Contents

Special issue on telecommunications network strategy and planning

cinkler@tmit.bme.hu

Networks 2010 *(www.networks2010.pl)*, the 14th International Telecommunications Network Strategy and Planning Symposium will take place on September 27-30, 2010 in Warsaw, Poland. For this occasion we have edited a Special Issue of the Journal of Infocommunications on Telecommunications Network Strategy and Planning that consists of six selected papers from Networks 2008 *(www.networks2008.org)*, the 13th International Telecommunications Network Strategy and Planning Symposium.

Networks 2008 was a five-day professional and scientific event held from September 28, 2008 in Budapest, Hungary, organized jointly by the Scientific Association for Infocommunications *(www.hte.hu)* and the Department of Telecommunications and Media Informatics *(www. tmit.bme.hu)* of the Budapest University of Technology and Economics (BME-TMIT). Altogether 90 papers were presented in 21 technical sessions. 26 countries were represented by authors with the most papers coming from Hungary, then Japan, US, Germany, Canada, Spain and UK in descending order. There were eight keynote talks held by influential telecom experts in four plenary sessions, as well as 11 half-day tutorials in two days. Throughout the conference there was also a related exhibition. The conference had a total of 335 registered participants from 30 countries, including 74 students that took advantage of the special offer to them of free attendance at the tutorials. All the Networks 2008 papers are available through IEEE *(ieeexplore.ieee.org)*, while all the presentations and tutorials are available at: *http://www.networks2008.org/online_offline_presentations.*

The theme of the conference, which was "Convergence in Progress", was reflected by the presentations from various perspectives. Not only new scientific methods for network optimisation and planning were presented, but also very practical approaches, experiences and case studies on introducing new solutions, and on the convergence of telecommunications, information and media technologies, as well as of fixed and mobile communications.

For the closing ceremony of the conference three papers were selected for the Best Paper Award. These are the first three papers of our Special Issue. The first, the US paper titled "FTTH Network Economics – Key Parameters Impacting Technology Decisions" performed a thorough analysis of technology and economic aspects in access network deployment.

The second, the French paper titled "Migrating to a Next Generation WDM Core Network" focuses on core networks and suggests a migration strategy towards switched wavelength-division multiplexed networks based on techno-economic analysis.

The third paper by authors from Spain titled "All-optical networks and switching technologies for a 3D videoconference system with the feeling of presence" discusses how an optical network should be designed and what optical technology should be chosen in order to meet strict higher-layer service requirements.

The fourth paper by Japanese authors is entitled "Reducing total call-blocking rates by flow admission control based on equality of heterogeneous traffic". This teletraffic paper proposes a new connection admission control scheme that guarantees quality of services (low blocking), which is of particular interest for the increasingly popular video services.

The fifth paper by Polish authors titled "Flow optimization in IP networks with fast proactive recovery" proposes and evaluates new methods for recovering IP networks after a failure.

And finally the last, sixth paper, is a joint paper of authors from different countries including Hungary, titled "Network Resilience Requirements and Algorithms for Multicasting and Broadcasting Digital TV" which discusses and evaluates methods for protecting multicast TV services upon a failure to minimise the impact onto the quality experienced by the users.

We hope that our selection of papers clearly represents the main areas of interest of recent Networks conferences. Enjoy reading them!

Tibor Cinkler,
BME-TMIT, Hungary
Oscar Gonzalez-Soto,
ITU Consultant, Spain
Gyula Sallai,
BME-TMIT, HTE, Hungary
Rati C. Thanawala,
Alcatel-Lucent, USA
Andy Valdar,
University College London, UK

Guest Editors

TIBOR CINKLER *(cinkler@tmit.bme.hu)* has received M.Sc.('94) and Ph.D.('99) degrees from the Budapest University of Technology and Economics (BME), Hungary, where he is currently associate professor at the Department of Telecommunications and Media Informatics (TMIT). His research interests focus on optimisation of routing, traffic engineering, design, configuration, dimensioning and resilience of IP, Ethernet, MPLS, ngSDH, OTN and particularly of heterogeneous GMPLS-controlled WDM-based multilayer networks. He is author of over 180 refereed scientific publications and of 4 patents. He teaches various related courses on networking and optimization at the university, as well as for companies and also gives tutorials at conferences and summer schools. He has been involved in numerous related European and Hungarian projects and he took part in the work of Technical Program Committees of numerous conferences. He was the Technical Program Committee Chair of Networks2008, the International Telecommunications Network Strategy and Planning Symposium, held in Budapest in 2008.

OSCAR GONZÁLEZ SOTO *(oscar.gonzalez-soto@ ties.itu.int)* graduated as Telecommunicación Engineer at the Escuela Superior de Ingenieros de Telecomunicación de Madrid in 1969. Performed master seminars in Strategy for the International Telecom Industry and Business Management Development at the London Business School and Euroforum – INSEAD. Currently he is a Consultant for Strategic Planning and Assessment in the Telecom sector. He is associated to the ITU-BDT for the programs in Network Planning and for the International Centres of Excellence. Previously has been Director of Network Planning and Economics at the Alcatel centre for network design and integration at Paris HQ, after holding different positions as manager for techno-economical evaluations and network planning development (Strategy centre in Brussels), Telecom planning tools, Performance analysis and Teletraffic methodologies (Research centre in Madrid). He participated in the organization of more than 15 international conferences and has published more than 50 papers at international fora. Currently is member of the Advisory Council for the International Teletraffic Congress, of the International Management Committee for the NETWORKS International Symposium and the IEEE.

GYULA SALLAI received MSc degree from the Budapest University of Technology and Economics (BME) in 1968, PhD and DSc degrees from the Hungarian Academy of Sciences (MTA) in 1976 and 1989 resp., all in telecommunications. He was appointed as honorary professor in 1990, as full professor in 1997. His professional life is related to the telecommunications, then the ICT. He was senior researcher in telecommunication network planning, then research director, strategic executive director, later deputy CEO responsible for telecommunication services with the Hungarian Telecom Company; then international vice president, after that executive vice president for the ICT regulation and scarce resource management with the Communication Authority of Hungary. Since 2002 he is the head of the Department of Telecommunications and Media Informatics at the BME. He was the vice-rector of the BME as well. He is also the chairman of the Telecommunication Committee of the MTA, the president of the Hungarian Scientific Association for Infocommunications (HTE) and a member of the Hungarian Academy of Engineering. Recently his main research area is the ICT management and regulatory issues. He was the general chair of the International Telecommunications Network Strategy and Planning Symposium, held in Budapest in 2008 (Networks2008).

RATI C. THANAWALA is Network Planning, Performance and Economic Analysis Vice President at Bell Labs. Her organization of Bell Labs scientists and engineers works with operators in the communications industry worldwide, and the U.S. Government, supporting advanced technology planning for evolution to Next Generation Networks. The work includes end-to-end network modeling and network design, performance, reliability and operations systems engineering, and business cases for customers and product managers evaluating product and network evolution scenarios. Rati is also a member of the Homeland Security Standards Panel created by the U.S. Department of Homeland Security in cooperation with the American National Standards Institute to align the cutting-edge efforts of the standards community with urgent national priorities of homeland security. Rati has over 20 years of experience in introduction of new technologies into telecommunication networks, leading innovations for new products and services, and managing global organizations across a variety of R&D, Systems Engineering, and Business Management functions. Rati holds a BS in Mathematics from Lucknow University, India and a Masters and Ph.D. in Computer Science from Yale University. She attended the Program for Management Development at Harvard University.

ANDY VALDAR obtained his honours degree in electronic and electrical engineering at Loughborough University of Technology in 1969. Then he joined BT (then the GPO) as an open-competition executive engineer in the Network Planning Department. Some 18 months after joining BT, Andy was awarded a bursary to study full time at Essex University for an MSc in telecommunications systems. On returning to BT he became involved in the pioneering work concerned with the introduction of digital switching in BT's network. In 1977, Andy joined the UN ITU agency for a three-year teaching and course development assignment in India and subsequently undertook ITU teaching assignments in Bangladesh, India, and Swaziland. After returning to the UK in 1980, Andy's career in BT has ranged from network planning, development of international standards and technical strategy, marketing, product management and new product development. During the last nine years of his time in BT he became General Manager of network and technology strategy. In 1999 Andy left BT to take up the role with University College London as Academic Director of the BT Masters Programme, which led to an MSc in Telecommunications Business. He then moved to take over the directorate of BT's other MSc Programme – the BT MSc. Andy is currently an active participant in international telecommunications conferences, chairman of the board of Editors on the Journal of the Institute of Telecommunications Professionals (ITP), and is the author of a recently published best selling text book explaining telecommunications networks.

Infocommunications Journal

Editorial Office (Subscription and Advertisements)**:**
Scientific Association for Infocommunications
H-1055 Budapest, Kossuth Lajos tér 6-8, Room: 422
Mail Address: 1372 Budapest Pf. 451. Hungary
Phone: +36 1 353 1027, Fax: +36 1 353 0451
E-mail: info@hte.hu, Web: www.hte.hu

Articles can be sent also to the following address:
Budapest University of Technology and Economics
Department of Telecommunications
Tel.: +36 1 463 3261, E-mail: szabo@hit.bme.hu

Subscription rates for foreign subscribers:
4 issues 50 USD, single copies 15 USD + postage

Publisher: PÉTER NAGY • Manager: ANDRÁS DANKÓ

HU ISSN 2061-2079 • Layout: MATT DTP Bt. • Printed by: Regiszter Kft.

FTTH network economics: Key parameters impacting technology decisions

SAMRAT KULKARNI, MOHAMED EL-SAYED, PAUL GAGEN, BETH POLONSKY

Network Planning - Bell Labs, Alcatel-Lucent Technologies, USA
svk@alcatel-lucent.com

Keywords: FTTH, GPON, EPON, P2P, network economics

Different technology options are available to operators today for their Fiber-to-the-home (FTTH) network deployment strategy decisions. Gigabit-Passive Optical Network (GPON), Ethernet Passive Optical Network (EPON), Active Ethernet (AE) and Point-to-Point Ethernet (P2P) are the major competing technologies. There are a number of technical, economic and business drivers that impact the right choice for each specific network situation. When modeling network economics, it is important to consider a Total Cost of Ownership (TCO) model to enable operators to rightly evaluate these choices, instead of comparisons of only specific cost elements (e.g., port costs). On the other hand, in a network model with a large number of parameters it is often challenging to identify the key parameters that are critical to the decision-making. An operator runs the risk of picking an incorrect technology strategy if any of these key parameters are not identified and cost optimized.

