstalk noise, transmitter jitter, and thermal noise in the channel. A key system-level requirement for the receiver is its jitter tolerance capability, which enables it to correctly decode a data signal launched at the transmitter with a time-jittered clock. Other receiver design requirements include minimum and maximum signal level handling capability, channel loss and crosstalk handling capability, and line input termination impedance match/ESD protection. Interconnect standards do not normally specify a particular receive architecture, but may provide an informative reference model, such as a 1- to 5-tap DFE-based receiver with CTE, to be used with reference channels to help guide the receiver implementation.

Receiver Datapath — The received signal at the IC input propagates through an ESD protection and line termination network similar to that used at the transmitter, then passes to a first stage buffer amplifier (Fig. 4b). Depending on the system design, this stage may incorporate variable gain, variable high-frequency peaking (to realize a CTE), or both. An important requirement of receiver systems that perform equalization is the need to maintain linearity of the signal. Linearity can be maintained through the use of a variable-gain amplifier (VGA) driven by an automatic-gain control (AGC) loop designed to maintain signal swings within the linear operating range of the circuits in the datapath.

A variable CTE function in the receiver increases the maximum loss handling capability of the equalization system by providing added high-frequency boost beyond that achievable by the transmitter FFE de-emphasis alone. Common implementations of this equalizer employ one or more capacitively degenerated amplifiers (Fig. 6a) to selectively increase gain at high frequency. A key feature of this CTE topology is its ability to add gain at high frequency while maintaining no DC gain loss, thereby maintaining the highest possible low-frequency SNR in the datapath and reducing the need for following gain stages to restore needed signal swing. Passive inductors may also be used in shunt and/or series peaking configurations to assist in neutralizing the impact of parasitic wiring and intrinsic device capacitances to help achieve the needed high frequency gain.

Receiver Decision-Feedback Equalizer — The receiver DFE uses past data bit decisions to drive polarity controls of an ISI cancellation summer, as shown in the diagram in Fig. 3. The DFE ISI cancellation summers can be realized using resistively [6, 7] or capacitively [8] loaded current-mode logic (CML) stages. A block diagram of a resistively loaded CML DFE summer with a single ISI cancellation tap is shown in Fig. 6b. In this circuit, a CML XOR gate structure is used to multiply the sign of the data history bit with the sign of the channel ISI to steer a current proportional to the ISI magnitude to the load resistors of the DFE summer stage. The DFE tap weights are normally dynamically adapted [6] using an ISI minimization constraint. The adaptation process adjusts the sign and magnitude of the ISI cancellation taps in the DFE to achieve zero average correlation between previously received data bits and the ISI at data sampling time. Because the DFE relies on past data decisions to drive the tap polarity controls, it is only able to compensate post-cursor ISI (Fig. 2a).

A practical issue in the realization of DFE at high data rates is the need to detect a data bit and apply the sign of this bit (for NRZ signaling) to the associated ISI cancellation tap in the DFE summer before the next data bit can be decoded. If it is not feasible to perform this operation in real time for the first and/or second previous data bits, they can be evaluated using a look-ahead speculation approach [15] that latches decisions from all possible ISI sign combinations of the speculated data bits in parallel. Later data bits (second or third previous and beyond) have sufficient time to apply their signs to the associated DFE taps dynamically. A description of a hybrid speculative/dynamic feedback DFE design is given in [6] for a 6 Gb/s DFE realized in 130 nm CMOS; this technique has also been applied in [7, 8] for 90 nm and 65 nm designs operating up to 11 Gb/s.

Receiver Sense Amplifier — Following CTE/DFE correction, the equalized analog data signal is converted to a binary level using a high-precision sense amplifier. The sense amplifier must reliably detect overdrive levels in the range of 10 mVp or lower to achieve the best possible receiver sensitivity. Two basic circuits suitable for use as high-precision sense amplifiers are the CML and differential-cascode voltage switch (DCVS) based binary latches, as shown in Figs. 6c and 6d, respectively. The CML latch has the benefit of very short clock-to-Q delay (delay time from latch clock to data out) since it operates by fast current steering, producing a single-ended output swing lower than the rail-to-rail swing of the DCVS latch. An example of the high-data-rate capability of a CML-based latch is the 40 Gb/s transceiver described in [16], which achieves 10 mV sensitivity CML latches clocked at 20 GHz.

A large advantage of the DCVS latch over the CML latch is its inherent ability to offer power scaling vs. operating frequency due to its switched-CMOS architecture, which avoids the need for CML quiescent bias currents. The CML latch design cannot easily achieve rate-proportional power scaling without complicated variable biasing/load resistor schemes, which can negatively impact the high-frequency performance of the latch. The DCVS latch also has the additional benefit that its memory is cleared by reset switches for every new sample, so it has inherently low sampling hysteresis and therefore high sensitivity potential. These considerations point to the DCVS latch as a preferable sense amplifier for application in high-sensitivity multi-protocol/broadband line receivers, as long as the regeneration time of the cross-coupled inverter (highlighted in Fig. 6d) is fast enough in the application CMOS technology to achieve the needed clock-to-Q delay at the highest operating data rate of the I/O core.

Clock and Data Recovery — The receiver is responsible for generating a low-jitter sample clock that is aligned with the received data stream to enable optimum decoding performance. The data clock can be derived from either an explicit clock line from a transmitter (used in a forwarded-clock system for short C2C/memory links), or a local low-jitter PLL, as shown in the system diagram in Fig. 4. In either case, a CDR loop [13, chapter 9] is needed in the receiver to align the frequency and/or phase of the local high-speed PLL clock to achieve an optimum sampling clock for data detection. The most general CDR design is fully asynchronous (meaning it must track frequency offsets ranging up to approximately ±200 ppm in non-spread-

A variable CTE function in the receiver increases the maximum loss handling capability of the equalization system by providing added high-frequency boost beyond that achievable by the transmitter FFE de-emphasis alone.

I/O core system designs are evolving to employ more digital/switched-CMOS dominant approaches as CMOS technology migrates to ever smaller geometries and faster devices with lower supply voltages and intrinsic device characteristics that favor digital-based designs over linear analog circuit implementations.

spectrum clock systems and up to ±2000 ppm or more in spread-spectrum systems) with a requirement to lock to the incoming data and track jitter with a bandwidth of approximately BAUD/1667 [17] to meet the jitter-tolerance mask requirements of many serial interconnect standards. The BAUD/1667 CDR tracking bandwidth represents a trade-off between filtering out unwanted data-dependent jitter in the received signal (for which a low CDR tracking bandwidth is desired to avoid incorrectly biasing the recovered clock) and tracking uncorrelated high-frequency jitter arising from the transmit/receive PLL clocks (for which a high tracking bandwidth is desired). Interconnect standards may specify different CDR tracking bandwidths depending on the assumed clock jitter and data-dependent jitter the standard intends to cover.

The CDR tracking bandwidth is determined by its jitter transfer function, which is defined as the magnitude ratio and phase difference between sinusoidal jitter on the CDR output clock and input data as a function of jitter frequency. Within the CDR loop tracking bandwidth, this ratio stays close to unity (unity magnitude ratio and no phase shift), enabling the local data sampling clock to track jitter on the incoming data. This provides the receiver with large jitter tolerance (JTOL) within the CDR loop tracking bandwidth. At higher jitter frequencies, the CDR output clock jitter begins to lose both amplitude match and phase coherence with the input data jitter due to the lowpass characteristic of the CDR loop, and jitter tolerance drops. The system JTOL can also be degraded by CDR jitter generation, which is defined as jitter appearing on the CDR output clock that originates within the receiver system. As an example, jitter might arise from high-frequency noise on the local PLL clock due to insufficient power supply isolation from clock trees or other circuits within the IC.

To track frequency offsets with no steady-state phase error, a second-order CDR loop is used. A common second-order CDR implementation employs an early/late phase detector, followed by a proportional-integral loop filter that drives a clock phase integrator as shown in Fig. 4d. The clock phase integrator can be realized as either a dedicated voltage-controlled oscillator (VCO) or a digital integrator that modulates the phase of a fixed-frequency VCO using a quadrature generator and phase mixer [6, 7] as shown in Fig. 4d. The CDR jitter transfer function is optimized by selecting appropriate proportional loop gain K_p and integral loop gain K_i values to achieve the desired loop bandwidth. Jitter transfer peaking is one of the most problematic issues to address at ultra-high data rates that require wide (BAUD/1667 = 15 MHz at 25 Gb/s) jitter tracking bandwidth. Jitter peaking arises due to the destabilizing influence of latency in the CDR loop implementation. Loop latency is primarily added through clocked logic delays in the early/late phase detector, loop filter, and phase interpolator control logic if used. Excessive CDR loop latency results in frequency-dependent amplification of jitter and degraded receiver jitter tolerance.

Summary and Conclusions

This article has presented a short survey of the high-data-rate chip interconnect space, including an overview of the channel environment and a summary of system requirements and circuit design trade-offs for key elements of the I/O cores. The chip interconnect space is roughly divided into two main categories, short-range chip-to-chip and long-range board-to-board. Data rates are evolving from the range of 5–10 Gb/s in wide use today to projections of over 25 Gb/s in the near future to efficiently support applications such as 100 GbE. Key design constraints for the I/O core system include minimization of die area and power draw, while maintaining uncorrected line bit error rates in the range of 1E-12 and lower. The realization of robust system designs at ultra-high data rates will require co-optimization of both the physical channel to minimize loss, reflections, and crosstalk, and the I/O core to operate reliably with bit periods approaching 40 ps and smaller in 25 Gb/s+ systems. Key elements of the I/O core system design include a low-jitter PLL, high-precision CDR, and line equalization subsystems. Example circuit trade-offs with basic building block circuit topologies for each of these key subsystems have been discussed. I/O core system designs are evolving to employ more digital/switched-CMOS dominant approaches as CMOS technology migrates to ever smaller geometries and faster devices with lower supply voltages and intrinsic device characteristics that favor digital-based designs over linear analog circuit implementations.

Acknowledgments

The author would like to thank his colleagues in both IBM Research and the IBM System and Technology Group for their extensive help over the years in his research and development efforts in the area of high-data-rate I/O systems. The author would also like to thank Daniel Friedman for his assistance in reviewing the initial draft of this manuscript.

References

[1] A. Healy, "Challenges and Solutions for Standards-Based Serial 10 Gb/s Backplane Ethernet," *IEEE CICC*, 2007, pp. 139–44.

[2] J. D'Ambrosia, D. Stauffer, and C. Cole, "CEI-28G: Paving the way for 100 Gigabit," Optical Internetworking Forum; http://www.oiforum.com/public/documents/OIF_CEI-28G_WP_Final.pdf

[3] A. F. Benner, P. K. Pepeljugoski, and R. J. Recio, "A Roadmap to 100G Ethernet at the Enterprise Data Center," *IEEE Commun. Mag.*, vol. 45, no. 11, Nov. 2007, pp. 10–17.

[4] D. G. Kam *et al.*, "Is 25 Gb/s On-Board Signaling Viable?," *IEEE Trans. Advanced Packaging*, vol. 32, no. 2, May 2009, pp. 328–44.

[5] J. Zhang *et al.*, "Chip-to-Chip Communication Beyond 25 Gb/s — Modeling and Realization," *DesignCon*, 2010.

[6] T. Beukema *et al.*, "A 6.4-Gb/s CMOS SerDes Core with Feed-Forward and Decision-Feedback Equalization," *IEEE J. Solid-State Circuits*, vol. 40, no. 12, Dec. 2005, pp. 2633–45.

[7] J. F. Bulzacchelli *et al.*, "A 10-Gb/s 5-Tap DFE/4-Tap FFE Transceiver in 90-nm CMOS Technology," *IEEE J. Solid-State Circuits*, vol. 41, no. 12, Dec. 2006, pp. 2885–2900.

[8] J. F. Bulzacchelli *et al.*, "A 78 mW 11.1 Gb/s 5-Tap DFE Receiver with Digitally Calibrated Current-Integrating Summers in 65 nm CMOS," *IEEE ISSCC '09*, Digest Tech. Papers, pp. 368–70.
[9] A. M. Niknejad *et al.*, "Nanoscale CMOS for mm-Wave Applications," *CSIC*, 2007, pp. 1–4.
[10] M. Kossel *et al.*, "A T-Coil-Enhanced 8.5 Gb/s High-Swing SST Transmitter in 65 nm Bulk CMOS With <16 dB Return Loss Over 10 GHz Bandwidth," *IEEE J. Solid-State Circuits*, vol. 43, no. 12, Dec. 2008, pp. 2905–20.
[11] T. Toifl, M. Schmatz, and C. Menolfi, "Low-Complexity Adaptive Equalization for High-Speed Chip-to-Chip Communication Paths by Zero-Forcing of Jitter Components," *IEEE Trans. Commun.*, vol. 54, no. 9, Sept. 2006, pp. 1554–57.
[12] A. L. S. Loke *et al.*, "A Versatile 90-nm CMOS Charge-Pump PLL for SerDes Transmitter Clocking," *IEEE J. Solid-State Circuits*, vol. 41, no. 8, Aug. 2006, pp. 1894–1907.
[13] B. Razavi, *Design of Integrated Circuits for Optical Communications*, McGraw-Hill, 2003.
[14] A. Hajimiri, S. Limotyrakis, and T. H. Lee , "Jitter and Phase Noise in Ring Oscillators," *IEEE J. Solid-State Circuits*, vol. 34, no. 6, June 1999, pp. 790–804.
[15] S. Kasturia and J. H. Winters, "Techniques for high-Speed Implementation of Nonlinear Cancellation," *IEEE JSAC*, vol. 9, no. 5, June 1991, pp. 711–17
[16] J.-K. Kim *et al.*, "A Fully Integrated 0.13u CMOS 40-Gb/s Serial Link Transceiver," *IEEE J. Solid-State Circuits*, vol. 44, no. 5, May 2009, pp. 1510–21.
[17] "Fibre Channel-Methodologies for Jitter and Signal Quality Specification (FC-MJSQ)," T11.2 project 1316-DT Rev. 14, Annex C, June 19, 2004, pp. 197–99.

Biography

Troy Beukema (troyb@us.ibm.com) received his B.S.E.E and M.S.E.E degrees from Michigan Technological University, Houghton, in 1984 and 1988, respectively. Since 1996 he has been with IBM Research in Yorktown Heights, New York, where he has focused on both wireless and wireline system designs in the Communication Technology department. His most recent activity at IBM is in the area of high-data-rate SERDES system design and analysis. He has developed line equalization and clock-and-data recovery system designs for backplane I/Os operating from 6 to 25 Gb/s spanning CMOS technology generations from 130 nm to 32 nm. His ongoing research interests include both ultra-high-data-rate SERDES and low-power ADC-based I/O core designs.

Optical Technology for Energy Efficient I/O in High Performance Computing

Ian A. Young, Edris M. Mohammed, Jason T. S. Liao, and Alexandra M. Kern, Intel Corporation
Samuel Palermo, Texas A&M University
Bruce A. Block, Miriam R. Reshotko, and Peter L. D. Chang, Intel Corporation

ABSTRACT

Future high-performance computing systems will require optical I/O to achieve their aggressive bandwidth requirements of multiple terabytes per second with energy efficiency better than 1 pJ/b. Near-term optical I/O solutions will integrate optical and electrical components in the package, but longer-term solutions will integrate photonic elements directly into the CMOS chip to further improve bandwidth and energy efficiency. The presented near-term optical I/O uses a customized package to assemble CMOS integrated transceiver circuits, discrete VCSEL/detector arrays, and polymer waveguides. Circuit simulations predict this architecture will achieve energy efficiency better than 1 pJ/b at the 16 nm CMOS technology node. Monolithic photonic CMOS process technology enables higher bandwidth and improved energy efficiency for chip-to-chip optical I/O through integration of electro-optical polymer based modulators, silicon nitride waveguides, and polycrystalline germanium (Ge) detectors into a CMOS logic process. Experimental results for the photonic CMOS ring resonator (RR) modulators and Ge detectors demonstrate performance at up to 40 Gb/s and analysis predicts that photonic CMOS will eventually enable energy efficiency of 0.3 pJ/b with 16 nm CMOS. Optical interconnect technologies with multilane communication or wavelength-division multiplexing will further increase bandwidth to provide the multiple-terabyte-per-second optical interconnect solution that enables scaling of high-performance computing into and beyond the tera-scale era.

INTRODUCTION

The microprocessor architecture transition from multicore to many-core will increase chip-to-chip input/output (I/O) bandwidth demands at processor/memory interfaces and in multiprocessor systems. Near-term projections, shown in Fig. 1, estimate that CPU-to-memory interconnects will require 100 Gbytes/s bandwidth in 2012–2013. Future many-core architectures will require bandwidths from 200 Gbytes/s to 1.0 Tbyte/s and begin the era of tera-scale computing.

To meet these bandwidth demands, traditional chip-to-chip electrical interconnect techniques will require increased transceiver circuit complexity and costlier materials. However, due to electrical channel loss, increasing I/O bandwidth in electrical links eventually comes at the cost of reducing interconnect link length, reducing signal integrity, or increasing power consumption.

In contrast, optical interconnects have negligible frequency-dependent loss and low crosstalk. Performance is independent of link length (for lengths of interest in chip-to-chip I/O), and little or no equalization is required. This motivates chip-to-chip I/O architects to consider optical I/O as a means of scaling data rates in a power-efficient manner.

ELECTRICAL LINK ISSUES

Figure 2 shows the components of a typical high-speed electrical link, including the transmitter, receiver, timing system, and channel. A phase-locked loop (PLL) frequency synthesizer generates the transmit serialization clocks, and the receiver timing system provides the serial data sampling clocks. The design complexity of the transmitter and receiver increases to include additional equalization circuitry as data rates scale above electrical channel bandwidths.

Electrical channel frequency characteristics are dependent on channel length. An inset in Fig. 2 shows the channel response for three typical electrical channels, a 17-in server backplane channel with two connectors, a 7-in desktop channel with no connectors, and an 8-in high-performance cable channel. The frequency-dependent loss exponentially increases with channel length, as illustrated by the loss difference between a 17-in backplane channel and a 7-in desktop channel. Attenuation and dispersion in these low-pass channels introduces intersymbol interference (ISI) in high-speed data patterns. Equalization can cancel ISI and open the received data eye, but requires additional circuit complexity, which increases I/O power and area. Equalization is typically implemented with a progressive combination of transmitter (Tx) feed-forward equalization (FFE), receiver (Rx) continuous-time linear equalizers (CTLEs), and decision feedback equalization (DFE).

S. Palermo was with Intel Corp. and is now with Texas A&M University.

Communications | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | qMags

A detailed circuit simulation study shows that electrical link bandwidth is limited by either the channel, at the frequency beyond which loss cannot be overcome with equalization, or the complementary metal oxide semiconductor (CMOS) technology, at the frequency beyond which the required equalization techniques cannot be implemented in an energy-efficient manner [1]. In 45 nm CMOS with constant 1 V_{pp} transmit signaling, power efficiency of transmitter and receiver front-end circuits initially improves as the data rate increases. This trend reverses, and power efficiency begins to decline with data rate as more complex equalization becomes necessary. While transmit equalization can be implemented with little additional energy, a CTLE with sufficient gain-bandwidth product requires significant power, so the energy efficiency degrades rapidly once a CTLE becomes necessary. Ultimately, the maximum data rates in all three channels shown in Fig. 2 are limited by the equalization circuit speed, as the 45 nm technology cannot support efficient DFE in the 20 Gb/s range.

Circuit simulation estimates based on a predicative 16 nm CMOS technology node reveal that the faster transistors remove the CMOS technology limitations and allow efficient implementation of all equalization circuitry necessary to operate the two conventional electrical channels at their fundamental limits [1]. Channel loss, transmit peak power constraints, receiver sensitivity, and jitter eventually limit the maximum data rate at which the desired 10^{-12} bit error rate (BER) can be achieved in backplane and desktop channels, even though significant equalization is used. For the shorter lengths, a high-performance low-loss flex cable channel is still technology-limited because it does not require DFE until the data rate exceeds 40 Gb/s, at which point a DFE cannot be implemented efficiently in the projected 16 nm node.

Optical I/O Implementation Using a Hybrid MCM Package

For the near term, the proposed 12-channel optical transceiver architecture allows package integration of low-cost high-performance optical components in existing microprocessor package technology. This hybrid architecture integrates CMOS and discrete optical components in a multichip module (MCM) package. In this architecture a multichannel optical transceiver chip, an 850 nm 10 Gb/s GaAs vertical-cavity surface-emitting laser (VCSEL) 1 × 12 linear array (or n × 12 array), and a PIN photodiode 1 × 12 linear array (or n × 12 array) are flip-chip mounted on a standard microprocessor organic land grid array (OLGA) package substrate. The CMOS drivers and receivers on the transceiver chip are electrically coupled to the VCSELs and photodiodes with very short transmission lines routed on the top surface of the package. The VCSEL and photodiode arrays are optically coupled to on-package integrated polymer waveguide arrays with metalized 45° mirrors. The waveguides couple the optical signals from the VCSELs and photodetectors to standard multiterminal (MT) fiber optic connectors, which connect to 1 × 12 (or n × 12 array) waveguide or fiber arrays to couple the light off-chip.

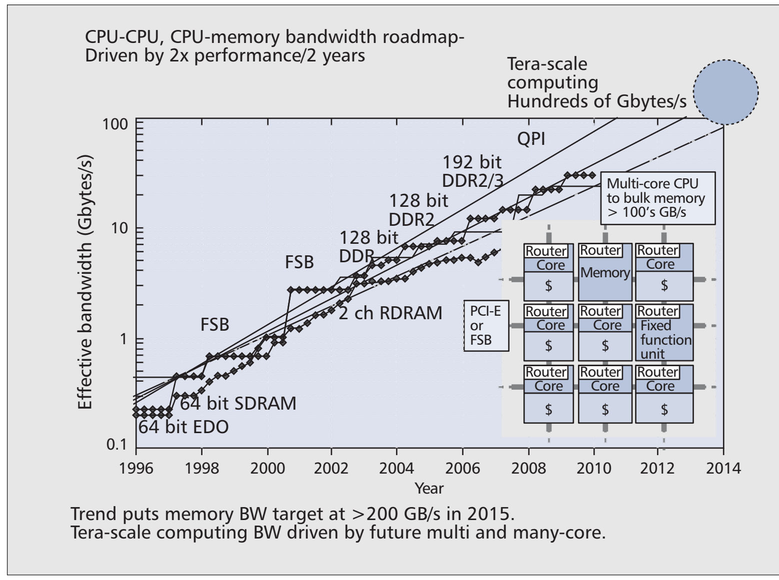


Figure 1. *Historical CPU trend — I/O bandwidth vs. year.*

The Transceiver Chip Architecture

The transceiver architecture for optical I/O shares many common features with the typical electrical transceiver shown in Fig. 2, including the serializer, deserializer, clock generation, and clock recovery. Transmitter clocks are generated with a PLL, and receiver clocks are either recovered from the data with a CDR (as in the presented link) or forwarded from the transmitter. Equalization complexity is significantly reduced compared to electrical links, but new circuits are required to perform the electrical-optical-electrical conversion. New package technologies are required to integrate the optical and electrical components of the link. The following sections describe the package and circuit innovations that enable a 10 Gb/s hybrid optical link in 90 nm CMOS.

Package Architecture

The package architecture allows the integration of low-cost high-performance optical components with standard microprocessor flip-chip OLGA package technology [2]. Figure 3 shows a photograph of the fully assembled optical transceiver package and a drawing illustrating the subcomponents. The package substrate is a stack of laminated copper layers separated by a dielectric. A trench is fabricated in the substrate to accommodate the multimode polymer waveguides, which have square apertures with a total height of 100 μm, core dimension of 35 μm × 35 μm, and pitch of 250 μm. The 12-channel polymer waveguide array is 10 mm long and 3 mm wide. A standard 12-channel MT optical connector on one end of the waveguide array connects to a fiber optical cable to couple light in and out of the package. An array of 45° mirrors on the other end of the waveguide bend the optical signal 90° in order to couple into and out of the VCSEL and photodiode arrays, which are flip-chip bonded face down onto the package. The 45° mirror cut is formed either by microtome or laser ablation and its reverse side is metalized. The loss from this 45° mirror is 0.3 dB. The high-speed electrical lines used to connect the CMOS chip to the optical

Current VCSEL technology is rated for 10 Gb/s, beyond which the VCSELs are bandwidth-limited with a slow transient tail due to intrinsic and extrinsic parasitic effects such as carrier diffusion and device parasitic capacitance. Pre-emphasis can compensate for these effects and increase the achievable data rate.

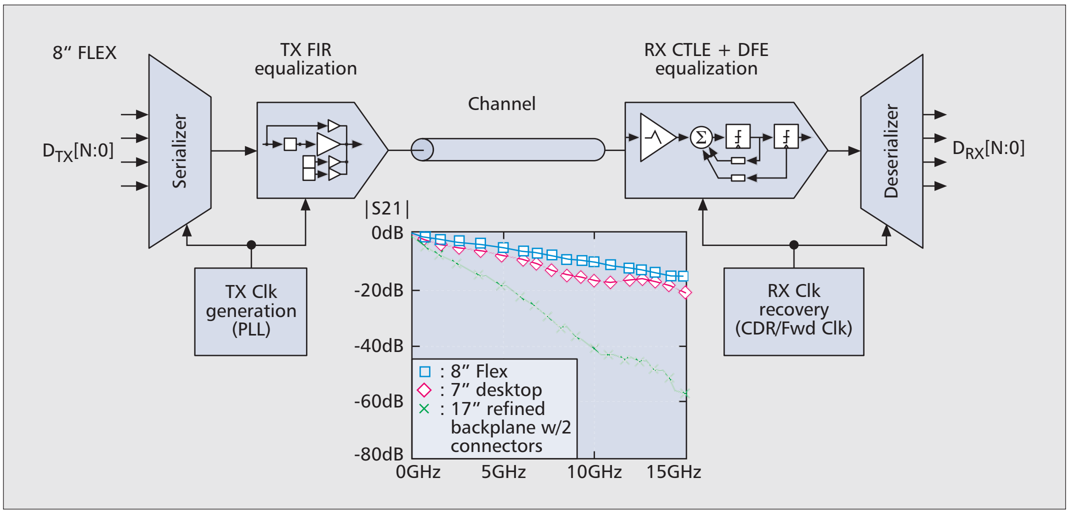


Figure 2. *High-speed electrical link block diagram showing serializer, TX PLL, TX finite impulse response (FIR) equalizer, RX continuous-time linear equalizer (CTLE), RX decision feedback equalizer (DFE), CDR, and deserialize.*

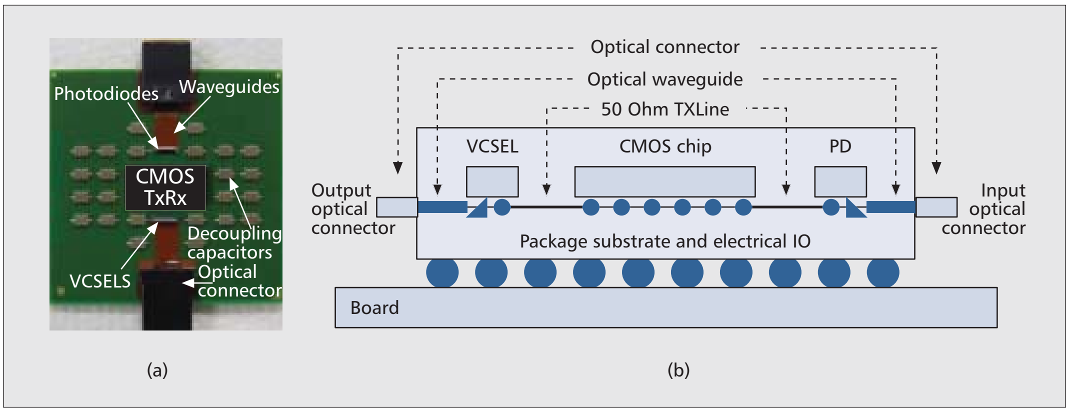


Figure 3. *a) A fully assembled optical transceiver unit; b) a schematic side view of the same unit, showing the optical coupling scheme of VCSELs/photo-detectors to waveguides through a 45° mirror.*

components are routed as controlled impedance (50Ω single-ended or 90Ω differential) microstrip traces on the top surface of the substrate where they have the best high-frequency characteristics. The optical signal is transmitted by an oxide-confined 850-nm 10 Gb/s 1 × 12 VCSEL array with peak optical output power greater than 3 mW (~5 dBm) and received by a 10 Gb/s, 1 × 12 GaAs PIN detector array with a diameter of 75 μm, a capacitance of 330 fF, a 3 dB bandwidth of 8 GHz, and a responsivity of 0.6 A/W. The total optical loss budget for the end-to-end link includes VCSEL and photodiode coupling loss through the 45° mirrors, propagation loss through the waveguide, MT connector loss, and Fresnel losses at the interfaces in the connectors. The total optical loss budget calculated for the complete link is 10 dB [2]. Improvements in optical coupling for this hybrid package architecture are in development to reduce the optical loss budget to as low as 6.8 dB.