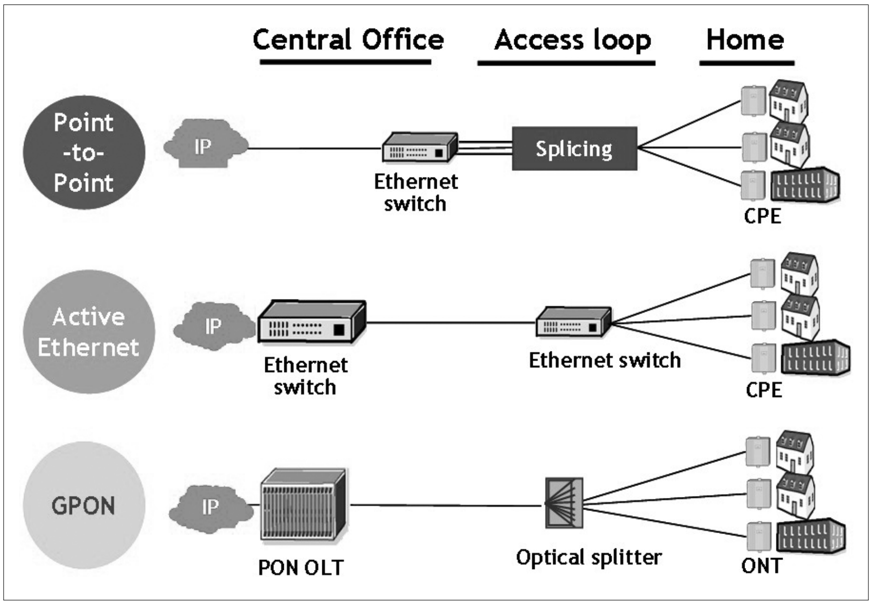
In this paper, we present the results obtained by modeling the capital investments and operations expenses incurred for some operator cases, and identify the key parameters that impact FTTH economics covering these major technologies. We show sensitivity analysis to identify the critical parameters. The methodology and results will enable operators quickly make the right FTTH technology deployment decisions.

1. Introduction

The number of households with fiber-optic network connections will grow by nearly 43% worldwide in 2008 and will continue to grow at rates above 30% a year through 2012, when the number of fiber-connected households will reach nearly 90 million globally according to a recent report by Heavy Reading [1]. Hence it is no surprise that almost all the network operators around the world are evaluating the different FTTH technology options today. Deploying a FTTH network requires significant upfront capital investments and it is absolutely critical for an operator to build a detailed network economic model and pick the right technology that optimizes their capital expenses, operations expenses and payback period. Service Provider network requirements and topologies vary considerably; hence network modeling and solutions need to be tailored to specific service provider situations.

Figure 1 shows the key FTTH technology options that exist today and are already being deployed by some of the major operators in the world. GPON [2] and EPON [3] optimize the outside plant (OSP) by using a passive splitter which provides bandwidth aggregation, requires less maintenance and doesn't have any power requirements like an active network element. The Active Ethernet solution achieves optimization in the OSP by using an Ethernet switch for aggregation, but requires hardened cabinets and remote power supply. The Point-to-point solution also uses Ethernet switching and aggregation, however all the Ethernet switches are deployed in the Central Office (CO). These COs, also known as Points-of-Presence (PoPs) tend to be closer to the subscriber.

Figure 1. Key FTTH technology options



In the next section, we discuss the details of a comprehensive and flexible network cost model that compares these FTTH technology options, and quantifies the savings through case studies. This model also includes a task-based operations analysis. Results of three case studies are presented, and sensitivity analyses of the results are applied to identify the key parameters.

2. Modeling framework

The network economic modeling framework includes capital investments (CAPEX) and operations expenses (OPEX) optimization for the technology options and across scenarios applicable to typical service provider networks. The services revenues supportable by these access options are assumed to be common and hence are not included in this model Also, the scope is related to cost-elements, and does not cover other aspects (such as performance, standards etc).

Typical operator scenarios include:
- Type of subscriber:
  Single-family residential (SFR),
  Multi-Dwelling Unit (MDU) and Enterprises.
- Subscriber housing density:
  Loop lengths from the CO,
  number of houses per square-km;
- Network build type: Greenfield, Overbuild.
- Fiber cost type:
  Leased vs. own, one-time fee vs. recurring.
- Outside plant construction type:
  Aerial, Buried, Conduit, Sewer etc.
- Splitting levels:
  1-Tier centralized and/or 2-Tier distributed for PON.

Typical cost elements are:
- Hardware and software for Central Office (CO), OSP, and Customer Premises Equipment (CPE), active equipment and operating support systems. List prices are prorated based on experience curves (market averages and up to 10% annual cost reductions) and equipment discounts (0-50%) for sensitivity analysis.
- Cost of the OSP: feeder, distribution and drop fiber; civil works for the structures, trenches, installation and splicing; cabinets, splitters, fiber management points and patch-panels.
- Power and space/housing costs: Costs to setup active nodes, realtor fees, provisioning of AC, ongoing energy costs and floor space rental.
- Activation costs such as truck roll to OSP, customer service visit, service activation in CO.
- Other operations cost such as provisioning and maintenance activities.

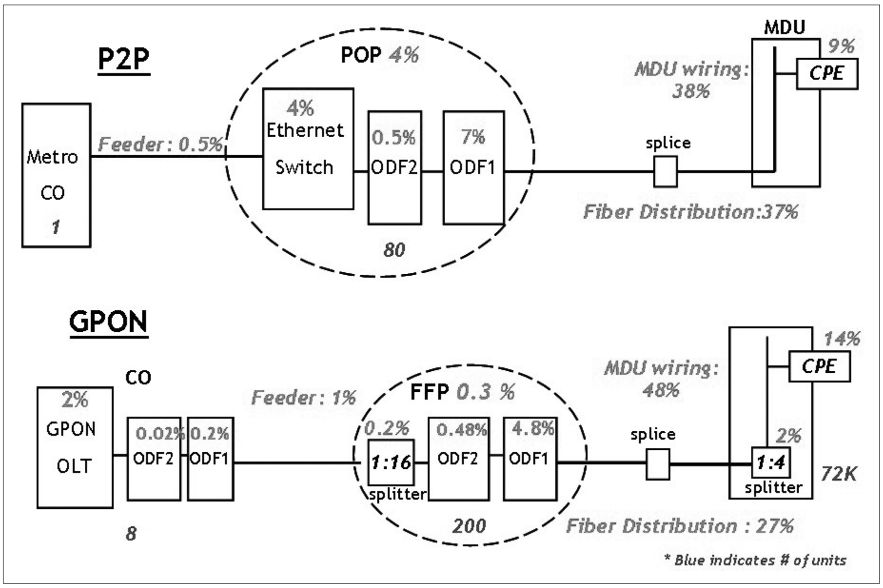


Figure 2. CAPEX Breakdown at 20% take-rate

3. Case Studies

In this section, we discuss and summarize results from three network modeling case studies as below. For all the three case studies, it can be assumed that the CAPEX/subscriber and OPEX/subscriber have been optimized for each technology solution by assuming a reasonable OSP model and based on the specific scenario and cost parameters. Then we use sensitivity analysis techniques to identify the key network cost parameters.

3.1 Case Study 1: GPON vs P2P for a dense urban city

This case study compares the costs of deploying GPON and P2P in a dense urban MDU subscriber base. There are close to a million households (HH) passed in an area of roughly 100 square kms. The average size of an MDU is assumed to be 16HH. A GPON operator has 8 CO locations to serve these HH and 200 Fiber Flexibility Points (FFPs) where splitters are located, whilst the P2P operator is deploying 80 new PoPs. The civil works is assumed to use existing structures such as the sewer in the city thereby eliminating most of the trenching and duct costs. In our modeling analysis the take-rate is varied from 0-100%.

Figure 2 shows the P2P and GPON 2-Tier architectures and the corresponding CAPEX cost components for a take-rate of 20%. The 2-Tier GPON architecture assumes a splitter in the basement of the building.

It is observed that the bulk of the CAPEX/sub is in the fiber distribution and MDU wiring. The CPE accounts for the next highest cost component followed by Ethernet switch cost and the GPON OLT. The remaining network elements do not contribute significantly to the overall cost.

We also find that GPON provides a saving of about 20% compared to P2P at a take-rate of 20%, and the savings are positive over the entire range of take-rates up

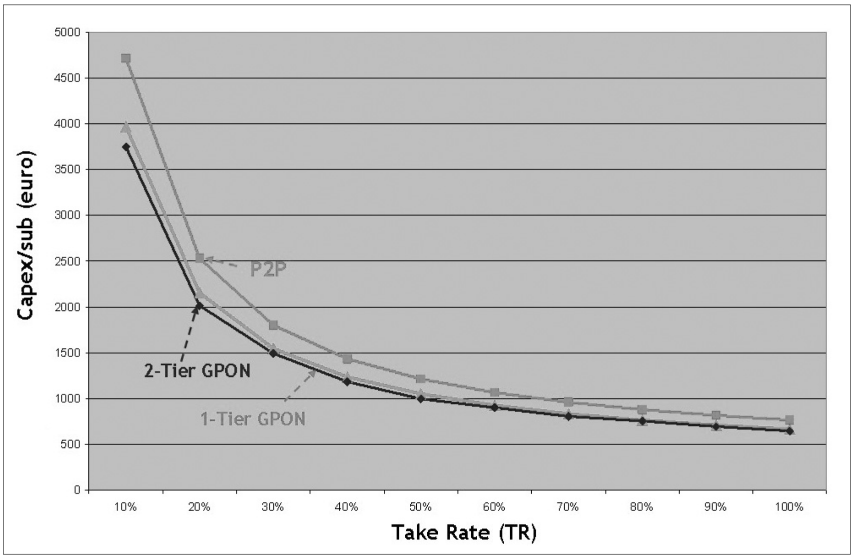


Figure 3. CAPEX/sub vs. Take-rate

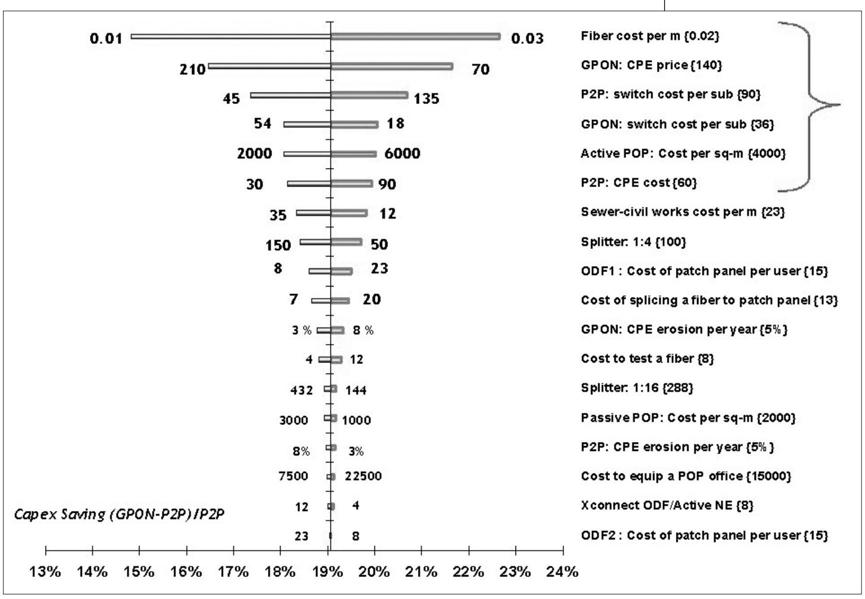
to 100% *(Figure 3)*. Also 2- Tier GPON is cost-effective by 0-10% over 1-Tier GPON and the savings are higher at lower take-rates.