Circuits: VCSEL Driver and Transimpedance Amplifier

Current VCSEL technology is rated for 10 Gb/s, beyond which the VCSELs are bandwidth-limited with a slow transient tail due to intrinsic and extrinsic parasitic effects such as carrier diffusion and device parasitic capacitance. Pre-emphasis can compensate for these effects and increase the achievable data rate. The VCSEL driver described in [1] directly generates dual-edge pre-emphasis with sub-bit-period pre-emphasis waveform timing precision. The pre-emphasized current waveform is generated by summing the main modulation current with a delayed and weighted peaking current in order to produce pre-emphasis pulses at each data transition. Typical average currents provided to the VCSELs range from 6 to 10 mA, which corresponds to an average optical power of 1.5 to 2 mW. The VCSEL driver is output terminated and connected to the VCSEL through a 50Ω microstrip transmission line routed on the top surface of the package. As the VCSEL technology develops for higher modulation speed (using quantum dots rather than quantum well technology), high-data-rate VCSELs at 20 Gb/s and higher will still benefit from these pre-emphasis techniques to further extend data rates.

The transimpedance amplifier (TIA) uses the differential symmetric-feedback topology [1],

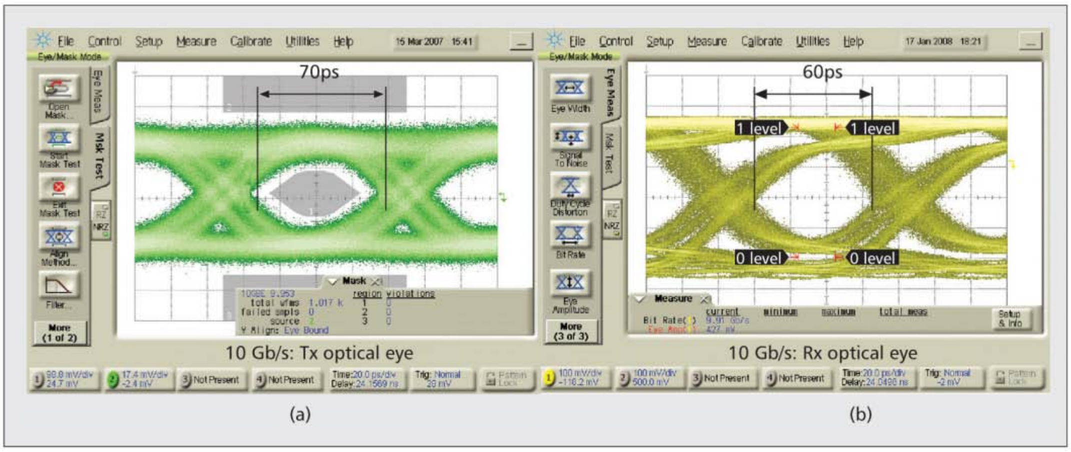


Figure 4. *Optical eye diagrams for 10 Gb/s tested with fully packaged: a) transmitter optical output; b) receiver optical input.*

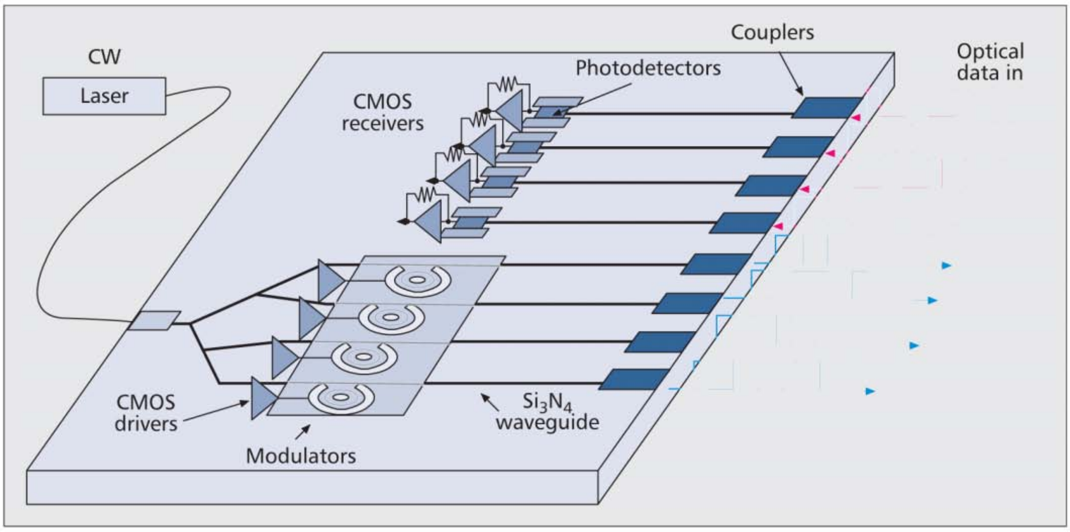


Figure 5. *Photonics optical interconnect architecture.*

In a photonic CMOS process for integrated optical links, the additional process steps required for photonics must not degrade or interfere with the front-end CMOS transistor performance. Furthermore, the process must allow fabrication of all required optical components on the same die.

which converts the single-ended input current to a differential output voltage to help mitigate supply noise at subsequent gain stages and provides a data rate above 12.5 Gb/s when the total input parasitic capacitance C_p is less than 250 fF. The TIA receives a single-ended photocurrent of 200 μA from the photodiode and generates a differential 2 × 50 mV_{pp} output that is fed to a following limiting amplifier (LA), which converts it to a CML level output. The LA consists of a cascade of CML buffers. In the packaged transceiver the combined capacitance of the photodiode, metal pad, bump, and ESD could be as high as 500 fF. This capacitance limits the maximum data rate that can be measured for the packaged receiver channel. The same TIA tested electrically with wafer probing had an open electrical eye diagram at 18 Gb/s for an input capacitance of 90 fF. This indicates there is a strong dependence of bandwidth on the input parasitic capacitance.

Experimental Results

10 Gb/s optical measurement results are shown in Fig. 4 for a fully assembled transmitter and receiver. For the transmitter measurement, external differential electrical pseudorandom data was sourced into the chip to drive the CMOS pre-emphasis VCSEL driver, and the VCSELs were biased with an average current of 7 mA. The measured transmitter optical eye opening was 70 ps. The receiver demonstrated an open electrical eye for optical 10 Gb/s input data. The electrical received signal eye opening was 60 ps with a peak-to-peak jitter of 30 ps. The individual transmitter and TIA receiver circuits are capable of operation at up to 18 Gb/s [2].

Photonic CMOS Optical I/O

Architecture

In the longer term, monolithic integration of photonic elements in a CMOS process will enable significant improvements in I/O performance, energy efficiency, and cost. The proposed monolithic photonic CMOS process, illustrated in Fig. 5, integrates modulators, waveguides, and detectors on top of the metal interconnect layers in the far back-end of a standard CMOS process. Light from a continuous-wave (CW) source is coupled onto the die and modulated using integrated waveguide-based modulators driven by on-chip circuits, such that the electrical signals do not

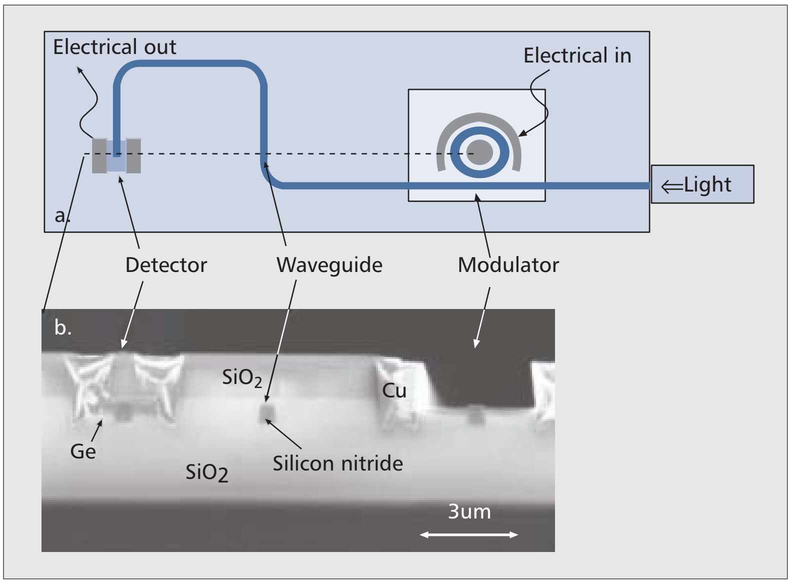


Figure 6. *a) Schematic of top view of full on-die optical link showing bus waveguide connecting modulator to photo detector; b) cross section SEM image (along the dotted line in 6a) showing optical components in one piece of silicon.*

leave the die. The modulated light is coupled off the die through a fiber or waveguide to a receiving chip, where it is coupled through an integrated waveguide into a compact photodetector. The photodetector output current is converted to a full-swing electrical signal by a TIA and an LA.

Monolithic integration of photonics onto the microprocessor will reduce the power and the cost of I/O. Integration reduces the capacitive load on the driver and receiver circuits and leads to higher bandwidth and lower power. Parasitic capacitance is reduced because integration of the circuits and optical devices on the same die removes the bump, package, and ESD capacitance from the signal path. The intrinsic device capacitance of integrated optical components is smaller than the capacitance of discrete alternatives. Static power consumption is reduced because small integrated optical devices do not require termination, in comparison to larger discrete alternatives such as Mach-Zender interferometers which require 50 Ohm termination for high-speed operation. Cost is reduced by decreasing the required number of discrete optical components.

In a photonic CMOS process for integrated optical links, the additional process steps required for photonics must not degrade or interfere with the front-end CMOS transistor performance. Furthermore, the process must allow fabrication of all required optical components on the same die. The optical components in previously demonstrated integrated optical links were fabricated in the front-end of a semiconductor-on-insulator (SOI) CMOS process [3], which constrains the transistor processing. The presented experimental photonic process is based on a silicon nitride single-mode waveguide with silicon dioxide cladding and provides waveguides, electro-optic (EO) polymer ring resonator (RR) modulators [2, 4], and waveguide-embedded metal-semiconductor-metal (MSM) detectors fabricated from polycrystalline germanium on the bulk CMOS (not SOI) process back-end compatible silicon-dioxide dielectric [4, 5].

Fabrication

The photonic elements are added to the CMOS process in the metal interconnect fabrication back-end section of the process after all the high temperature front-end processing of transistors is completed. The waveguides are formed with a 450 nm silicon nitride layer deposited by plasma-enhanced chemical vapor deposition (PECVD) on the SiO2 interlayer dielectric (ILD) and patterned with photolithography and plasma dry etch. This shared waveguide layer is used to build all the waveguides, RRs, and coupling waveguides for the active electro-optic devices. After patterning the waveguides, silicon dioxide cladding is deposited, and three subsequent lithography steps define the detector regions, the electrodes for all active devices, and modulator regions. The photodetector regions are filled with polycrystalline germanium in a damascene process, the detector electrodes are formed in a standard copper damascene process, and then the modulators are formed by depositing EO polymer cladding over the ring resonators in the regions defined for the modulators. The additional cost to add photonic devices to the CMOS process is low, since only four additional photolithography steps are required.

Figure 6 shows both a top view and an SEM cross-section of the modulator, waveguide, and detector constituting a complete optical link. A single patterned silicon nitride layer forms all of the waveguides in the active and passive components. Similarly, one metal layer forms all the electrodes for both the modulator and photodetector. Furthermore, this optical layer is compatible with standard microprocessor CMOS as it is created on an amorphous ILD and can therefore be fabricated in the back-end metal interconnect section of the CMOS process. In order to stay within the thermal budget for standard back-end processing, all steps in the process flow must occur below ~450°C.

Experimental Results

Waveguide — The waveguide is the foundation for the proposed photonic CMOS technology [6]. The waveguide is processed as a 450 nm PECVD silicon nitride film deposited on a 2 μm silicon dioxide undercladding layer at 400°C. The waveguide is patterned using conventional 248 nm lithography and plasma etching. Loss measurements at 1310 nm using the cut-back method show that the silicon nitride waveguide loss is ~1 dB/cm for waveguides with a width of 0.5 μm. This loss is sufficiently low for on-die applications where the total waveguide length is on the order of 1 cm.

Modulator — The electro-optic cladding RR modulator and photodetector share the high index contrast waveguide fabrication process. The modulator design is based on a high-performance ring resonator built with a silicon nitride waveguide and ring. Copper damascene electrodes are fabricated around the ring, and the top cladding is removed and replaced with the EO polymer. This work uses a proprietary chromophor-doped EO polymer [6]. The modulator design is optimized for a quality factor (Q) between 5000 and 10,000: high enough that a small resonance shift

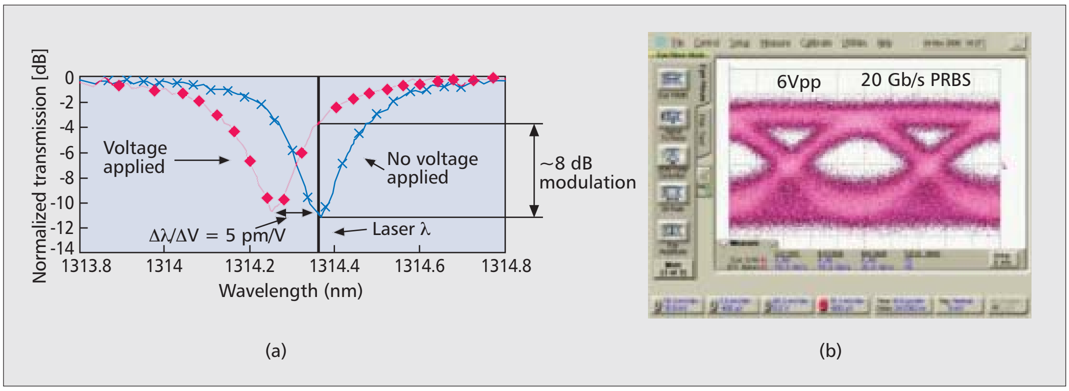


Figure 7. *a) Resonance spectra obtained with –20 V (♦) and 20 V (×) bias on the EO modulator; b) 20 Gb/s PRBS eye diagram of EO polymer modulator.*

Measurements at a higher frequency showed that 40 Gb/s operation is within reach. To improve the noise performance, band-gap engineering can be used to create a Schottky barrier at the metal/germanium contact in order to reduce the dark current further.

results in a large modulation depth, but not so high that the modulator is unable to switch at high data rates. The EO polymer is poled before wafer processing is completed using an electric field of 100 V/cm around the glass temperature of 143°C. The electrodes have a 4.5 μm gap centered around the waveguide ring, and the ring has a radius of 28 μm. An SEM image of the modulator is shown on the right of Fig. 6b. The resonance spectrum of a typical modulator under +20 V and –20 V bias is shown in Fig. 7a. The resonance shift calculated with a linear fit to the resonance frequencies measured at +20 V and –20 V bias is 5 pm/V. The measured Q was ~7000, and the resonance depth was ~11 dB. The highest measured modulation depth for a 10 GHz clock input with a 6 V swing was 8 dB. A 20 Gb/s pseudo-random binary sequence (PRBS) eye diagram for a typical device was measured [4] (Fig. 7b).

Photodetector — Unlike a PIN detector, the lateral MSM detector requires only one lithography step to form the contacts. An evanescently coupled waveguide, shown on the left of Fig. 6b, efficiently couples the light into the absorbing active material of the photodetector. The polycrystalline germanium in the detector was deposited by CVD processing at 600°C. Fabrication of a photodiode from polycrystalline germanium deposited on ILD is an important step toward compatibility with a standard back-end (BE) CMOS process. Measurements at a higher frequency showed that 40 Gb/s operation is within reach. To improve the noise performance, band-gap engineering can be used to create a Schottky barrier at the metal/germanium contact in order to reduce the dark current further. Another important step toward BE compatibility is lowering the process temperature. Devices were fabricated using PVD Ge at 350˚C, and the best measured devices had a dark current of 77 μA with open PRBS eyes at both 20 and 40 Gb/s [4].

Optical Link Modeling and Comparisons

The optical I/O link power efficiency is a strong function of the received optical power, which is determined by the transmit power and the link optical loss budget. A feasible best case value for the hybrid optical link budget is –6.8 dB with some packaging improvements. This is dominated by coupling losses from the VCSEL and photodetector to the multimode fiber (MMF) and the finite extinction ratio penalty. The hybrid optical I/O link budget is calculated using the following assumptions:

Average TX power	**3.0 dBm**
VCSEL to MMF coupling	–1.1 dB
MMF to photodetector coupling	–1.1 dB
Extinction ratio (7.3 dB) penalty	–1.6 dB
Margin	–3.0 dB
Link budget	**–6.8 dB**
Required RX sensitivity	**–3.8 dBm**

The integrated optical link budget is nearly 9 dB worse than the hybrid optical link budget due to the coupling loss between the off-chip single-mode fiber and the on-chip single-mode waveguide, and the extra coupling loss from the off-chip CW laser. However, the integrated photodetector's ultra-low capacitance allows the integrated optical receiver to achieve approximately 13 dB of sensitivity improvement at the same bandwidth, which results in significant system power savings. The integrated optical I/O link budget is calculated using the following assumptions:

Average VCSEL TX power	**3.0 dBm**
Source laser to SMF coupling	–2.0 dB
SMF to modulator coupling	–2.0 dB
Modulator loss	–2.0 dB
Modulator to SMF coupling	–2.0 dB
SMF to photodetector coupling	–3.0 dB
Extinction ratio (8.0 dB) penalty	–1.4 dB
Margin	–3.0 dB
Link budget	**–15.4 dB**
Required RX sensitivity	**–12.4 dB**

Circuit simulation-based power efficiency estimates of both transmit and receive front-end circuits for these two optical I/O architectures was performed for CMOS technologies starting from a 45 nm node and ending with a predicative 16 nm CMOS node [1]. A current-mode VCSEL driver and a simple CMOS inverter-based voltage-mode modulator driver are modeled for the hybrid and integrated optical systems, respectively. In both systems a TIA is followed by simple differential-pair LA stages to realize the optical receiver. The models are constructed with the circuits optimized to provide the minimum bandwidth necessary for a particular data rate, and

The comparison reveals that the hybrid optical architecture is equal to or better in power efficiency than both the electrical backplane channel and the desktop channel at data rates near where RX equalization becomes necessary.

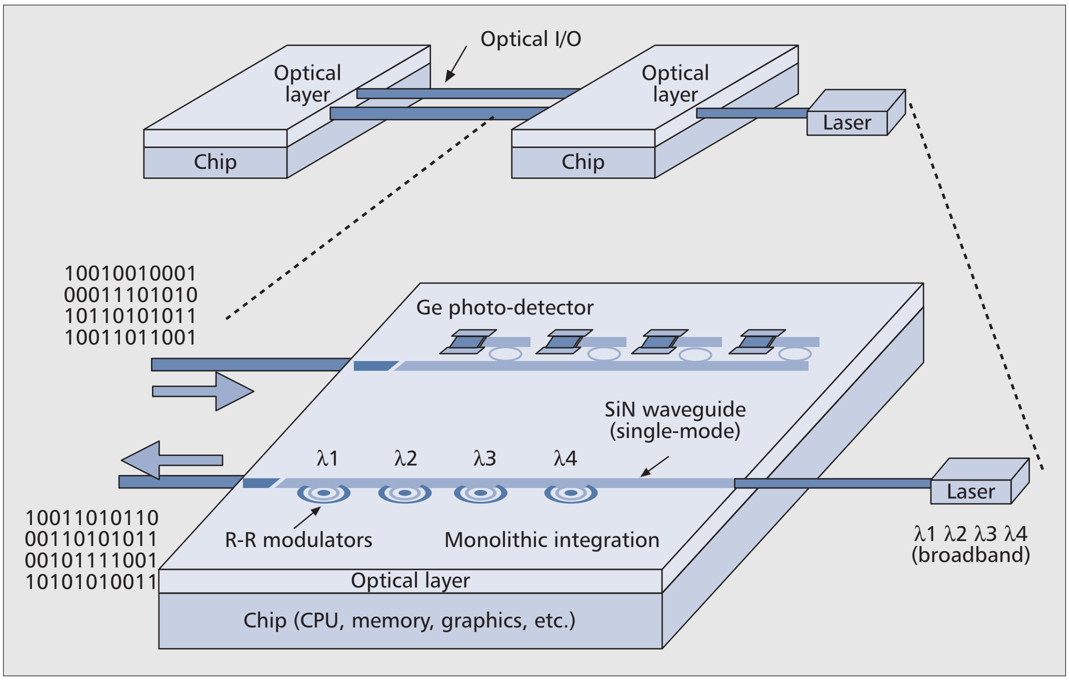


Figure 8. *Photonic CMOS enabled wavelength-division multiplexing architecture.*

thus approximate a power optimal solution. The hybrid optical link power efficiency initially improves as the data rate increases due to the assumed-constant 3 dBm optical power from the 850 nm VCSEL. Power efficiency degrades from the optimum at higher data rates due to the optical RX amplifier gain-bandwidth requirements. As technology scales, this optimum occurs at a higher data rate due to the increased transistor f_T. This analysis predicts that hybrid optical data transmission at less than 1 pJ/b will be realized in the future. Assuming a 1310 nm CW laser source with 3 dBm optical TX power, the integrated optical link power efficiency displays similar behavior at a much lower power level due to low capacitance of the modulator and photodetector allowing for very efficient optical drivers and receivers. Ultra-low receiver input capacitance enables a TIA-based receiver without any LA stages to provide sufficient sensitivity at data rates exceeding 30 Gb/s. The data rate at which extra LA stages become necessary scales with the improved CMOS technology f_T. These projections indicate that photonic CMOS could enable integrated optical interconnect to reach 0.3 pJ/b.

The power-performance analysis of the hybrid optical link was compared with electrical link systems that employ all three electrical channels discussed early in the electrical link analysis. The comparison reveals that the hybrid optical architecture is equal to or better in power efficiency than both the electrical backplane and desktop channels at data rates near where RX equalization becomes necessary. This data rate is dependent on the channel loss characteristics and is 13 Gb/s and 19 Gb/s for the 17 in backplane and 7 in desktop channels, respectively. While the hybrid optical link cannot outperform the high-performance electrical cable channel at the 45 nm node, the increased gain-bandwidth offered by the 16 nm node allows the hybrid optical link to become comparable near 40 Gb/s. Note that this assumes the availability of 40 Gb/s-class VCSELs, which are currently emerging from research [7]. The reduced parasitics offered by the integrated photonics with CMOS optical architecture allows it to achieve superior power efficiency over the majority of data rates compared to the three electrical channels as well as the hybrid optical architecture. This assumes further improvements in modulator EO polymer performance to enable sufficient optical modulation depth at voltage modulation levels compatible with CMOS inverter-based drivers [4, 6].

FUTURE DIRECTIONS

As CMOS scaling continues in the future, larger numbers of CPU cores will be integrated on the microprocessor chip, and it will become necessary to provide interconnect scaling to higher bandwidth between cores on chip, and between these cores and the off-chip DRAM. Wavelength-division multiplexed (WDM) links transmit multiple wavelengths through the same waveguide in order to increase the aggregate optical data transmission. A photonic CMOS architecture for optical WDM of signals monolithically integrated on chip is shown in Fig. 8. The RR modulator selectively modulates a single wavelength from a multiwavelength source, and eliminates the need for separate optical demultiplexers and multiplexers. At the receiver, passive RR optical filters can demultiplex the optical data by selecting a single unique wavelength for detection at each photodetector. Since the photonic CMOS RR modulators have such a narrow tuning range (Fig. 7), the WDM wavelengths can be spaced at less than 1 nm (100 GHz in optical frequency with a reference of 230 THz). Thus, the RR technology provides the means for bandwidth to scale by adding more wavelengths to each waveguide channel [1].

SUMMARY

This work provides a comparison of electrical I/O to optical I/O for chip-to-chip interconnect. While electrical interconnect will continue to use more sophisticated equalization techniques to overcome the loss of the interconnect channel, the high data rate and long interconnect lengths required by future many-core processors will require the introduction of optical interconnect. Optical interconnect for CPUs will first be introduced with optical package-to-package I/O using hybrid MCM single-package technology. In the long term, monolithic integration of optical components will provide terabyte-per-second interconnect data rates with the required energy efficiency at less than 1 pJ/bit.

REFERENCES

[1] I. A. Young *et al.*, "Optical I/O Technology for Tera-Scale Computing," *IEEE JSSC*, Jan. 2010, pp. 235–48.

[2] E. Mohammed *et al.*, "Optical Hybrid Package with an 8-Channel 18GT/s CMOS Transceiver for Chip-to-Chip Optical Interconnect," *Proc. SPIE*, vol. 6899, Feb. 2008.

[3] B. Analui *et al.*, "A Fully Integrated 20-Gb/s Optoelectronic Transceiver Implemented in a Standard 0.13m CMOS SOI Technology," *IEEE JSSC*, Dec. 2006, pp. 2945–55.

[4] I. A. Young *et al.*, "Integration of Nano-Photonic Devices for CMOS Chip-to-Chip Optical I/O," *CLEO/QELS Conf. Digest Tech. Papers*, May 16–21, 2010.

[5] M. R. Reshotko *et al.*, "Waveguide Coupled Ge-on-Oxide Photodetectors for Integrated Optical Links," *IEEE/LEOS Int'l. Conf. Group IV Photonics*, Sept. 17–19, 2008, pp. 182–84.

[6] B. A. Block *et al.*, "Electro-Optic Polymer Cladding Ring Resonator Modulators," *Optics Express*, vol. 16, no. 22, Oct. 2008, pp. 18326–33.

[7] T. Anan *et al.*, "High-Speed 1.1-mm-Range InGaAs VCSELs," *Optical Fiber Communication/National Fiber Optic Engineers Conf.*, Feb. 24–28, 2008, pp. 1–3.