To identify the key cost parameters, we look at the results of a single parameter sensitivity analysis (Tornado analysis) as shown in *Figure 4.*

Sensitivity analysis shows, the Top-5 CAPEX parameters are:
- Fiber-cost per meter,
- GPON CPE cost,
- Ethernet switch cost,
- Real-estate cost for Ethernet switches deployed outside the CO,
- GPON OLT cost.

The remaining parameters have an impact on the overall CAPEX but do not swing the GPON vs. P2P decision as much.

Figure 4. CAPEX Tornado analysis



The tornado analysis does not capture the interaction between the parameters. Hence a 1000-iteration Monte-Carlo analysis was performed with a ±50% variation in the value of the cost parameters with all parameters varied randomly per iteration, and histogram of results plotted. The x-axis of *Figure 5* shows the percent savings of GPON over P2P. Even with this wide range of variation, GPON still provided a significant cost advantage over P2P making it the technology of choice.

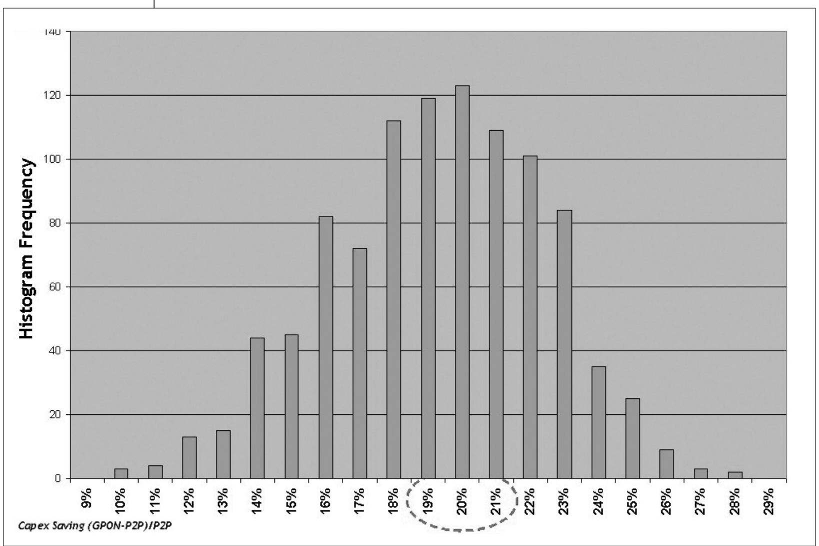
Operations Cost Modeling Results

The following operations cost elements are considered in this model.

• Unplanned Maintenance:

Repair activity based on equipment quantities and FIT data. Operation tasks include: testing, fault isolation and equipment repair (Truck roll).

Figure 5. CAPEX Monte-Carlo analysis



• Planned Maintenance:

Calculated based on equipment quantity, maintenance interval and effort, equipment clustering and location density. Fiber maintenance based on total length of cables and yearly per meter. Operation tasks include: battery replacement, fan filter replacement, drive time and paperwork to document preventative maintenance, fiber inspection/cleaning and debris removal.

• Centralized NOC Staffing:

Surveillance staff estimated based on total number of active devices. Operation tasks include: 24x7 fault monitoring, remote diagnostics and trouble ticket creation.

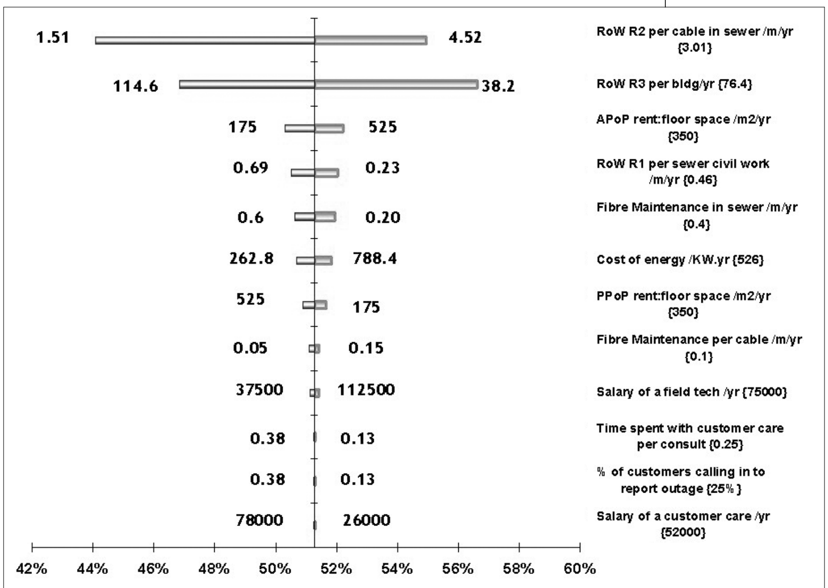
• Differences in Customer Provisioning and Disconnect Scenarios:

Cost of connecting/disconnecting a customer based on equipment locations and utilization. Operation Tasks include: CPE installation, in-building fiber connection, PoP/FFP connections, testing and inventory updates.

• Customer Care:

Estimated based on failure incidence (calculated for unplanned maintenance) and number of customer impacts/incidents. Operation Tasks include: customer care call handling.

Figure 7. OPEX Sensitivity Analysis



The OPEX modeling results show a saving of 55-60% for GPON compared to P2P over a wide range of take-rates *(Figure 6)*. These savings are mainly due to the higher Right-of-Way (RoW) expenses for P2P given the large amount of fiber infrastructure deployed on Day 1.

The Right-of-Way is a yearly recurring expense that the operator in this case would need to pay to the governing entity to use the civil works infrastructure while laying out fiber cables.

Typical components of RoW are a fixed cost to access the civil works in the OSP and in the building, and a variable cost as a function of the number of cables run. Furthermore, the cost of maintenance and management is higher in P2P compared to GPON because of the higher number of fiber-pairs deployed.

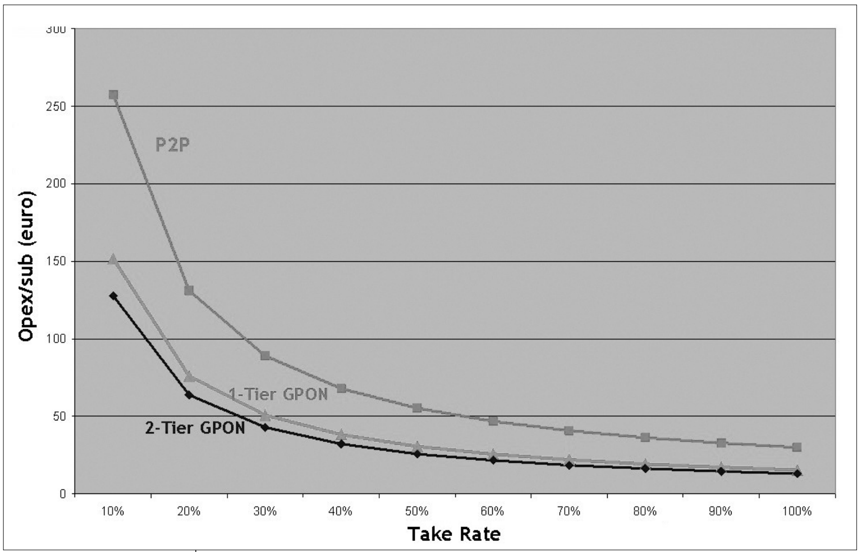


Figure 6. OPEX per sub vs. Take-rate

The sensitivity analysis *(Figure 7)* shows that the Top-3 OPEX impacting parameters are:

– Right-of-Way charges,

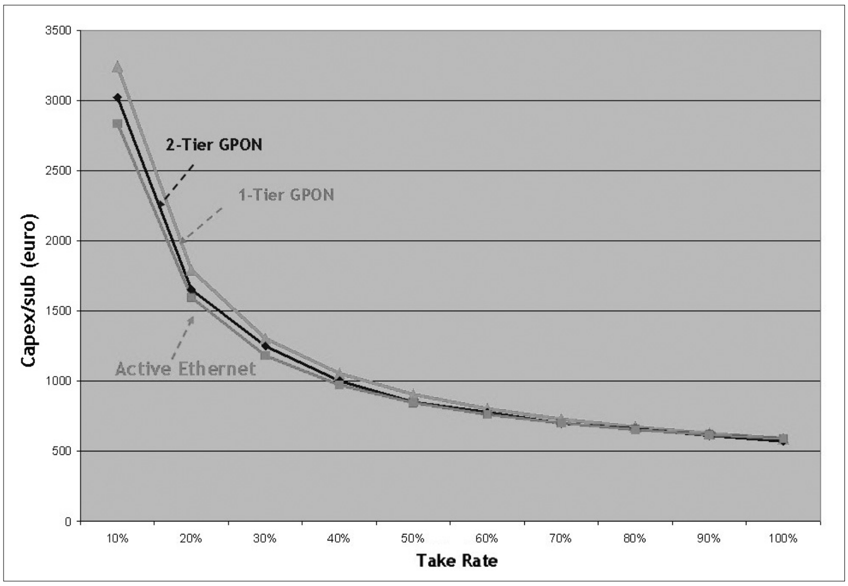
– cost of energy (KW/hr),

– fiber maintenance costs.

A Monte-Carlo analysis was also performed that confirmed the results that GPON provides significant OPEX savings compared to P2P, and details are not presented here for brevity.

Also, although this case study looked at a dense urban MDU deployment, it has been found that the same trend in terms of the key parameters and GPON savings are applicable to a single family residential urban and sub-urban model with reasonable population densities and operator deployment scenarios.

Figure 8. GPON, AE CAPEX vs. Take-rate



3.2 Case Study 2: GPON vs. Active Ethernet for a dense urban city

An operator that has deployed DSLAMs to provide ADSL/VDSL broadband access to some end users may consider this scenario when they decide to migrate to a FTTH last mile, by provisioning fiber loops to the cabinet. Also, technology exists today to install Ethernet cards in an existing (DSLAM) street cabinet. We compare the CAPEX/sub and OPEX/sub for this operator to an operator deploying only GPON to serve the fiber subscribers.

The same dense urban MDU subscriber model as in Case Study 1 is used here. We assume for the Active Ethernet model, buried fiber civil work is needed in the distribution network only (cabinet to sub) since fiber already exists to the cabinet to backhaul the DSLAM traffic Assuming typical serving areas of 250-300 per cabinet, about 4500 cabinets are needed.

Figure 8 shows that Active Ethernet has a saving of about 5% compared to GPON 2-Tier and the savings diminish with increased take-rate. Given the range of savings (<5%), it is argued that neither technology is the clear winner in terms of CAPEX/sub.