BIOGRAPHIES

IAN. A. YOUNG [M'78, SM'96, F'99] (ian.young@intel.com) is a senior fellow and director of Advanced Circuits and Technology Integration in the Technology and Manufacturing Group at Intel Corporation. He does research and development of mixed-signal circuits for microprocessor and SOC products along with process technology development. He joined Intel in 1983. Starting with the development of circuits for a 1 Mb DRAM, he then led the design of three generations of SRAM products and manufacturing test vehicles, and developed the original PLL-based clocking circuit in a microprocessor while working on the 50 MHz Intel486 processor design. He subsequently developed the core PLL clocking circuit building blocks used in each generation of Intel microprocessors through the 0.13 μm 3.2 GHz Pentium 4. He is currently directing the research and development of analog mixed-signal and RF circuits for high-speed serial I/O in advanced logic process, RF wireless circuits in advanced SOC process, and researching chip-to-chip optical I/O technology. He received his Bachelor's and Master's degrees in rlectrical rngineering from the University of Melbourne, Australia, in 1972 and 1975. He received his Ph.D. in electrical engineering from the University of California, Berkeley in 1978. He was a member of the Symposium on VLSI Circuits Technical Program Committee from 1991 to 1996, serving as Program Committee Chairman in 1995–1996 and Symposium Chairman in 1997–1998. He was a member of the ISSCC Technical Program Committee from 1992 to 2005, serving as Digital Subcommittee Chair from 1997 to 2003, Technical Program Committee Vice-Chair in 2004, and Chair in 2005.

EDRIS MOHAMMED is a research engineer at Intel Corporation responsible for system design, integration, and characterization of hybrid optical interconnect at Intel's components research division. He joined Intel Corporation in 2001. Before joining Intel, he was a member of technical staff at Lucent Technologies where he worked as a test and reliability engineer for 980 nm pump laser diodes. His research interests include high-speed optoelectronic devices, package integration, and passive optical and electrical devices. He received a B.Sc. in physics from Addis Ababa University in 1987, an M.Sc. in physics from Florida A&M University in 1995, and an M.Sc. in electrical engineering (optoelectronics) and a Ph.D. in applied physics from Georgia Institute of Technology in 2000. Since 2005 he has been serving as an associate editor of optical interconnects for *SPIE Optical Engineering Journal*.

JASON T. S. LIAO [M'09] received his M.S. degree in communication engineering from National Chiao Tung University, Taiwan, in 1996 and his Ph.D. degree in electrical engineering from the University of California, San Diego in 2004. In 2005 he joined Intel Corporation, Hillsboro, Oregon, as a senior electrical engineer. His research focuses on high-speed optical I/O and receiver clocking techniques for high-speed electrical I/O links.

ALEXANDRA M. KERN [S'01, M'07] received A.B. and B.E. degrees from Dartmouth College in 2002, and S.M. and Ph.D. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology in 2004 and 2007. In June 2007 she joined the Logic Technology Development group at Intel Corporation, where she is involved in the development of high-speed electrical serial links in advanced CMOS technology. Her research interests include electrical and optical interconnect for on-chip and chip-to-chip applications, and analog/mixed-signal circuit design techniques for deep-submicron CMOS.

SAMUEL PALERMO [S'98, M'07] received B.S. and M.S. degrees in electrical engineering from Texas A&M University, College Station in 1997 and 1999, respectively, and his Ph.D. degree in electrical engineering from Stanford University, California, in 2007. From 1999 to 2000, he was with Texas Instruments, Dallas, where he worked on the design of mixed-signal integrated circuits for high-speed serial data communication. From 2006 to 2008 he was with Intel Corporation, where he worked on high-speed optical and electrical I/O architectures. In 2009 he joined the Electrical and Computer Engineering Department of Texas A&M University, where he is currently an assistant professor. His research interests include high-speed electrical and optical links, clock recovery systems, and techniques for device variability compensation. He is a member of Eta Kappa Nu.

BRUCE A. BLOCK joined Intel in 1998 after receiving his Ph.D. degree in materials science and engineering from Northwestern University in 1997 and a B.S. degree in materials science and engineering from Cornell University in 1989. From 1989 to 1992 he worked for IBM Corporation, East Fishkill, New York, in the Advanced Packaging Laboratory. He has been working since 2000 on the process integration and characterization of CMOS-compatible optical devices in the Components Research Organization. He has developed many passive and active devices based on silicon nitride waveguides, including waveguide coupled germanium photodetectors, electro-optic polymer modulators, ring resonator filters, and surface plasmon polariton-based polarizers and detectors.

MIRIAM R. RESHOTKO [M'07] joined Intel's Components Research department in 2001 after completing a Ph.D. degree in physics at the Hebrew University of Jerusalem, Israel. She received her M.Sc. degree in physics from the same university, and her B.A. in physics from Cornell University. Her current research centers on integrated CMOS compatible optoelectronic devices for optical interconnects. Specific areas of focus are development of monolithically integrated high-speed waveguide coupled photodetectors for various interconnect applications, and process integration for discrete optical devices, including waveguides, modulators, photodetectors, and optical links.

PETER L. D. CHANG [M'92] received his B.S. degree in physics from National Taiwan University in 1977 and his Ph.D. in theoretical solid state physics from the State University of New York at Stony Brook in 1985. From 1985 to 1991 he conducted research in the Physics and ECE Departments of the University of California, Santa Barbara. From 1991 to 1994 he was with Lockheed Sanders working on MMIC. He joined Intel in 1994, developing 0.25 μm flash memory technology. Sine 1996 he has been involved in R&D and high volume manufacturing of logic and SRAM process. He is currently with the Components Research of Technology, Manufacturing and Enterprise Services Group. His research interests are in high-density memory and optical interconnects for high memory bandwidth required for future generations of CPUs.

Optical interconnect for CPUs will first be introduced with optical package-to-package I/O using hybrid MCM single-package technology. In the long term, monolithic integration of optical components will provide TB/s interconnect data rates with the required energy efficiency at less than 1 pJ/bit.

Communications | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | qMags

Wireless Proximity Interfaces with a Pulse-Based Inductive Coupling Technique

Hiroki Ishikuro and Tadahiro Kuroda, Keio University

ABSTRACT

The rapid performance progress in processors and memory cores by technology scaling requires further improvement in interface bandwidth. However, interface bandwidth is not keeping up with the processing speed of the core and is becoming a bottleneck in system performance. To fill the performance gap, wideband low-power low-cost interfaces are strongly demanded. A wireless proximity interface that uses inductive coupling is one such interface expected to be used for interchip links in high-performance 3D system integration. Inductive coupling interfaces use the magnetic near-field induced by micro-coils. The coils (channels) can be arranged in a dense array because magnetic near-field localizes in the proximity of each coil, and crosstalk between the channels is small. Therefore, inductive coupling interfaces are suitable for wideband low-cost proximity communication. An inductive coupling interface can also realize highly reliable communication with low power consumption. Evaluation systems developed to study the performance of inductive coupling interfaces have demonstrated the feasibility of the interfaces in a wide range of applications.

INTRODUCTION

Recent progress in portable equipment demands further improvement of bus bandwidth between the processor and memory. In a conventional system on a board, the performance of the processor core has been improved by 70 percent per year. On the other hand, the improvement of input/output (I/O) bandwidth has stayed at 28 percent per year. If the technology is scaled by a factor of k, the number of transistors increases by a factor of $1/k^2$, whereas the amount of I/O increases only by a factor of $1/k$. As a result, the performance gap between the bus bandwidth and processor core increases as the technology is scaled.

3D system integration, such as system in a package (SiP) and package on package, is attracting much attention. 3D system integration improves I/O bandwidth and reduces power consumption to some extent because it eliminates the parasitic load of the bus trace on a printed circuit board (PCB). Commercially available 3D chip stacking, for example, in a flush memory package utilizes bonding wires for chip-to-chip interconnects. In this case the parasitic inductance of bonding wires and parasitic capacitance of the electronic static discharge (ESD) protection device determine the maximum data rate and minimum power consumption. Furthermore, the I/O pads and buffers should be placed on the periphery of the chip for the wire bonding. Therefore, the amount of I/O increases only by a factor of $1/k$.

The next step to improve I/O bandwidth and reduce power consumption is to realize area array interfaces for 3D system integration. Area array interfaces place the I/O channels in a 2D array. The amount of I/O in the area array interface increases by a factor of $1/k^2$. Therefore, the area array interfaces become a solution to fill the performance gap between the I/O bandwidth and the chip core. Through silicon via (TSV) is one of the candidates for the area array interfaces in 3D system integration and has been intensively studied [1, 2]. However, an additional fabrication process is required to form the TSV, which increases the chip cost. Furthermore, the connection relies on mechanical contact, which brings a reliability problem as the amount of I/O increases. Therefore, the application of TSV technology is limited to some area such as an image sensor. To expand the use of 3D system integration, another approach to realize the wideband low-power low-cost area array interface is essential.

The goal of our group is to develop such a wideband low-power low-cost area array interface. For this purpose, we chose to use a pulse-based inductive coupling technique with micro-coils for the chip-to-chip interface [3]. Inductive coupling using magnetic near-field is most suitable for wireless area array interfaces. The strong localization of the magnetic field in an inductive coupling channel makes it possible to arrange a dense I/O array [4]. The interface eliminates an ESD protection device and can reduce parasitic capacitance, which is favorable for the improvement of bandwidth and reduction of power consumption. Furthermore, inductive

0163-6804/10/$25.00 © 2010 IEEE

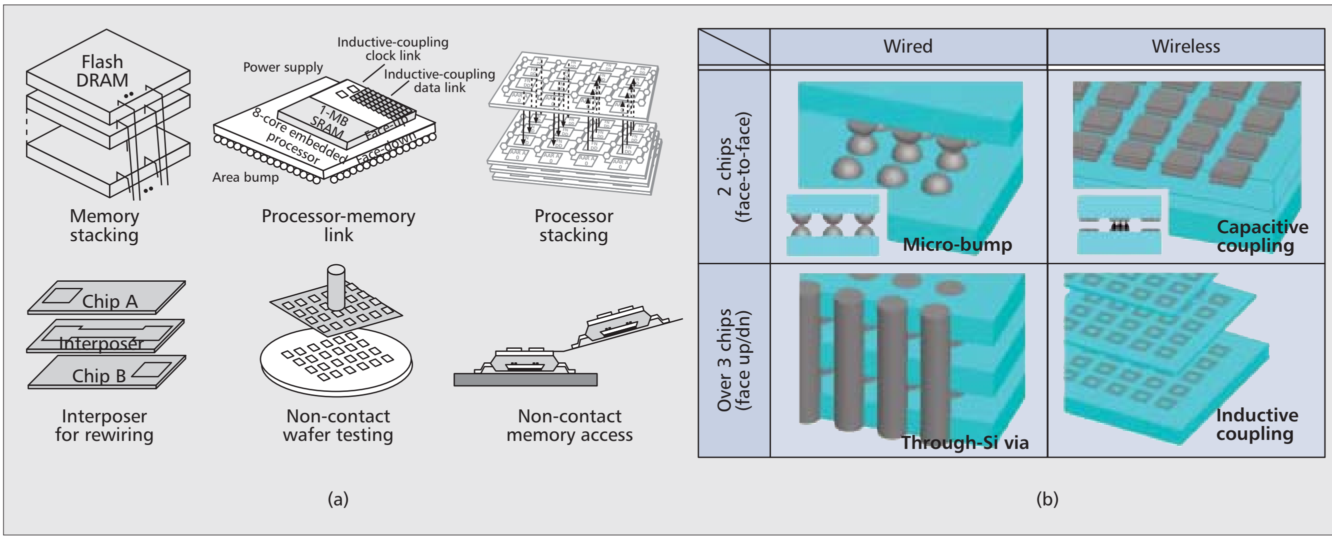


Figure 1. *a) Applications of the pulse-based inductive coupling wireless interface; b) area array interfaces for 3D system integration.*

coupling interfaces are cost effective because they do not require an additional fabrication process, and connections do not rely on mechanical contacts, which improve the yield. The pulse-based technique enables the use of simple small transceiver circuits, which is also favorable for low-power and low-cost implementation.

In this article pulse-based inductive coupling wireless interfaces for 3D system integration are introduced. In the next section various area array interfaces are compared, and the features of inductive coupling using magnetic near-field are studied. The signaling and transceiver circuits of inductive coupling interfaces are then explained. After that, two applications of inductive coupling interfaces developed by our group are introduced.

Comparison of Area Array Interfaces

The application of the inductive coupling interface will cover a wide area. Figure 1a shows an example. 3D system integration such as memory stacking [5], processor to memory link [6], and processor stacking [7] will bring high-performance small systems. A wireless interposer [8] will bring flexibility to chip assembly. Non-contact wafer test [9] will make it possible to test many chips with massive I/O counts and reduce the test cost. Wireless interface through a large-scale integration (LSI) package can be used for package-on-package, non-contact memory card [10], and non-contact bus probing for software debugging [11].

Area array interfaces for 3D system integration can be categorized as shown in Fig. 1b. Micro-bump [12] and TSV [1, 2] are wired interfaces. Capacitive coupling [13] and inductive coupling links are wireless interfaces. The micro-bump and capacitive coupling link can only be applied to face-to-face chip stacking, and the number of chips is limited to two. On the other hand, TSV and the inductive coupling link can be applied to both the face-up face-down chip stacking, and more than three chips can be stacked.

The micro-bump and TSV require additional fabrication processes, which increases the chip cost. Furthermore, the wired interfaces rely on mechanical contact, and the yield of chip stacking decreases as the amount of I/O increases. On the other hand, a wireless interface can be fabricated by a standard digital complementary metal oxide semiconductor (CMOS) process. The wireless interface does not require mechanical contact and is free from contact failure. Furthermore, wireless interfaces have advantages in bandwidth and power consumption because they do not require an ESD protection device, which has large parasitic capacitance. Another advantage of wireless interfaces is that they can be used through LSI package communications and realize detachable interfaces.

Inductive coupling is more suitable as a wireless interface than capacitive coupling [14]. For example, assume that the wafer thickness is 60 μm, which is typical for 3D chip stacking. In the capacitive coupling link the coupling coefficient between the transmitter and receiver begins to decrease as the resistivity of the Si substrate becomes less than 10 Ωcm and decreases to 1/10 when the resistivity decreases to 1 Ωcm. In the case of the inductive coupling link, the coupling between the transmitter and receiver does not degrade if the substrate resistivity is higher than 10^{-2} Ωcm. The charge supplied by displacement current under the metal plate shields the electric field in the capacitive coupling link, whereas the eddy current shields the magnetic field in the inductive coupling link. In the case of capacitive coupling, the parasitic capacitance between the metal plate and Si-substrate is on the order of 10 fF. If the signal frequency is 1 GHz, the impedance of this parasitic capacitance becomes 17 kΩ. Considering that the signal amplitude is around 1 V, displacement current of several times 10 μA can supply the charge to shield the electric field. On the other hand, the inductance and series resistance of the coil in an inductive coupling link are 10 nH and several times 10 Ω,

Communications | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | qMags

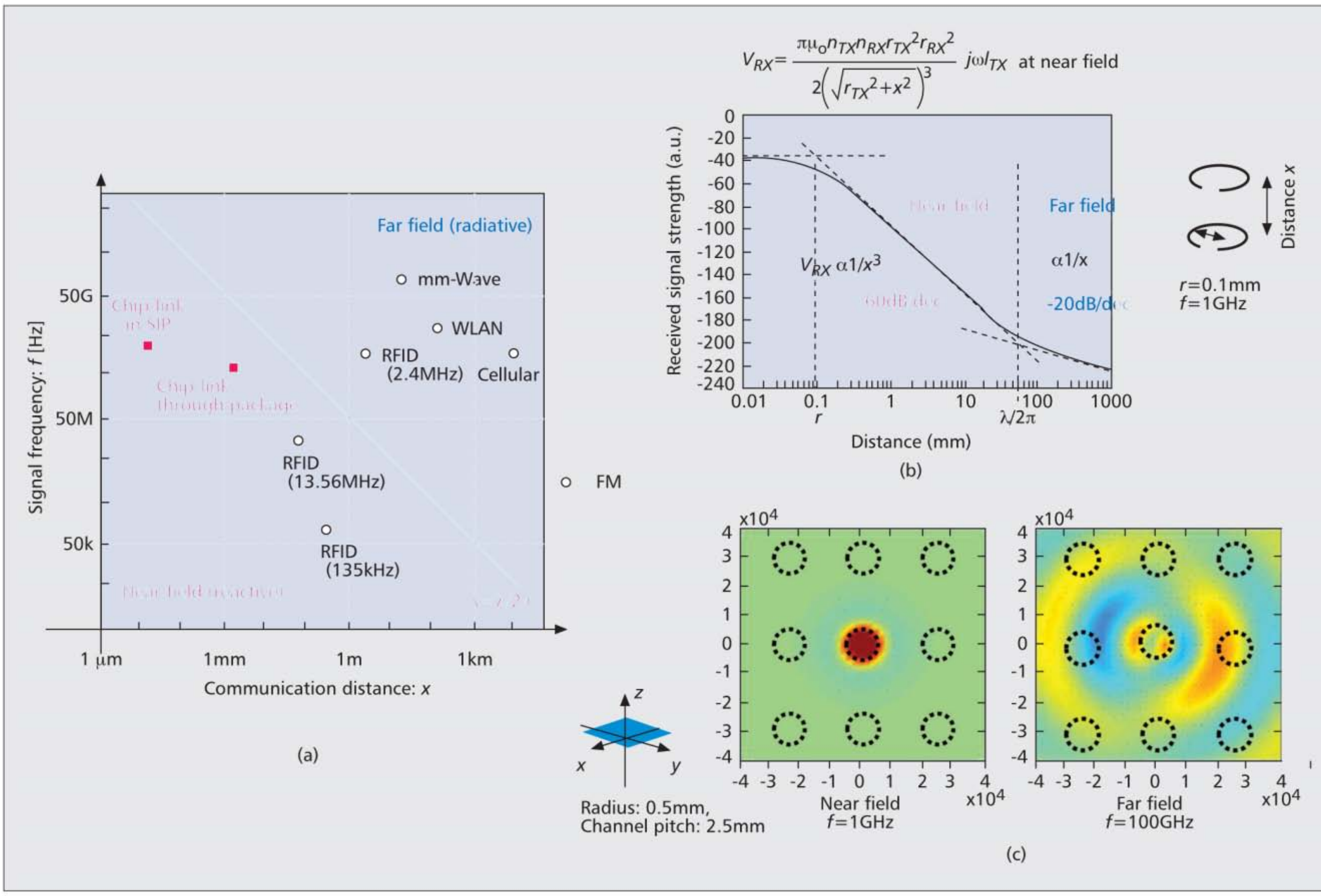


Figure 2. *a) Classification of wireless standards into near-field and far-field; b) received signal strength as a function of distance from transmitter; c) magnetic field strength of near-field and far-field.*

respectively. The impedance of the coil becomes around 100 Ω, and signal current in the transmitter coil is in milli-amperes. Therefore, the eddy current required to shield the magnetic field in the inductive coupling link is several milli-amperes. This is why the shielding effects differ between capacitive and inductive coupling. Since the usual Si substrate resistivity is several to several tens of ohm-centimeters, a capacitive coupling link can only be applied to face-to-face chip stacking, whereas the inductive coupling link can be applied to both face-up and face-down chip stacking.

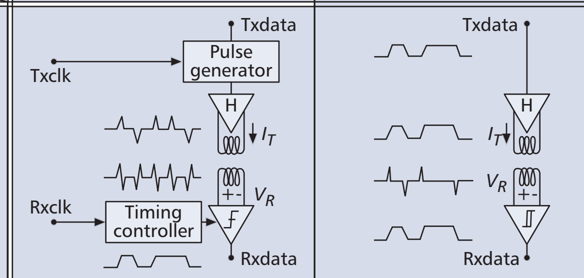
Wireless Interfaces Using Magnetic Near-Field

Figure 2a shows the relation between signal or carrier frequency and communication distance of several wireless standards. The electromagnetic field within λ/2π of a transmitter is called near-field, where λ is the wavelength. The electromagnetic field beyond λ/2π from a transmitter is called far-field or radiative field. In the figure the wireless standards located left below the line $x = \lambda/2\pi$ use near-field, whereas the standards located to the upper right of the line use far-field. For example, wireless LAN and cellular phone use far-field, and a contactless IC card (RF-ID) uses near-field. Our wireless chip-to-chip interfaces are near-field communication. The communication distance ranges from several tens to several millimeters. The signal frequency ranges from several to several tens of gigahertz.

Figure 2b shows the relation between the magnetic field and distance from transmitter. The radius of a transmitter coil is 100 mm, and signal frequency is 1 GHz. In this example the boundary between near-field and far-field is 50 mm. The figure indicates that the magnetic field rapidly decreases as the distance exceeds the size of the transmitter coil. In this region the magnetic field changes in proportion to $1/x^3$. If the distance exceeds 50 mm, the relation between the field strength and distance changes to $1/x$. The strong dependence of near-field strength on distance becomes a disadvantage for long-distance communications. However, in proximity communication it provides an advantage because one can place many channels thanks to the small crosstalk between channels.

Figure 3c illustrates the magnetic field strength in the same plane of the transmitter coil. The transmitter coil is indicated as a dashed circle. In this simulation application of an interface through-LSI-package is assumed, and the coil size is determined. The communication distance is 1 mm. The radius and pitch of the coil are 0.5 mm and 2.5 mm, respectively. The left figure shows near-field communication when the signal frequency is 1 GHz and coil size is much smaller than λ/2π. The right figure shows far-field communication when the signal frequency is 100 GHz and coil size is equal to $\lambda/2\pi$. From these figures, it is clearly shown that the magnetic field localized in the proximity of the transmitter coil in the case of near-field communication. Therefore, many transmitter coils can be densely arranged to form an area array interface. On the other hand, the magnetic field travels long distance in the case of far-field communication. In this case it is difficult to form an area array interface because of the crosstalk between each channel.

In the near-field communications, if the size of the coil (D) is assumed to be three times larger than communication distance (x), the required channel pitches in the 1D and 2D arrays are $2 \times D$ and $3 \times D$, respectively, to establish a reliable communication (bit error ratio [BER] less than 10^{-14}) [15].



	(a)	(b)
Circuit topology		
Data rate	1 Gb/s	11 Gb/s
Energy/bit	0.4 pJ/b	1.4 pJ/b
Timing control	Needed	Not needed
Coil diameter	2 × communication distance	3 × communication distance

in 180 nm CMOS

Figure 3. *Transceiver circuit of pulse-based inductive-coupling interfaces: a) synchronous scheme; b) asynchronous scheme.*

Transceiver Circuits for Pulse-Based Inductive Coupling Interfaces

For inductive coupling interfaces, we adopted a pulse-based data transfer scheme. The pulse-based scheme eliminates the local oscillator and makes it possible to use a simple small circuit with low-power operation. The transceiver architecture can be classified into synchronous and asynchronous schemes (Fig. 3).

In the synchronous scheme (Fig. 3a) the current pulses are generated by a transmitter circuit and fed into the transmitter coil. The polarity of the pulse is determined by the transmitting binary data. That is, positive pulse is generated if the transmitting data is 1, and negative pulse is generated if the data is 0. On the receiver side, a time derivative of the transmitted pulse waveform is induced at the receiver coil as a voltage signal. The received signal has double pulse. The polarity of the front peak is decided by the clocked comparator at the rising edge of the timing clock and then recovered to non-return-to-zero (NRZ) data. This clock is supplied from a dedicated clock channel or clock recovery circuit. Since both the transmitter and receiver (clocked comparator) consume power only at the edge of the clock, the transceiver operates at low power. Furthermore, the receiver is activated only at the clock rising edge, so the influence of the noise is small. However, precise timing control is required to decide the polarity of the received short pulses.

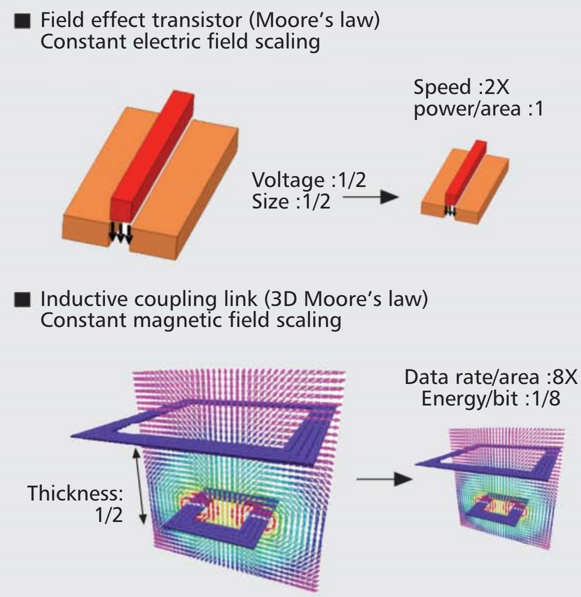
In the asynchronous scheme (Fig. 3b) the current signal whose waveform is the same as the transmitting NRZ data is fed into the transmitter coil. At the receiver side, the pulse signal is induced at the transition point of the transmitted data. The original NRZ data can be recovered by using a hysteresis comparator. Since the data can be recovered asynchronously, the precise clock timing control for symbol detection is not required. On the other hand, the receiver circuit is always active and can be disturbed by the noise. The power consumption becomes large compared with the synchronous receiver.

The size, bandwidth, and power consumption of the inductive coupling transceiver can be scaled as the device technology is scaled. Just like the constant electric field scaling plays an important role in the scaling of a metal oxide semiconductor field effect transistor (MOSFET), constant magnetic field scaling is important for scaling of the inductive coupling transceiver. The scaling scenario is summarized in Fig. 4. In magnetic field scaling both device size and supply voltage are scaled to $1/\alpha$. Chip thickness should be scaled to $1/\alpha$ and the coil turn number increased by $\alpha^{0.8}$. Not only the electric field of FET but also the magnetic field of the inductive coupling channel can be kept constant. The operation of both the MOSFET and inductive coupling channels is ensured to be constant before and after the scaling. As a result, data rate can be increased by α, channel number is increased by α^2, aggregated data rate per area is increased by α^3, and energy per bit is reduced to $1/\alpha^3$. In some technologies the top metal layer for global wiring or number of metal layers may not be scaled at the same ratio as the other parameters. In this case the coil turn number will not increase at the rate of $\alpha^{0.8}$, and the inductance of the coil will decrease. To keep the received signal strength, the scaling of signal current will become a value between $1/\alpha$ and 1. Therefore, improvement of energy per bit falls between $1/\alpha^2$ and $1/\alpha^3$. Thinning a chip has not been challenged before simply because there was no demand. Good opportunities will be found if this scaling scenario motivates technology development. In *International Technology Roadmap for Semiconductors 2007* (ITRS2007) the chip thickness of for system-in-a-package (SiP) application is expected to scale at the same factor as device scaling.

Applications of Pulse-Based Inductive Coupling Interfaces

As described earlier, the inductive coupling interface can be applied to many applications. In this section two such applications are intro-

There is a trade-off between power dissipation and timing margin. Since power dissipation in a transmitter is in proportion to the square of the pulse width, the narrower the pulse is, the smaller the power dissipation becomes.



Transistor size	[x]	$1/\alpha$
Power supply voltage	[V]	$1/\alpha$
Chip thickness	[T]	$1/\alpha$
Coil turn number (layer #)	[n]	$\alpha^{0.8}$
Current	[I]	$1/\alpha$
Circuit delay time	$[t]\sim[CV/I]$	$1/\alpha$
Coil diameter	$[D]\sim[1/x]$	$1/\alpha$
Self inductance	$[L]\sim[n^2D^{1.6}]$	1
Magnetic coupling coefficient	[k]	1
Receive signal	$[V_r]\sim[kn^2D^{1.6}(I/t)]$	1
Crosstalk	$[V_{RS}/V_{RN}]$	1
Data rate / channel	[1/f]	α
Channel number / area	$[1/D^2]$	α^2
Aggregated data rate / area	$[1/tD^2]$	α^3
Energy / bit	[ItV]	$1/\alpha^3$

Figure 4. *The magnetic field scaling scenario.*

duced. The first one is the stacking of processor and memory [6]. The second one is a detachable through-LSI-package interface [11].