Figure 9 shows the breakdown of CAPEX/sub at 20% take-rate. The AE solution has zero housing cost for the PoP since the OSP cabinets are re-used (unlike in Case Study 1), and remaining costs balance out. Therefore CAPEX is not a key differentiator in this case.

Now considering the OPEX/sub as shown in *Figure 10,* we find that GPON provides large savings compared to P2P. The OPEX savings for 2-Tier GPON increase with higher take-rates and are in the range of 5-58% savings annually. Therefore, if the operator plans to target for a 30% or higher subscriber take-rate, then GPON should be the technology of choice.

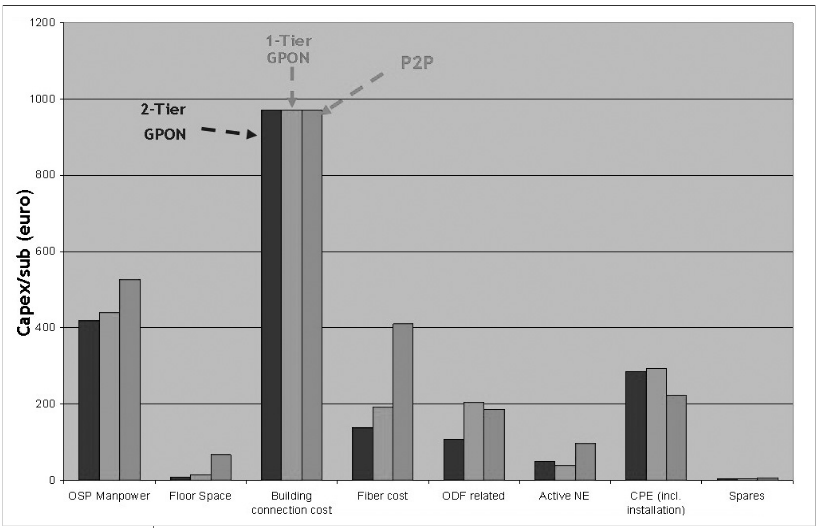


Figure 9. GPON vs. AE CAPEX (20% TR)

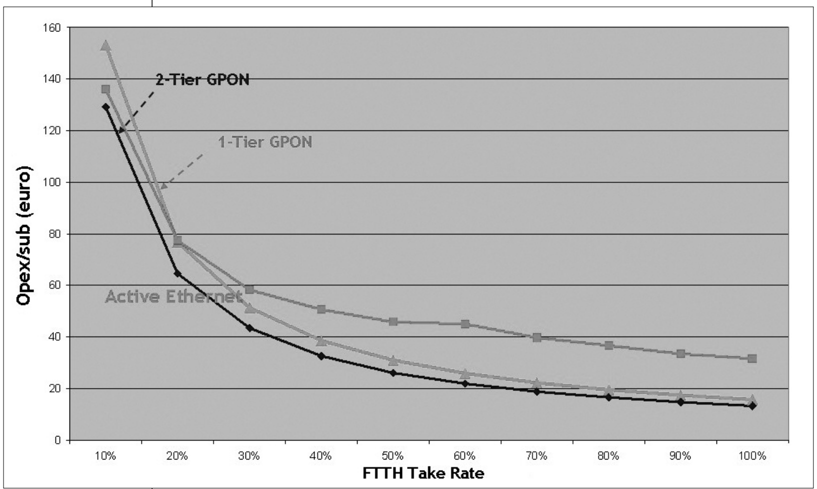


Figure 10. GPON, AE OPEX/sub vs. Take-rate

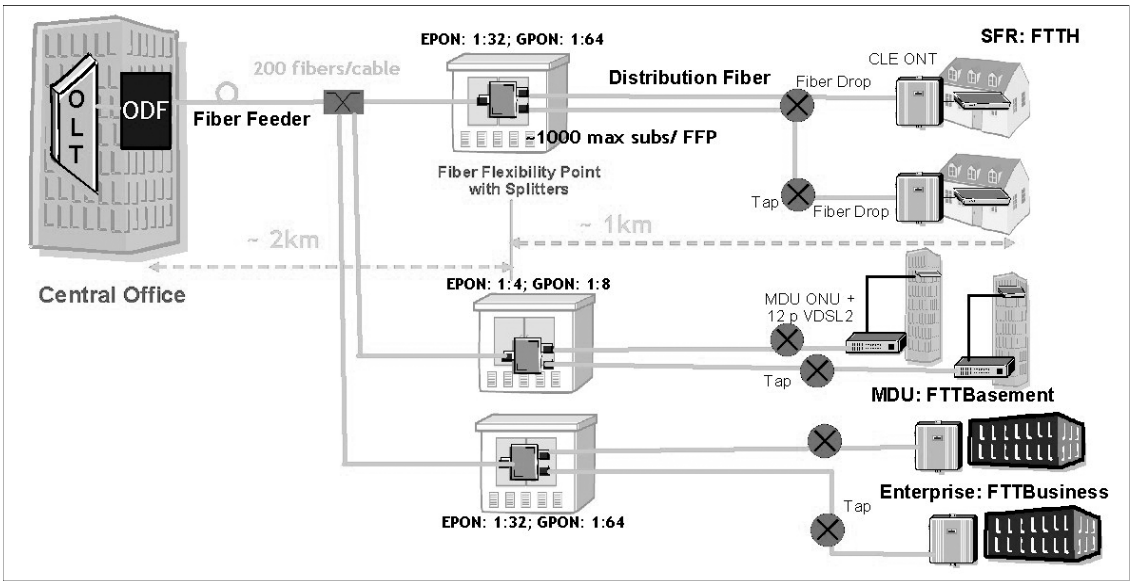


Figure 11. GPON vs. EPON network model

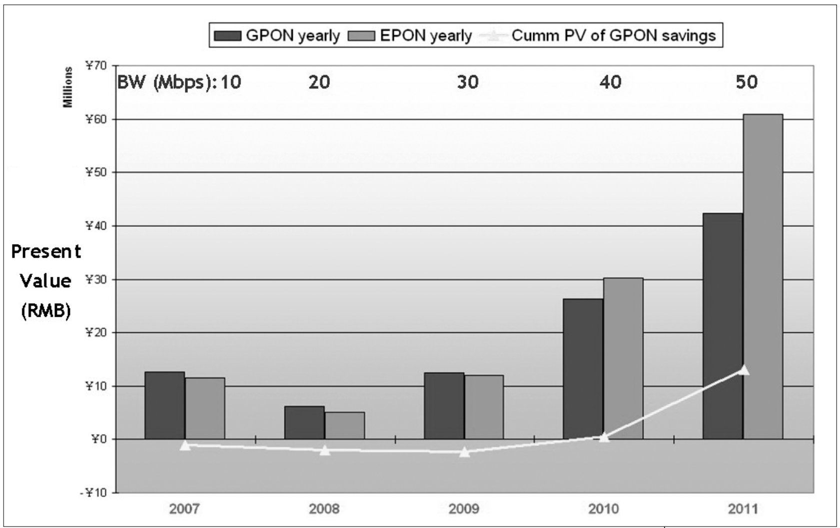


Figure 12. CAPEX for SFR (year-over-year)

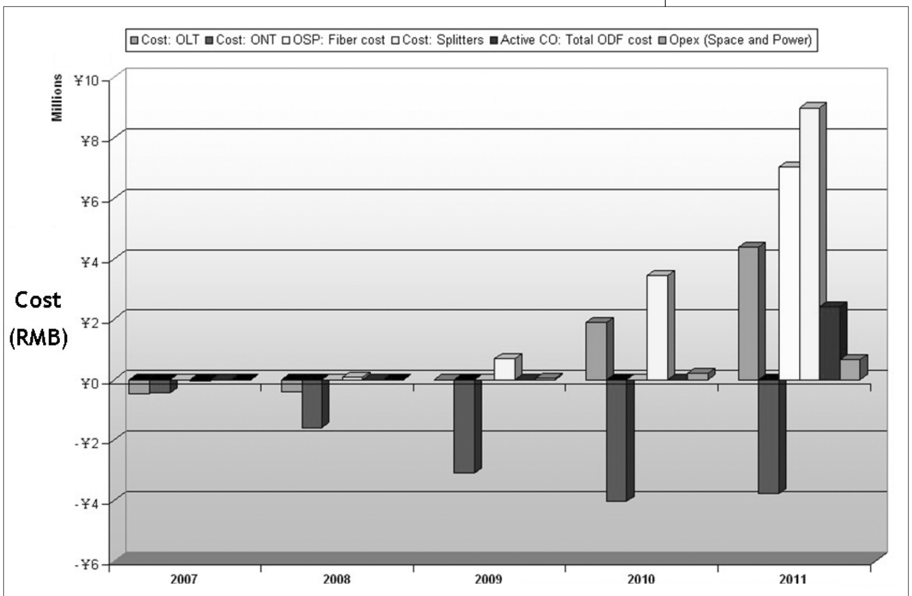


Figure 13. SFR CAPEX Delta (EPON-GPON)

We note that this case is really a special case of Case Study 1 with the assumption that the Ethernet switch is deployed in the OSP using a hardened cabinet instead of a PoP. The relevant key parameters identified through sensitivity analysis in Case Study 1 apply here as well.

3.3 GPON vs. EPON

This case study is for an operator deciding between GPON and EPON. We model an operator deploying a network in an urban city. The model assumes a deployment period of five years (2007-2011). *Figure 11* shows the architectures for the SFR model which is FTTH, MDU model which assumes Fiber-to-the-Building and VDSL2 inside the building, and enterprises served by fiber (Fiber-to-the-Business).

Each type is modeled independently. The MDU case assumes copper loops inside the building are used instead of fiber all the way. Services bandwidth is assumed to grow from 10 Mbps/year starting with 10 Mbps in 2007 to 50 Mbps in 2011.

Cost items modeled include: active NE (CO switch, CPE), passive components (splitter, ODF, fiber) and OPEX (space, power). It can be noted that both technologies use the same OSP infrastructure (civil works etc.) and that cost is ignored in this model.

Comparing the CAPEX/sub *(Figure 12)* shows that EPON provides a lower start-up cost in the initial years, but requires significant investment in future years.