Figure 5a shows the developed evaluation system, and die photographs of the processor and memory stacking with an inductive coupling interface. The performance bottleneck by the bus bandwidth between the processor and memory can be resolved by the wideband wireless interface. The processor chip and memory chip were fabricated in their optimal process to reduce the cost. The processor chip was designed and fabricated in a 90 nm CMOS process. The chip has eight cores and was flip-chip mounted on the package by using area bump. The 1 Mbyte SRAM chip was fabricated using a 65 nm CMOS process and stacked on the processor chip in a face-up manner. The power is supplied to the memory chip by bonded wire, and data communication between the processor chip and memory chip is performed by an inductive coupling wireless interface. The thickness of both chips is 50 μm. The size of the coil is 120 μm, and the channel pitch is 243 μm. Total layout area for the inductive coupling link is 2.82 mm^2.

The block diagram of the processor and memory link system is shown in Fig. 5b. Both the uplink and downlink buses have 16 data channels and 1 clock channel. The bandwidth of each data channel is 600 Mb/s. Therefore, total bus bandwidth is 19.2 Gb/s (0.6 Gb/s × 16 channels × 2). Area normalized by bandwidth is 0.15 mm^2/Gb/s, which is 1/3 that of a conventional DDR2 interface in the same technology.

In this system a synchronous data detection scheme was adopted to reduce power consumption. There is a trade-off between power dissipation and timing margin. Since power dissipation in a transmitter is in proportion to the square of the pulse width, the narrower the pulse, the smaller the power dissipation. The timing margin for sampling the narrow pulse, however, will be reduced. Low-power design requires precise timing control of the clock signal for highly reliable data communication. Adaptive circuits and systems are required to adjust the timing for the following reasons:

- Timing jitter caused by process, power supply voltage, and temperature (PVT) variations, especially in a clock path with long latency caused by global process variations
- Interchannel skews caused by local process variation and layout parasitics

The timing jitter under PVT variations can be monitored and calibrated by a coarse timing control unit. The interchannel deskew can be performed by a fine timing control unit that is implemented in each channel. The timing adjustment controlled by the processor is carried out in the following sequence. First, the control register sets a loopback path in the SRAM for a test mode (an SRAM through mode). Second, pass/fail information, much like a shmoo plot, is stored in a register for both the uplink and downlink by changing the coarse timing. Third, the coarse timing is set such that the timing margin becomes the largest when all the channels pass. For each channel, fine timing is tuned next such that the timing margin becomes the largest.

Figure 5c shows the measured shmoo plot and bathtub curve. After optimizing the timing by setting the control register at the center of the shmoo plot, tolerance against VDD and temperature changes was measured. In the measurement, no single bit failed under ±5 percent VDD variations, and temperature ranges from 25°C to 55°C were observed. The power efficiency is 1 pJ/b, which is 1/30 that of the conventional DDR2 interface. The VDD tolerance can be improved from ±5 to ±10 percent by widening the pulse width from 180 to 320 ps at a cost of an increase in power efficiency from 1 to 2.5 pJ/b (still 1/12 of DDR2).

Figure 6a shows the developed detachable wireless interface. This interface aims at proces-

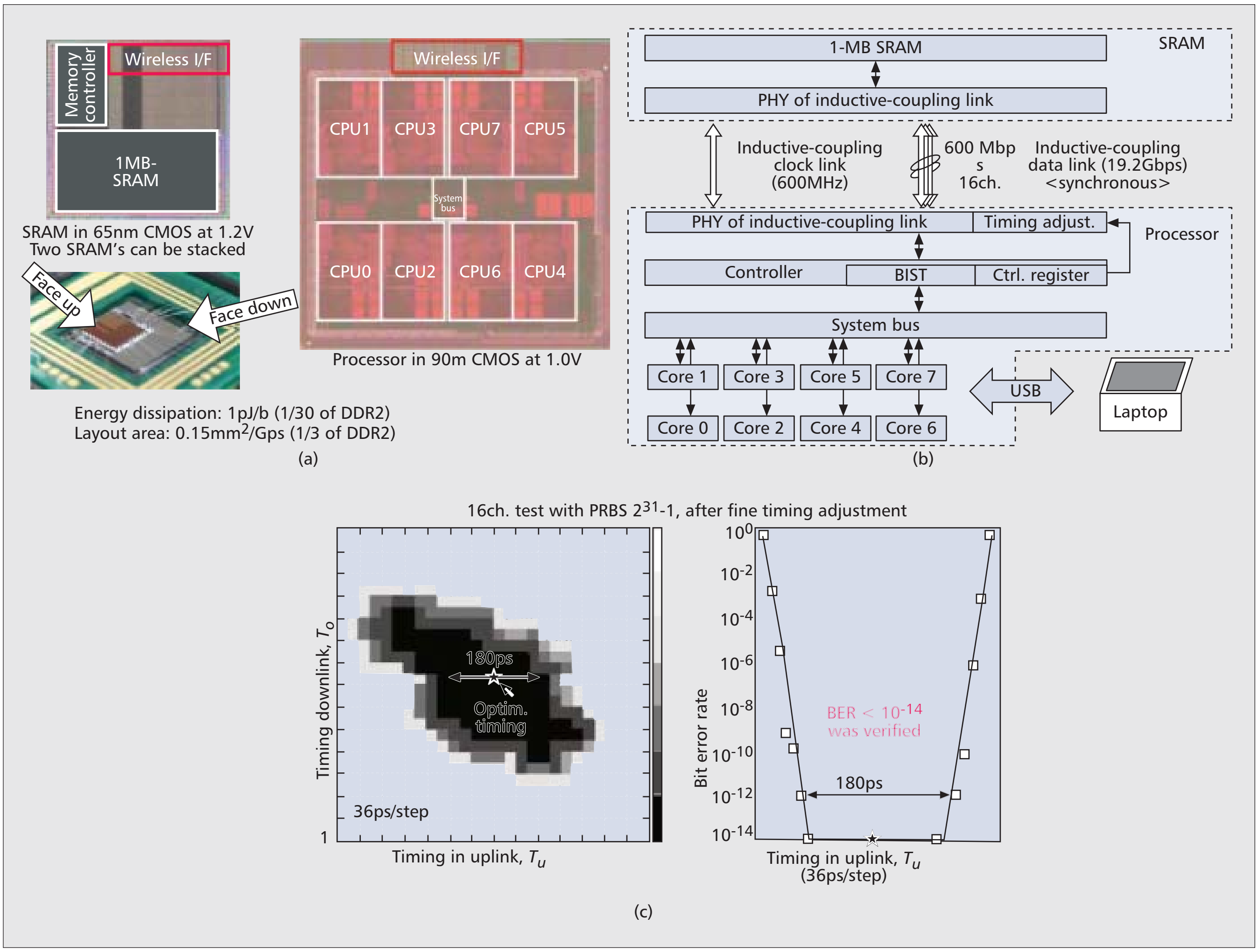


Figure 5. *a) Test chip for processor and cache memory link; b) block diagram of processor and cache memory link; c) measured results of shmoo plot and bathtub curve.*

sor bus monitoring for firmware debugging. The wireless probe consists of inductors that are patterned in the flexible circuit board (FCB) and a transceiver IC. Combination of the FCB inductors and the probe IC is cost effective because the inductor pattern can be optimally customized according to the target LSI, whereas the probe IC can be designed for general use. The wireless probe can be controlled by a PC via an in-circuit emulator (ICE) with a USB 2.0 interface. In this study an MCU chip housed in a 1 mm thick SSOP package was chosen as the target LSI. The target LSI also contains a pulse transceiver with on-chip inductors formed by the top metal layer. The MCU core and the transceiver part are designed in a 0.25 μm CMOS process.

Figure 6b shows a block diagram of the pulse-based transceiver. An asynchronous channel for clock link and synchronous channels for full-duplex data up/downlinks can be established. The system clock of the target LSI is transmitted to the wireless probe, recovered by a hysteresis comparator, and then used for data synchronization in the probe IC. The hysteresis level can be controlled by changing the bias current of a latch. The clock skew between the transmitter and the receiver is adjusted by a controllable delay line. Since both the target LSI and probe IC use the same clock edge, the transmitted data can be robustly recovered even if the system clock has large jitter.

In pulse-based inductive coupling communication, trade-offs between the communication range and data rate should be considered. The communication range can be extended by increasing the inductor size. The larger inductor, however, requires a longer pulse, because the larger inductor has lower self-resonant frequency. Since the maximum attainable data rate is in inverse proportion to the pulse width, the larger inductor has a lower data rate. Furthermore, the probing system should have high alignment tolerance for easy handling. In this study the inductors in the FCB and target LSI are 1.0 mm and 0.6 mm each side, respectively. The self-resonant frequency is 2 GHz. Taking that into account, the pulse width is set to 1 ns.

To extend the communication range under the size limitation of the inductor, received signals are amplified by 30 dB using a 2-stage complementary differential amplifier prior to the comparator. DC offsets of the amplifier and comparator cause a serious problem during sig-

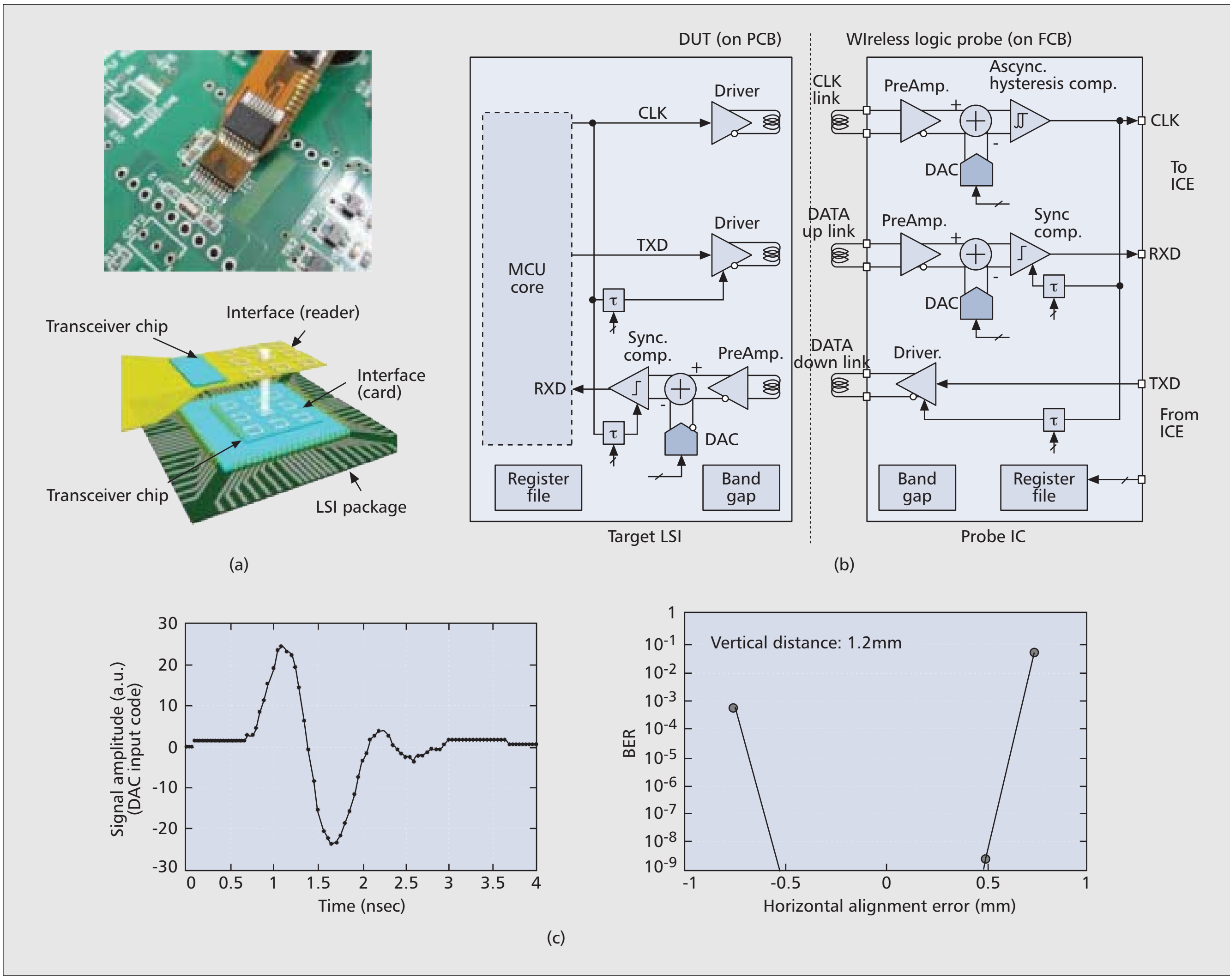


Figure 6. *a) Test system of through LSI package interface; b) block diagram of through LSI package interface; c) measured results of received waveform and alignment tolerance.*

nal detection. For compatibility with the standard CMOS process, a metal-insulator-metal (MIM) capacitor that has an excellent frequency characteristic is not available for DC offset elimination. In this receiver a 6-bit current steering digital-to-analog converter (DAC) is used for offset cancellation. During offset calibration, the transmitter is powered off, and the comparator in the clock receiver operates without hysteresis. The DAC setting for the offsets canceling can be determined by scanning the DAC input value and detecting the point where the polarity of the comparator output changes.

In Fig. 6c measured received pulse waveform and alignment tolerance are shown. From this figure it can be seen that received pulse waveform has double peak pulse with small ringing. The alignment tolerance of 0.5 mm allows manual attachment of the probe.

Since this application does not require high bandwidth, the data rate of the interface is 20 Mb/s. Recently developed, our detachable wireless interface [10], aimed at a non-contact memory card, has achieved 2.5 Gb/s with power consumption of 15 mW (power efficiency is 6 pJ/bit).

Conclusion

A wireless interface that uses inductive -coupling is suitable for a wideband low-power low-cost area array interface. The crosstalk between channels is small because it uses a magnetic near-field whose field strength rapidly decays. The channel pitch can be reduced to three times the coil size, and a dense array of I/O channels can be realized easily. Pulse-based communication makes it possible to use simple small transceiver circuits. Inductive coupling is compatible with device scaling.

Low-power operation (1 pJ/bit) with a small interface (1/30 the size of a conventional DDR2 interface) has been demonstrated by the developed evaluation system of processor and memory stacking. Another application for the chip access through LSI package was demonstrated. The bandwidth is on the order of a gigabit per second at communication distances longer than 1 mm.

References

[1] U. Kang *et al.*, "8Gb 3D DDR3 DRAM Using Through-Silicon-via Technology," *IEEE ISSCC Dig. Tech. Papers*, Feb. 2009, pp. 130–31.

[2] H. Yoshikawa *et al.*, "Chip Scale Camera Module (CSCM) using Through-Silicon-Via (TSV)," *IEEE ISSCC Dig. Tech. Papers*, Feb. 2009, pp. 476–77.
[3] D. Mizoguchi *et al.*, "A 1.2 Gb/s/pin Wireless Superconnect Based on Inductive Inter-Chip Signaling (IIS)," *IEEE ISSCC Dig. Tech. Papers*, Feb. 2004, pp. 142–43.
[4] N. Miura *et al.*, "A 1 Tb/s 3 W Inductive-Coupling Transceiver for 3D-Stacked Inter-Chip Clock and Data Link," *IEEE J. Solid-State Circuits*, vol. 42, no. 1, Jan. 2007, pp. 111–22.
[5] M. Saito, N. Miura, and T. Kuroda, "A 2 Gb/s 1.8pJ/b/chip Inductive-Coupling Through-Chip Bus for 128-Die NAND-Flash Memory Stacking," *IEEE ISSCC Dig. Tech. Papers*, Feb. 2010, pp. 440–41.
[6] K. Niitsu *et al.*, "An Inductive-Coupling Link for 3D Integration of a 90nm CMOS Processor and a 65nm CMOS SRAM," *IEEE ISSCC Dig. Tech. Papers*, Feb. 2009, pp. 480–81.
[7] Y.Kohama *et al.*, "A Scalable 3D Processor by Homogeneous Chip Stacking with Inductive-Coupling Link," *Symp. VLSI Circuits Dig. Tech. Papers*, June 2009, pp. 94–95.
[8] S. Kawai, H. Ishikuro, and T. Kuroda, "A 4.7 Gb/s Inductive Coupling Interposer with Dual Mode Modem," *Symp. VLSI Circuits Dig. Tech. Papers*, June 2009, pp. 92–93.
[9] Y. Yoshida *et al.*, "Wireless DC Voltage Transmission using Inductive-Coupling Channel for Highly-Parallel Wafer-Level Testing," *IEEE ISSCC Dig. Tech. Papers*, Feb. 2009, pp. 470–71.
[10] S. Kawai, H. Ishikuro, and T. Kuroda, "A 2.5 Gb/s/ch 4PAM Inductive-Coupling Transceiver for Non-Contact Memory Card," *IEEE ISSCC Dig. Tech. Papers*, Feb. 2010, pp. 264–65.
[11] H. Ishikuro, T. Sugahara, and T. Kuroda, "An Attachable Wireless Chip Access Interface for Arbitrary Data Rate Using Pulse-Based Inductive-Coupling through LSI Package," *IEEE ISSCC Dig. Tech. Papers*, Feb. 2007, pp. 360–61.
[12] T. Ezaki *et al.*, "A 160 Gb/s Interface Design Configuration for Multichip LSI," *IEEE ISSCC Dig. Tech. Papers*, Feb. 2004, pp. 140–41.
[13] A. Fazzi *et al.*, "3-D Capacitive Interconnections for Wafer-Level and Die-Level Assembly," *IEEE J. Solid-State Circuits*, vol. 42, no. 10, Oct. 2007, pp. 2270–82.
[14] H. Ishikuro, N. Miura, and T. Kuroda, "Wideband Inductive-coupling Interface for High-Performance Portable System," *IEEE CICC '07*, Sept. 2007, pp. 13–20.
[15] N. Miura *et al.*, "Cross Talk Countermeasures in Inductive Inter-Chip Wireless Superconnect," *Proc. IEEE CICC '04*, Oct. 2004, pp. 99–102.

BIOGRAPHIES

HIROKI ISHIKURO (ishikuro@elec.keio.ac.jp) received his B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Tokyo, Japan, in 1994, 1996, and 1999, respectively. In 1999 he joined the System LSI Research and Development Center, Toshiba Corp., Kawasaki, Japan, where he was involved in the development of CMOS RF and mixed-signal circuits for wireless interface chips. In 2006 he joined the Department of Electrical Engineering at Keio University as an assistant professor and started research on high-speed inductive coupling links for 3D chip integration and non-contact connectors. He is currently an associate professor at Keio University, and focuses on mixed-signal circuit and system designs for extremely low-power wireless interfaces. He is a member of the Technical Program Committee for the Symposium on VLSI Circuits.

TADAHIRO KURODA [M'88, SM'00, F'06] received his Ph.D. degree in electrical engineering from the University of Tokyo in 1999. In 1982 he joined Toshiba Corporation, where he designed CMOS SRAMs, gate arrays, and standard cells. From 1988 to 1990 he was a visiting scholar at the University of California, Berkeley, where he conducted research in the field of VLSI CAD. In 1990 he returned to Toshiba, and engaged in the research and development of BiCMOS ASICs, ECL gate arrays, high-speed CMOS LSIs for telecommunications, and low-power CMOS LSIs for multimedia and mobile applications. He invented a variable threshold-voltage CMOS (VTCMOS) technology to control VTH through substrate bias, and applied it to a DCT core processor and a gate array in 1995. He also developed a variable supply-voltage scheme using an embedded DC-DC converter, and employed it in a microprocessor core and an MPEG-4 chip for the first time in the world in 1997. In 2000 he moved to Keio University, where he has been a professor since 2002. He has been a visiting professor at Hiroshima University, Japan, and the University of California, Berkeley. His research interests include low-power high-speed CMOS design for wireless and wireline communications, human computer interactions, and ubiquitous electronics. He has published more than 200 technical publications, including 60 invited papers and 21 books/chapters, and has filed more than 100 patents. He served as General Chairman of the Symposium on VLSI Circuits, Vice Chairman for ASP-DAC, Subcommittee Chair for A-SSCC, ICCAD, and SSDM, and program committee member for ISSCC, Symposium on VLSI Circuits, CICC, DAC, ASP-DAC, ISLPED, SSDM, ISQED, and other international conferences. He is a recipient of the 2005 P&I Patent of the Year Award, the 2006 LSI IP Design Award, the 2007 ASP-DAC Best Design Award, and the 2009 IEICE Achievement Award. He is an elected AdCom member for the IEEE Solid-State Circuits Society and an IEEE SSCS Distinguished Lecturer.

Advertisers' Index

ADVERTISING SALES OFFICES

Closing date for space reservation:
1st of the month prior to issue date

NATIONAL SALES OFFICE
Eric L. Levine
Advertising Sales Manager
IEEE Communications Magazine
3 Park Avenue, 17th Floor
New York, NY 10016
Tel: (212) 705-8920
Fax: (212) 705-8999
Email: e.levine@comsoc.org

SOUTHERN CALIFORNIA
Patrick Jagendorf
7202 S. Marina Pacifica Drive
Long Beach, CA 90803
Tel: (562) 795-9134
Fax: (562) 598-8242
Email: patrick@roymcdonald.com

NORTHERN CALIFORNIA
George Roman
4779 Luna Ridge Court
Las Vegas, NV 89129
Tel: (702) 515-7247
Fax: (702) 515-7248
Cell: (702) 280-1158
Email: george@roymcdonald.com

SOUTHEAST
Scott Rickles
560 Jacaranda Court
Alpharetta, GA 30022
Tel: (770) 664-4567
Fax: (770) 740-1399
Email: srickles@aol.com

EUROPE
Rachel DiSanto
Huson International Media
Cambridge House, Gogmore Lane
Chertsey, Surrey, KT16 9AP
ENGLAND
Tel: +44 1428608150
Fax: +44 1 1932564998
Email: rachel.disanto@husonmedia.com

NOVEMBER 2010 EDITORIAL PREVIEW

Energy Efficiency
New R&D Tools for Wireless Communications
Optical Communications
Automotive Networking & Applications

DECEMBER 2010 EDITORIAL PREVIEW

RF Communications
Consumer Communications: Gaming & Entertainment
Ad Hoc and Sensor Networks

Communications Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page qMags

Convergent Technologies

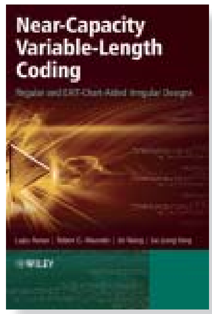
WILEY-BLACKWELL

Don't forget your IEEE member discount

Near-Capacity Variable-Length Coding

Regular and EXIT-Chart-Aided Irregular Designs

LAJOS HANZO, LIE-LIANG YANG, ROBERT G. MAUNDER and JIN WANG



Providing insights into the very latest research, the authors examine the design of diverse near-capacity VLC codes in the context of wireless telecommunications. It discusses the very latest research into VLC codes and introduces the novel concept of Irregular VLCs and their application in joint-source and channel coding.

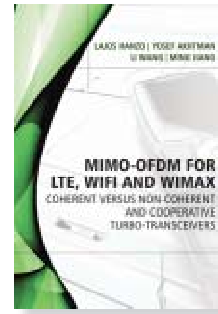
978-0-470-66520-6 • October 2010 • Hbk • 520pp • $140.00

Wiley-IEEE Press

MIMO-OFDM for LTE, WiFi and WiMAX

Coherent versus Non-coherent and Cooperative Turbo Transceivers

LAJOS HANZO, YOSEF AKHTMAN, LI WANG and MING JIANG



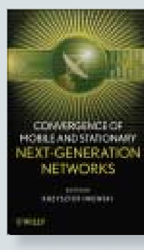
Provides an up-to-date portrayal of wireless transmission based on OFDM techniques augmented with Space-Time Block Codes (STBCs) and Spatial-Division Multiple Access (SDMA). The volume also offers an in-depth treatment of cutting-edge Cooperative Communications.

978-0-470-68669-0 • October 2010 • Hbk • 720pp • $170.00

Wiley-IEEE Press

Convergence of Mobile and Stationary Next Generation Networks

KRZYSZTOF INIEWSKI

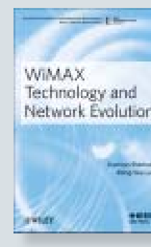


Featuring contributions from top industrial experts and academic professors, this authoritative work provides a comprehensive introduction to next-generation networks. It explains wireless networks such as wireless local area networks (WLAN), wireless personal area networks (WPAN), wireless access, 3G/4G cellular, and RF transmission.

978-0-470-54356-6 • October 2010 • Hbk • 808pp $115.00

WiMAX Technology and Network Evolution

KAMRAN ETEMAD and MING-YEE LAI



Written and edited by experts who have developed WiMAX technology and standards, this book presents an accurate, complete, and objective description of mobile WiMAX technology.

978-0-470-34387-6 • September 2010 • Pbk • 536pp $59.95

Wiley-IEEE Press ComSoc Guides to Communications Technologies

IP Telephony

Deploying VoIP Protocols and IMS Infrastructure, 2nd Edition

OLIVIER HERSENT

This book provides a comprehensive and practical overview of the technology behind Internet Telephony (IP), providing essential information to Network Engineers, Designers, and Managers who need to understand the protocols.

978-0-470-66584-8 • November 2010 • Hbk • 480pp $90.00

Satellite Technology

Principles and Applications, 2nd Edition

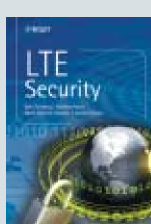
ANIL K. MAINI and VARSHA AGRAWAL

A comprehensive, single-source reference on satellite technology and its applications this book includes the latest developments on the topic. It covers the features and facilities of satellites and satellite launch vehicles, with an emphasis on the fundamental principles and concepts.

978-0-470-66024-9 • October 2010 • Hbk • 704pp $130.00

LTE Security

GÜNTHER HORN, DAN FORSBERG, WOLF-DIETRICH MOELLER, and VALTTERI NIEMI



Addressing the security solutions for LTE, this book shows how LTE security substantially extends GSM and 3G security. It also encompasses the architectural aspects, known as SAE, to give a comprehensive resource on the topic.

978-0-470-66103-1 • September 2010 • Hbk • 296pp $105.00

MIMO-OFDM Wireless Communications with MATLAB

YONG SOO CHO, JAEKWON KIM, WON YOUNG YANG and CHUNG-GU KANG

The authors provide a comprehensive introduction to the theory and practice of wireless channel modeling, OFDM, and MIMO, using MATLAB® programs to simulate the various techniques on MIMO-OFDM systems.

978-0-470-82561-7 • September 2010 • Hbk • 544pp $155.00

HOW TO ORDER

Wiley books are available through your Bookseller.
Alternatively send your order to:
John Wiley & Sons Inc., 111 River Street, Hoboken, NJ 07030, USA

For telephone orders:
Call 1-800-225-5945 (1-800-US-WILEY)
Fax: (212) 850-8888
E-mail: custserv@wiley.com

ORDER WILEY ONLINE...

- Visit the Wiley Electrical Engineering homepage: www.wiley.com/electrical
- Wiley Communications Technology Website: www.wiley.com/go/commstech/
- Amazon Wiley Communications Technology bookstore: www.amazon.com/professional/

Please note all prices correct at time of going to press but subject to change.

18015