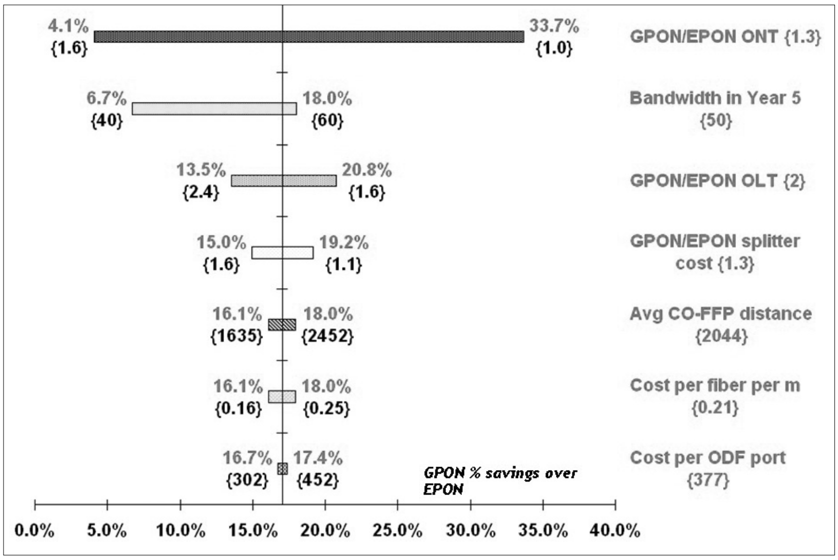
- Present Value of Savings of GPON over EPON = 17% (SFR), 19% (MDU) and 30% (Enterprise)

Figure 13 provides a breakdown of the key cost elements for the SFR case. When the bandwidth is low (<20 Mbps until 2008), EPON saves on all cost components. With increased bandwidth however, GPON scales better costwise whereas EPON needs more OLT ports, splitters, fibers etc.

This is because of the lower overheads and higher payload bit-rates in the GPON technology today compared to EPON.

Sensitivity analysis in *(Figure 14)* shows the key parameters impacting the economics.

Figure 14. GPON vs. EPON sensitivity analysis



The Top-2 critical parameters are:
– user bandwidth: if BW is below 30 Mbps, GPON doesn't save compared to EPON,
– GPON/EPON price ratios for the ONT and ONU.

Evaluating these key parameters correctly will enable the operator make the right technology decisions moving forward.

4. Conclusions

A detailed analysis of FTTH economics across a range of different scenarios and parameters was developed comparing GPON, EPON, P2P and Active Ethernet. Results for three real-world customer modeling case studies were presented with sensitivity analyses.

They are summarized as follows:

Case Study 1: (GPON vs. P2P network)

Over a wide range of take rates and parameters, GPON provides lower CAPEX/sub and OPEX/sub compared to P2P. This is primarily due to the significant OSP fiber investment needed on Day 1 for P2P.

- Average savings: CAPEX = 20% and OPEX = 55-60%
- 2-Tier GPON is more cost effective than 1-Tier (for MDU) by 0-10%.
- The specific results above apply to an example case of an overbuild FTTH network deployment in MDU; our studies show GPON savings apply to an urban/sub-urban SFR deployment case as well.

Sensitivity analysis shows that the Top-5 parameters impacting CAPEX are:
– fiber cost per meter,
– GPON CPE cost,
– Ethernet switch cost,
– real-estate/housing cost for Ethernet switches deployed outside the CO,
– GPON OLT cost.

Sensitivity analysis indicates that the Top-3 OPEX parameters are:
– Right-of-Way
– cost of energy
– fiber maintenance.

Case Study 2: (GPON vs. AE network)

Here an operator may consider serving FTTH subscribers directly from the DSLAM chassis. In such situations, the economics of Active Ethernet and GPON will change considerably.

- An Ethernet card in the DSLAM is expected to provide a cost effective solution for FTTH in low or medium fiber deployment situations (take-rates ~10-20%).
- The CAPEX/sub difference between GPON and AE are small (<5%), but OPEX/sub is a big differentiator. GPON provides OPEX savings from 5-58% with higher savings for increasing take-rates.
- In areas where no DSLAMs are deployed, GPON is expected to be more cost effective in general because of the additional cost of building the OSP cabinets for AE and significant OPEX savings from a passive outside plant.

Case Study 3: (GPON vs. EPON)

EPON provides a lower start-up cost, but requires significant investment in future years as demonstrated for all cases. A savings of 17% over EPON was obtained for the urban SFR model, 19% for MDU and 30% for the Enterprise model. Sensitivity analysis indicated that the two key parameters impacting the economics are:
– GPON/EPON ONT and ONU cost ratios,
– end year subscriber bandwidth.

Acknowledgments

The authors would like to acknowledge their colleagues Fabien Pinaud and Chen Jing for their contributions towards the case studies discussed in this paper.

Authors

SAMRAT KULKARNI is a Member of Technical Staff in the Bell Labs Network Planning, Performance and Economic Analysis Group with Alcatel-Lucent Technologies, based in Murray Hill, NJ. He has over 6 years of experience in optical transport network modeling and optimization of access and metro networks, specifically in the areas of FTTx transport technologies like GPON/EPON, xDSL, Active/P2P Ethernet and next-generation SONET/SDH, Ethernet-over-fiber and DWDM technologies. He has consulted extensively with service providers performing network and business modeling activities for operators all around the world which includes, providing optimized network designs, product configurations, cost modeling and systems engineering. His work has also been published in journals like the Bell Labs Technical Journal, and conferences such as OFC/NFOEC. He holds a Masters of Science (MS) in electrical engineering from Purdue University (2001) and also serves as a chair on the local IEEE NJ coast Light Wave systems chapter.

MOHAMED EL-SAYED is the manager of the Advanced Network Modeling & Optimization group at Bell Labs Research, Alcatel-Lucent, in Murray Hill, New Jersey, USA. He has over 25 years of experience in the area of Telecommunications R&D. He started his career in Bell Labs in the early eighties where he was involved in the architecture and design of packet switching systems and networks. He has subsequently contributed to many areas in data communication including TCP/IP, Satellite Communication, Virtual Private Networks, MPLS, VoIP etc. He took on responsibilities in a variety of leadership roles in Research, Network Architecture, Systems Engineering, Network Design & Optimization as well as Network Performance. Dr. El-Sayed's current interests include architecture and design of 3G & 4G Wireless networks, IP Multimedia Subsystem (IMS) and Applications, FTTH, Intelligent Optical Networks, IPTV and Next Gen Transport networks based on Carrier-Grade Ethernet/VPLS and MPLS. He has consulted with executives of many Service Providers in North America, Europe, Asia and the Middle East helping them with a variety of questions related to network evolution, technology introduction and network design and optimization. He has published many papers in professional journals and conferences and was a recipient of many Bell Labs, Alcatel-Lucent and AT&T awards. Dr. El-Sayed holds a Ph.D. in Systems Engineering in addition to a B.Sc. and an M. Sc. in Electrical Engineering.

PAUL GAGEN is the Director of Network Design, Operations, and Methods in Bell Labs. He is responsible for developing advanced network designs/tools/methods applied to a multi-technology, global array of customers' network projects. These include advanced optical, data, broadband, video, and wireless networks. End-to-end digital video operations and QoS metrics are among the current investigation topics. Paul has a background in applied research; optical fiber communications; broadband access including HFC; product and manufacturing process development including the Manufacturing Leadership Program; technical sales for network deployments as the Bell Labs Senior Field Rep; and holds twelve patents. He holds a BS and MS degrees in Engineering summa cum laude from The Cooper Union and New York University respectively, and is an alumnus of the Harvard Graduate School of Business – Program for Management Development. Paul is located in Holmdel, NJ and is a recent winner of Lucent awards including the Bell Labs President's Gold Award, and Teamwork Awards.

BETH POLONSKY is a member of the Network Operations Planning team with expertise in Operational Modeling, Analysis, and Solution Design. In her current assignment, she has provided operational analysis for new wireless technologies, converged network architectures, and FTTH alternatives. This analysis has supported the proving in of technology for customers in the telecom and cable industries. She is also responsible for NGN IP network transformation process design and implementation to facilitate successful customer network migration. Over the last decade, Beth has been involved with many aspects of network operations software including system architecture, design, systems engineering and customer support. She has worked on site with several large North American Telcos to design and customize OSS solutions to meet their specific needs. In addition, Beth has lab experience integrating OSS system with various network configurations to prove interoperability and feature functionality in support of customer deployments. Beth has a BS degree in Electrical Engineering from the University of Connecticut and an MS in Computer Science from the University of Virginia.

References

[1] FTTH Worldwide Technology Update & Market Forecast by Heavy Reading, Feb 2008.

[2] ITU-T Recommendation G.984.

[3] IEEE 802.3ah Ethernet in the First Mile standard.

[4] M.K. Weldon and F. X. Zane, "The Economics of Fiber to the Home Revisited," Bell Labs Tech. Journal 8:1, pp.181–206, (2003).

Call for Papers

Prospective authors are invited to submit original research papers for publication in the upcoming issues of our Infocommunications Journal.

Topics of interests include the following areas:

Data and network security • Digital broadcasting
Infocommunication services • Internet technologies and applications
Media informatics • Multimedia systems • Optical communications
Society-related issues • Space communications
Telecommunication software • Telecommunications economy and regulation
Testbeds and research infrastructures • Wireless and mobile communications

Theoretical and experimentation research results achieved within the framework of European ICT projects are particularly welcome.

From time to time we publish special issues and feature topics so please follow the announcements. Proposals for new special issues and feature topics are welcome.

Our journal is currently published quarterly and the editors try to keep the review and decision process as short as possible to ensure a timely publication of the paper, if accepted.

As for manuscript preparation and submission, please follow the guidelines published on our website:
http://www.hiradastechnika.hu/for_our_authors

Authors are requested to send their manuscripts via electronic mail (preferably) or on a CD by regular mail to the Editor-in-Chief:

Csaba A. Szabó
Dept. of Telecommunications, Budapest University of Technology and Economics
2 Magyar Tudósok krt., Budapest 1117, Hungary
E-mail: szabo@hit.bme.hu

Migrating to a next generation WDM core network

Olivier Renais, Gaëlle Yven, Esther Le Rouzic

Orange Labs, France

{olivier.renais, gaelle.yven, esther.lerouzic}@orange-ftgroup.com

Keywords: tunability, PXC, ROADM, WB, WSS, WDM

The main goal of this study is to examine the different options that arise when building a modern core transport network, considering the short to medium term evolutions planned by the manufacturers. After a short review of possible WDM network designs, a techno-economic analysis allows for benchmarking different architectural scenarios, providing some guidelines in building a future proof WDM core network. At last we highlight some of the challenges to relieve when migrating to an agile WDM infrastructure.

1. Introduction

The explosion of the long distance traffic and the convergence towards all over IP affect the requirements for the transport layer. FTTH may act as an accelerator of this phenomena, and the transport bandwidth required by backbone routers is quickly moving from 10 Gbps towards the 40 Gbps scale.

Network availability becomes also a critical issue, and whereas traditional implementations were based on protection at higher layer using diversely routed but unprotected transport resources, network operators start to investigate on multilayer resilience mechanisms to optimize the core network architecture. In this context, one must consider that the IP layer could benefit from protected transport bandwidth and/or reconfigurable topology.

This clearly makes WDM as the federating layer in the core transport network and WDM granularity seems to be adapted to the future needs, thus limiting the interest of deploying an additional intermediate layer in a pure aggregation perspective.

However following the explosion of the “Internet bubble”, the downturn lead the different WDM equipment manufacturers to reduce the development of new products and advanced features. Some of them dramatically cut down on their R&D resources, which resulted in the freezing of advanced feature development, especially in the core transport WDM market considered as less profitable than in the metro space. The so promised control plane and associated end-to-end transparency and switching are still roadmap features.

The availability of WSS as components at affordable prices and the reduction of Photonic Cross Connect (PXC) per port cost are marking a turning point in the WDM market. It will provide the advanced switching features, needed to build transparent, or at least hybrid transparent WDM networks (see [1] for definition). It is now more a matter of industrializing these functions and developing the associated software.

At the same time, France Telecom WDM network is undergoing major changes. Extended Long Haul (ELH) systems have been introduced to lower the cost of the overall network, while increasing the capacity of the network, using systems of higher capacity (80 to 96 versus 32 channels) and generalizing 10 Gbps usage on the overall core network.

However, as for most of the “Legacy” Service Providers, many existing links are still equipped with systems of previous generation. They need to be replaced for different reasons:

- They have reached their maximum capacity.
- The systems have become End Of Life and generate high maintenance costs.
- They were purchased several years ago, at the cost of the market at that time and need to be disinvested because they generate very high cost through the business rates.

It is then rather important to manage correctly the upgrade of the legacy core network, considering all the coming WDM system evolutions and benchmarking all the possible solutions.

The remainder of this paper is organized as follows: architecture alternatives are presented in Section 2, the architecture benchmarking in Section 3. The challenges of migration are discussed in Section 4. Finally, Section 5 concludes this study.

2. Architecture alternatives

The latest evolutions of WDM equipments allow for building more efficient networks. As a first step, we will review the various options that can be selected to deploy a WDM core network, and evaluate their advantages and drawbacks.

Opaque network

In a pure opaque network, the signal is regenerated electronically in each node where the traffic is accessed. Depending on the option selected to perform regeneration, we can separate different classes of opaque networks: it can be implemented through back to back transponders or through an external equipment of a higher layer with or without colored interfaces (Router, Carrier Ethernet switch, OEO Optical Cross Connect...).

As far as ELH and Reconfigurable Optical Add Drop Multiplexer (ROADM) technologies allow for optically bypassing nodes in a very economic way, this architectural option might appear as completely obsolete. However, one shall note that some manufacturers have worked on reducing the cost of regeneration through the massive integration of optical components, making these configurations cost effective [4]. The benefits come then from the ability of such systems to provide switching at the electrical layer, releasing engineering constraints.

Transparent network

In a pure transparent network, the signals are transmitted end-to-end without any electronic regeneration. Wavelength channels may be switched through optical devices (WSS/PXC...) or manually, using specific high density patch panels. Suppressing regeneration shall reduce dramatically the cost of the overall network as far as the cost associated with the enhancement of the system performances, extending its reach, remains acceptable [2].

In practice, it is almost impossible to build a fully transparent network at the scale of a European network, especially when an existing legacy infrastructure is in place (old cables with large PMD values).

Hybrid network

In the framework of this study, transparent and opaque architectures must just be considered as reference models. The evolution of the WDM equipments allows now for building hybrid architectures that correspond to intermediate solutions, where part of the previously highlighted drawbacks (regeneration cost/reduced ability to tolerate large physical impairments) are circumvented using the latest technology developments.

Hybrid opaque network is an alternative of the opaque network in which ROADMs of degree 2 are deployed in some of the add/drop sites to reduce the number of regenerations and simplify network operation. It corresponds to the way most of the long haul WDM networks are built today.

Hybrid transparent network architecture is a variation of the transparent network in which intermediate OEO regenerations are allowed to relieve the cost of the network and simplify its engineering. The end-to-end connectivity is mainly achieved through optical switching, either manual or automatic.

Regenerator positioning is calculated for each path, depending on the physical impairments that it experiences, and is no more constrained by the architecture of the nodes like in an opaque network. Thus, the regenerations can be distributed on the network.

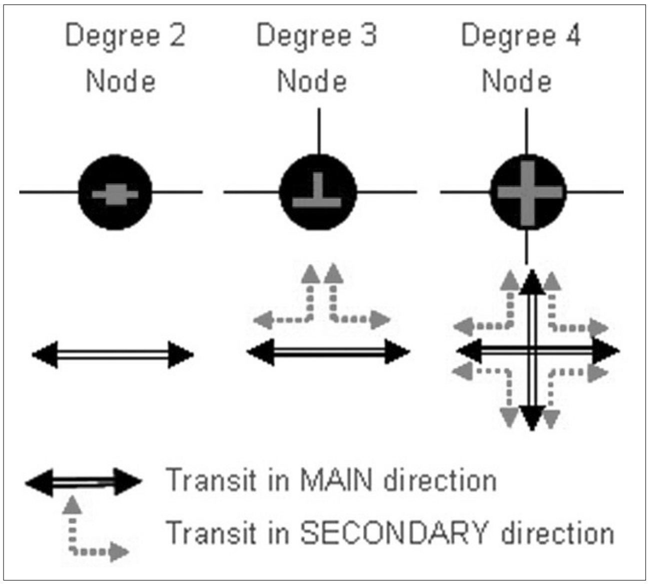
This approach does not prevent from defining specific points of regeneration in order to limit the number of nodes where regenerations are allowed. This will simplify the migration to a fully agile core network where regenerations might be needed to solve wavelength contention or performance issues after a channel has been rerouted. In this perspective, limiting the number of sites where regenerations are implemented allows a better sharing of resources, appropriately placing some pools of active and standby transponders.

3. Architectures benchmarking

A. Studied Scenarios

The most promising architectures have been benchmarked in the favorable context of a green field deployment. Five different scenarios have been defined according to the transit options. We will differentiate in the following sections two types of transit, main and secondary as depicted in *Figure 1.* Costs for main and secondary transits may be different according to scenario and equipment cost model.

Figure 1. Node transit type definition



The *first scenario "Hybrid opaque"* is used as a reference and corresponds to a hybrid opaque architecture, which can be deployed today considering existing WDM equipments. Wavelength Blockers (WB) are used to provide the transit function in main direction(s) for nodes of degree 2, 3 and 4. We consider that regenerators are used for the channels that transit from the main direction to a secondary direction.

The *second scenario, "Hybrid transparent"* corresponds to an architecture, where transits in secondary directions are performed using passive optical bypass. This function can be implemented with different elements (patchcords, optical amplifiers, multiplexers/demultiplexers...) depending on whether it is done at the band level or the wavelength level.

The *third scenario “WSS”*, is built on a hybrid transparent architecture where Wavelength Selective Switches (WSS) are used in nodes of degree 3 and 4 for the transit of the channels from one WDM link to another.

In the *fourth scenario “Fully tunable”*, we also consider the use of WSS to perform multiplexing and demultiplexing of the Add/Drop channels.

In the *fifth scenario,* the architecture of the nodes is selected node by node, so that the resulting network is the most cost effective one. Thus any configuration from the previous scenario can be selected, independently for each of the nodes.

B. Network topology

The network topology has been defined according to the existing France Telecom fiber infrastructure. All related physical constraints (site locations, attenuation, PMD, Chromatic Dispersion...) have been considered in the network model. The network model includes a total of 92 nodes spread in France and on the European Backbone Network (EBN) as follows:

Degree	1	2	3	4
Number of sites	16	48	25	3

The main characteristics of the topology are given as follows:

STUDIED NETWORK TOPOLOGY

Network size	92 nodes	99 links	
Maximum shortest path	15 hops	1982 km	
	Average	**Minimum**	**Maximum**
Linklength	226 km 3 spans	2 km 1 span	797 km 11 spans
Span length	72.91 km	2 km	129 km
Degree of node	2.15	1	4

Using the results of preliminary internal studies, the topology was designed by partitioning the network into several subnetworks. This approach allows for releasing engineering constraints, by reducing the network perimeter for each subnetwork. Since most of the traffic converges to the Paris area all subnetworks covering France include Parisian nodes. The Parisian nodes constitute gateways where regeneration allows for crossing the different subnetworks. In such subnetworks several paths allow for getting to Paris, which includes 8 nodes spread out among 4 main sites.

The proposed architecture is based on an analysis of the main traffic contributions and on the flow distribution. The subnetworks were defined according to this, in order to limit the probability for wavelength contention to happen, by reducing their perimeter and increasing their meshing level. The fiber infrastructure was also considered, so that the break points introduced by the network partitioning correspond as far as possible to the best compromise according the PMD level on the different links.

The site connectivity degree has been limited to 4, since the roadmaps provided by different WDM equipment manufacturers show a possible limitation in the short to medium term implementations.

Each demand that terminates in 2 different subnetworks shall be regenerated. There is one exception to the previous generic rule with the defined topology: 2 of the subnetworks are bridged and allow on one path some channels to cross them without being regenerated.

Figure 2 shows the topology which was used for the present study.

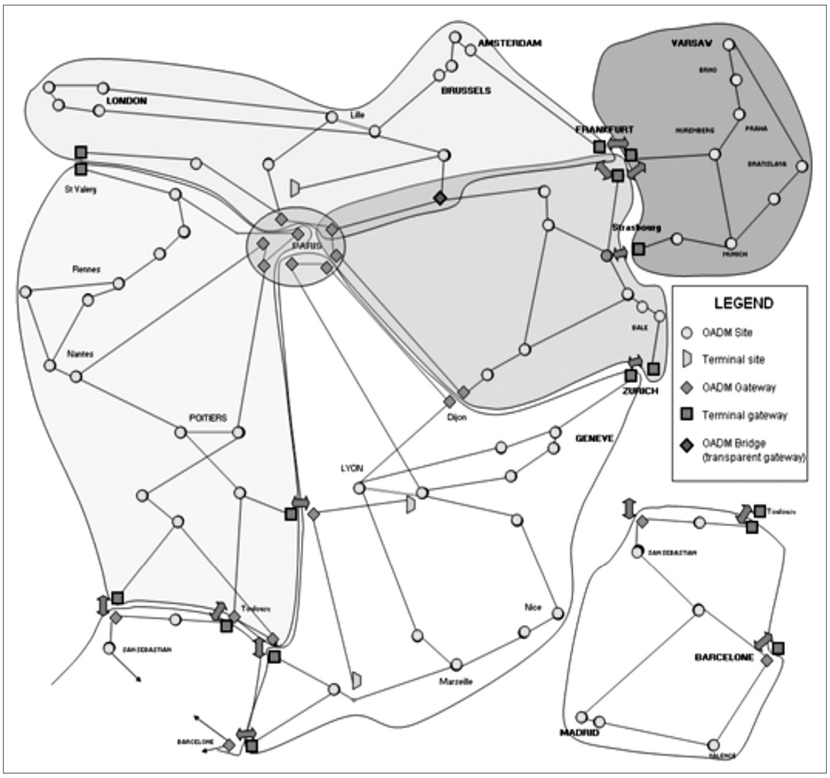
C. Traffic model

The traffic model was built considering that all the existing traffic has to be migrated (if not already supported by new systems) on the NG WDM network. The traffic contributions include ATM and SDH traffic which are quite stable, and L2 and L3 traffic that cross the core network. Part of this last corresponds to residential traffic; the other part corresponds to business services. *Figure 3* gives the traffic distribution from 2007 to 2010.

D. Cost assumptions

The costs used in this study correspond to the costs provided by the equipment manufacturers for the WDM systems sourced by France Telecom. They are revised regularly according to the contractual conditions. They include hardware, software, Installation and Commissioning (I&C) costs.

Figure 2.
Network topology: Hybrid transparent partitioned network



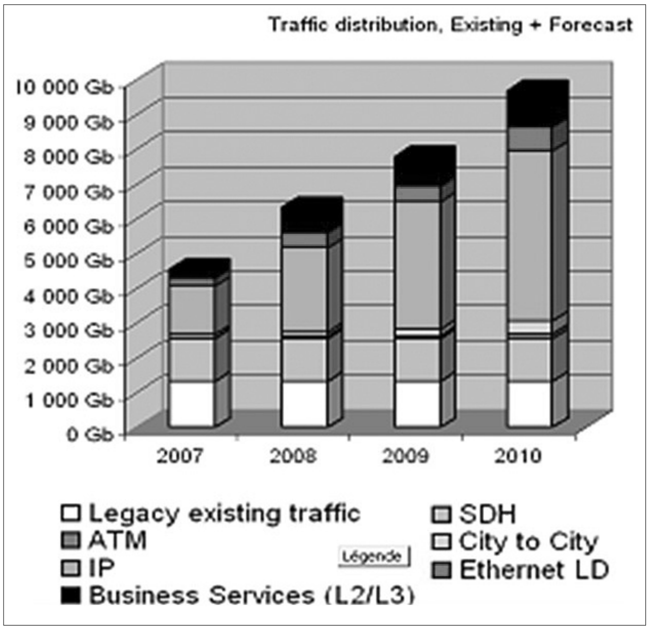
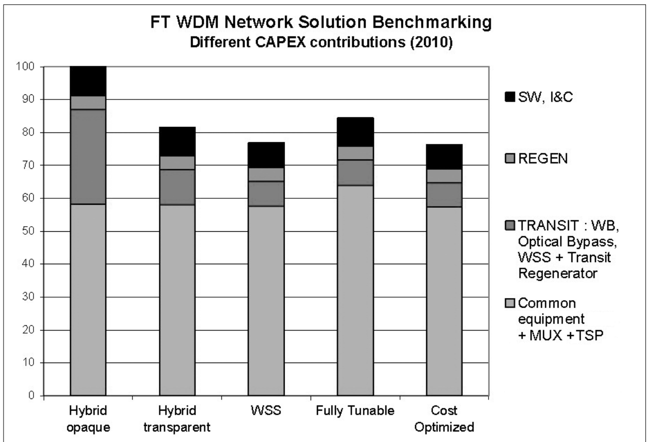


Figure 3. Traffic distribution



		Hybrid opaque	Hybrid transparent	WSS	Fully Tunable	Optimized
Traffic up to 2007		100	85	80	89	79
		Savings /Ref	15%	20%	11%	21%
Traffic up to 2010		100	82	77	84	76
		Savings /Ref	18%	23%	16%	24%

Figure 4. Comparison of the benchmarked scenarios

E. Methodology

A specific tool was designed for the purpose of this study. The inputs of this tool are the network topology and the traffic matrix. As an output, it provides for all the sites the characteristics needed to calculate the cost of the network for the different scenarios (Channel/band Mux/demux, number of transits on the main and the secondary axes, regenerators...).

The methodology uses an iterative algorithm which is schematically the following:

- Find routes (or rings for protected demands) with the least number of regenerators according to engineering limitations of equipments and nodes regeneration possibility;
- Assign wavelength(s) to each demand according to possible blocking with other demands and to band filling strategy.

Two strategies for band allocation have been considered:

- Minimization of the overall number of used bands in the network.
- Minimization of the number secondary transits handled at channel level (in opposition to transit that can be handled at the band level) in each node.

The first one represents the minimum required configuration, without any consideration on wavelength management complexity in the nodes. It is adapted to flexible networks where degree 2, 3 and 4 nodes allow for channel management at wavelength granularity (Scenario 3, 4 and 5).

The second one, on the contrary, considers wavelength management complexity in priority, to the detriment of the global number of bands used. It models the specialization of bands that is desirable when nodes are not equipped with reconfigurable structures: bands are dedicated to a specific network path.

F. Economic Assessment

The main results of the techno-economic benchmarking of the different architectures are presented in *Figure 4.*

The cost of each scenario is detailed for several contributions: common parts, modules providing transit features, regenerators, software (SW), Installation and commissioning (I&C).

From the comparison between the first and the second scenario, we see that implementing optical bypass for transit in secondary directions allows for significant savings (up to 18%) since it allows for avoiding unnecessary regenerations. Optical bypass can be performed at the channel or at the band level. Going down to the channel level implies to cascade more devices and to increase the cost of the bypass. It is also associated with an increased operational complexity. First and second scenarios were studied using the second wavelength allocation strategy, minimizing mixed direction transit in the same band. This solution corresponds to what can be deployed today.

WSS based ROADM seems to be a really interesting option since it adds flexibility in the network. Providing the ability to handle each channel path separately and remotely, with the modules deployed at day one, it eases the network planning since the wavelength allocation is no more connected to the band structure of the WDM equipments. Additionally, the third scenario brings significant CAPEX saving (5 points) in comparison with the second scenario.

The full tunability increases significantly the global cost of the network. The configuration studied here is the simplest option and corresponds to a limited number of WSS in the node. WSS are used in place of channel Mux/demux, to improve the node flexibility and reduce the commissioning time. Thus, the interest of this kind of configuration is somewhat limited, since it does not provide the advanced features needed in the perspective of migrating to a fully agile network. As far as the bands are efficiently filled thanks to the use of WSS to switch channels from one link to another (as it was with the third scenario), the overspend associated with the use of WSS to replace the multiplexers/demultiplexers is not compensated by the savings on the number of ports and the number of modules used to perform channel add and drop.

The optimized solution, where the configuration is selected on a site by site basis allows for limited savings in comparison with the 3rd scenario (WSS based ROADM). WSS based ROADM could be deployed in every site to simplify network planning & engineering for almost no additional cost. One should note that with respect to OPEX, on site commissioning is required at both ends to add transponders and connect them to the WDM and the client equipment for all scenarios. It might also be needed on intermediate site(s) if regeneration is required (back to back transponders patching).

Last but not least, the gain observed for solutions 2 and 3 slightly increases with the network load: the gain presented in *Figure 4* (for 2010) is 3 points higher than the gain observed with the traffic up to 2007.

4. Challenges foreseen in getting to a next en WDM networks

In a more and more competitive environment, the WDM network shall be as inexpensive as possible, at both the CAPEX and the OPEX level, and the flexibility shall allow for reducing production time. Reducing the total cost of ownership while increasing the flexibility is feasible but really depends on the target we set: if solution 3 allows for significant savings, further increasing the flexibility through fully tunable solutions makes these two objectives antinomic. The WDM network shall be designed considering specificities that differ from one network operator to another.

A. Challenges associated with transparency

Transponders and regenerators constitute the biggest cost item in a WDM network. The main reason for introducing transparency is to avoid unnecessary regenerations in order to reduce the total cost of ownership. Each time a channel needs to be regenerated, a break point is introduced in the network. Commissioning this channel currently implies human field operation in a location where the signal does not end. Rerouting a channel on an alternative longer path is also challenging as soon as regenerations are needed.

Nevertheless, increasing network transparency brings its lot of challenges and implies to deploy equipment with high performances.

At the equipment level, increasing the performance is a key element to reduce the network overall cost. There are three main axes we can play with: the capacity, the supported rates and the system's reach.

The highest is the capacity of the system; the lowest is the probability to get wavelength contention. Extending the capacity has a cost since it implies to improve the optical amplifier performances, or to reduce the system's power budget. Upgrading the capacity of the systems in place is most of the time traffic affecting. Another reasonable option to increase a link capacity is to add another trunk in parallel with the first one. In a hybrid transparent network, this translates in upgrading the node to a higher degree, which implies to have available ports on the WSS: 1:5 WSS may appear as rather limited considering this point, and most of the equipment manufacturers go for 1:9 components on the LH product range.

Increasing the rate of transported services helps in reducing the number of channels used in the network. It also structures the network architecture since the migration of a system to higher rates implies that the design has been made at day one considering engineering rules of tomorrow. As developed previously, adding regenerations to compensate for lower physical impairment tolerance is not a good option. In this context, the development of advanced modulation formats like DQPSK (Differential Quadrature Phase-Shift Keying) or Dual Polarization DQPSK which allow for higher tolerances in terms of CD and PMD make way for smooth migration since they solve part of the problem. Engineering constraint at 40 Gbps or even 100 Gbps could be lower than those at 10 Gbps for current technologies [7] in terms of CD and PMD tolerance. However, OSNR requirements will still be higher with this modulation format at 100 Gbps and even 40 Gbps.

Depending on their technology, cascaded WSS may degrade the signal with a wide spectral width, and their ability to support a mix of 50 GHz and 100 GHz spaced signals is an aspect to look carefully at when deploying a transparent network.

Regarding the reach, if Extended Long Haul (ELH) or Ultra Long Haul (ULH) systems have been introduced with the aim of lowering the cost of large backbone covering one or several states/country, one shall keep in mind that in the frame of highly meshed network, any possible path might be implemented. Even if the subnetworks cover apparently limited areas, the deployment of ELH or ULH systems may be justified to reduce the number of regenerations: in the case of France-Telecom EBN, using systems with a 15x25 dB reference power budget allowed for saving up to 230 regenerations compared with a system that has a 10x25 dB reference power budget.

At the network level, partitioning allows for releasing the constraint associated with the physical impairments and network engineering. This operation is complex and the main difficulty lies in sizing appropriately the subnetworks: the smallest the subnetworks are, the less constraint we get in terms of wavelength contention, and the weakest the physical impairments are. At the same time, reducing the perimeter of the subnetwork can lead to unnecessary regenerations needed when crossing 2 of them.

A part of the study was dedicated to the analysis of wavelength contention. Two simulations were launched for this purpose. In the first simulation, all the traffic from 2007 to 2010 is routed one shot, allowing for a global optimization. In the second one, the traffic is routed incrementally on a year by year basis, considering that the planning activity is made every year, using marketing forecasts.

No wavelength contention was observed in the two simulations. The link load reaches up to 81% on some links and the maximum number of bands is reached on several links. However, there was no need for channel sub-optimized rerouting or wavelength conversion to solve contention. No significant change was observed on the global cost of the network for the different solutions, but the global number of channel Mux/demux used increases by 10% from 1st to 2nd simulation. One should note that the target of suppressing secondary transit that require handling channels separately was not reached, which would translate in strong operational constraints for the solutions where no switching device is used.

A third simulation was launched in order to find the threshold where wavelength contention starts to occur. We then considered the traffic up to 2012 which forecasts a 50% growth of the pure IP traffic from year 2010. We then started to see the first signs of wavelength contention on 4 links, corresponding to the main axes of the IP hierarchal network. Whereas the link load is of 58 to 87%, no wavelength can be found for additional required paths. We can then consider that the network partitioning was rather efficient since these first signs of wavelength contention appear quite lately and concern a limited number of demands.

B. Interoperability

France Telecom, as many service providers, has a dual source purchasing policy. This contributes to let the competition act, leading to significant cost reductions while reducing the risks, in case of equipment wrong design, or procurement issues. It also guarantees the network perenniality in case one of the manufacturers is not able to meet its obligations.

However, this can lead to interoperability issues, especially in the context of a transparent mesh network, where interoperability is not only required at the transponder level. Hybrid nodes built from equipments of different manufacturers seem difficult to operate, because automatic level adjustments, gain and power control are required for a consistent end-to-end configuration and management: thus, considering nodes of degree 3 or 4 which interconnect WDM trunks from different vendors is not reasonable. Mixing the equipments of two different vendors in a transparent network would lead to a heterogeneous network where different systems cohabit in the same area with no way to make them interoperate. Partitioning solves interoperability issues as far as we consider deploying only one type of equipment in a subnetwork.

C. Flexibility

Degree 2 ROADM built with Wavelength Blockers and higher degree nodes equipped with WSS should simplify dramatically network planning, by providing switching at the wavelength level. These components remove the constraints associated with band allocation, uncoupling the wavelength grid from the band grid. A path will no more need to be strictly mapped to a band, since a band can be added and dropped several times and contains channels that do not experience the same path with this kind of architecture. ROADM can be fully configured remotely by software, simplifying network provisioning since external patching between channels Mux & Demux is no more needed at intermediate points where channels are bypassed.

Getting further in improving network flexibility and lowering the OPEX, some WDM equipment manufacturers propose to introduce TOADM, standing for “Tunable” OADM sites, extending provisioning automation up to the ends of the paths.

In traditional ROADM, multiplexing functions are still performed by Array Waveguide Gratings (AWG) devices and optical access ports are fixed: a color is assigned to each port. A way to get the full tunability is to replace the Mux/Demux by couplers or splitters and WSS. This architecture provides a full tunability of the site: tunable transponders are no more connected to a dedicated port through mux/demux. This solution has been modeled and studied in the fourth scenario. Its main advantage is that everything can be configured remotely as long as sufficient resources have been provisioned.

Figure 5 (on the next page) shows the structure of a degree 2N node. Only one transmission direction is represented on each main axis to simplify the diagram. Ii and Oi_{+1} represent WDM line’s input and outputs on main transmission directions i-i_{+1}. Amplifiers are used at both the inputs and outputs as in standard OADM sites to compensate for WDM link and site modules losses. WSS allow for switching channels from one link to another. Some splitters (after the input amplifiers) provide the broadcast function, feeding the WSS as well as the drop modules with the incoming WDM multiplex. The drop modules associated with each branch are built using splitters and 1:9 WSS. Amplifiers are required to compensate for the losses introduced by these last. The add module is built with several stages of combiners. The added resulting multiplexes are connected to the nodes outputs through the (2N-1):1WSS.

The number of WSS used is rather limited: WSS are used to provide switching from one link to another, and to replace Add/Drop multiplexers. In the simplest configuration, the number of WSS, N_{WSS}, is as follows.

$$N_{WSS} = N + \sum_{i=1}^{N} \left[Int\left(\frac{Chi}{9}\right) + 1 \right]$$

For 1x9 WSS devices used as demultiplexers, with Chi = Number of A/D channel on branch i, N the node degree.

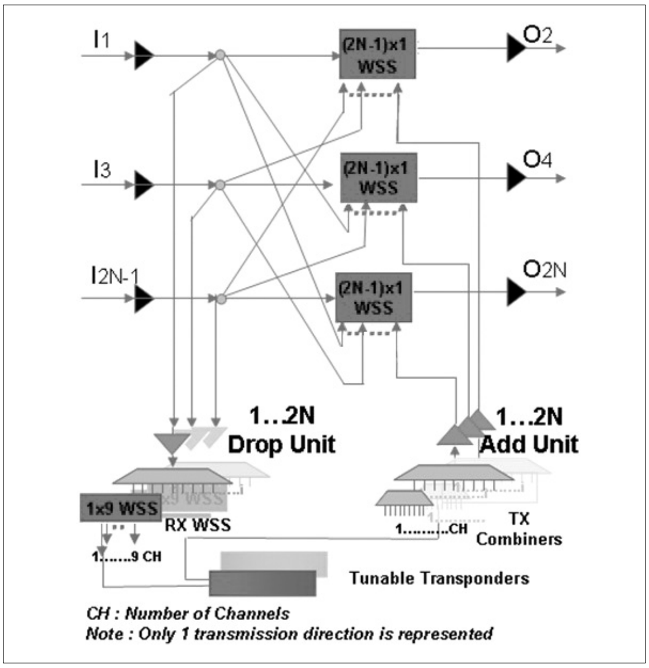


Figure 5. Degree 2N fully tunable ROADM

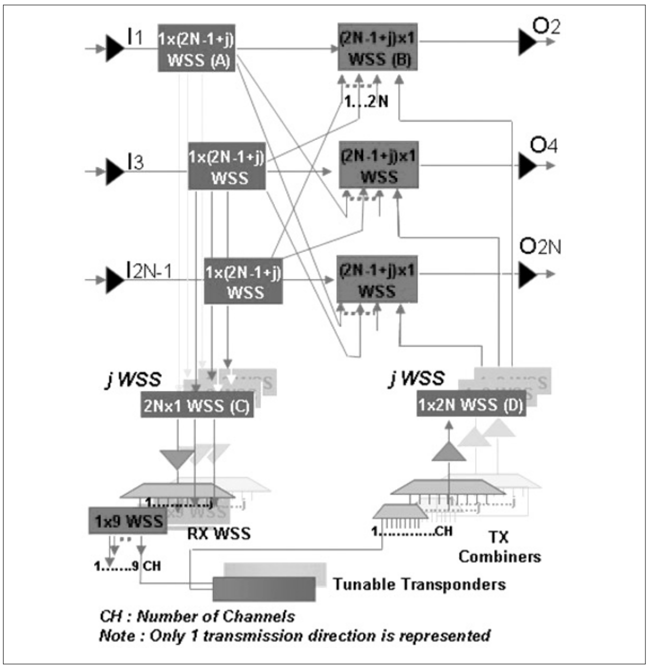


Figure 7. Directionless degree 2N fully tunable ROADM

If full tunability can help in reducing commissioning time, human on site operations are still required to connect client equipments and this configuration presents limited advantages [3]. Since a transponder is physically connected to one branch of a node; first and last hops can not be protected against failures through this configuration as shown in *Figure 6*.

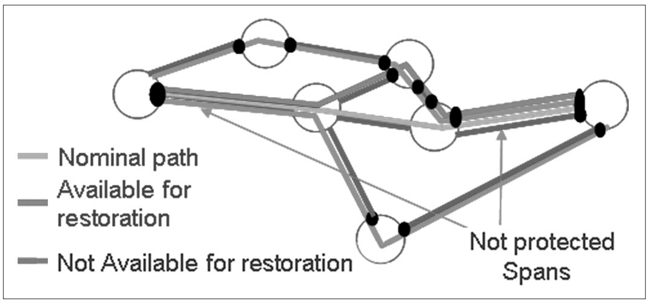


Figure 6. Protection limitations

The directionless feature allows for end to end protection or restoration. The corresponding configuration is presented in the *Figure 7.* The WSS which are added in respect to the previous configuration are represented in blue. The splitters that ensured the broadcast function are replaced by WSS (A). 2 stages of WSS are used in the Drop modules (A and C). The 1:2N WSS (D) of the Add and the Drop module provides the connectivity of the transponders to any of the node branches.

Some of the WSS are not strictly necessary since they could be replaced by combiners, but their use allows for adding flexibility and reducing losses on the different paths. One shall note that any channel must be unique at the output of C and at the input of D. Considering that the wavelengths will be provisioned sequentially and that each branch of the node will load according to the same planning rules, the probability that the same channels are used on different branches is high. Thus, we can assume that *j* will be close to 2N for the degree 2N node represented in Figure 7. This kind of architecture increases drastically the number of WSS. For a degree N node, the number of WSS will be:

$$N_{WSS} = AN + \sum_{i=1}^{N} \left[Int\left(\frac{Chi}{9}\right) + 1 \right]$$

With $A=2$ at the minimum and up to 4 if WSS are used at both the receiver (RX) side and the emitter (TX) side with $j=N$.

According to this, the extra cost for directionless full tunability could reach 45% of the global site cost, to which 15% must be added on the channel cost for each of them. The configuration becomes really complex and a high number of active equipments is crossed. It results in a significant increase of the power consumption and the footprint.

The configuration is directionless but is not fully non-blocking since a WSS and combiners used at the RX and TX side can not support twice the same channel. During a path reconfiguration (after a network failure), a transponder might be tuned to another wavelength to solve contention issues. It can not be tuned to the same wavelength as a transponder that is connected to the same WSS or combiner.

With flexible node architecture, a control plane could be used to provide shared protection or Dynamic Restoration (DR) as it is implemented in SDH networks. The savings on transport resources (capacity-distance) associated with the implementation of advanced restoration schemes at the SDH layer (VC-4) is in the order of 15 to 30% depending on the network characteristics [5,6]. Implementing DR at the WDM layer has an advantage since all the core traffic uses the WDM resources. The associated savings shall however be closer to the gain obtained with link restoration: 1 to 16% according to [5]. If protection is inherent to the SDH layer, additional modules must be added to standard WDM configurations to provide this feature, which would dramatically reduce the absolute cost savings as highlighted before. Additionally, handling dynamic restoration at the WDM layer implies to handle physical impairments. In this context, some pools of transponders need to be provisioned in almost every big site, not only for signal termination, but also for regenerations and wavelength conversions.

Focusing on costs and limiting us to the previously reviewed configurations which correspond to the solutions proposed by the different manufacturers, there is no choice but to accept that, full tunability will probably not meet savings expectations. Single stage WSS configurations will not bring dramatic OPEX savings since human field operations are still required on each end to connect client equipment, and reducing further the OPEX would imply the use of pools of transponders in most of the big sites. More complex WSS configurations add flexibility but increase dramatically the CAPEX: WSS costs will probably never reach the Mux/demux cost level.

We must then consider that a NG WDM network shall support different ranges of services. Some of them are not protected because higher layer protection mechanisms are already used. Some others shall be 1+1 protected and require a high level of SLA. The cost of the additional hardware needed to implement shared protection is significant since it implies to deploy configurations with 2 stages of WSS. However, if it does not bring the expected savings on the WDM layer, it contributes to further increasing the availability of the services. Additionally, two stage WSS configurations bring the possibility to reduce production lead-times, increasing the network operator's responsiveness.

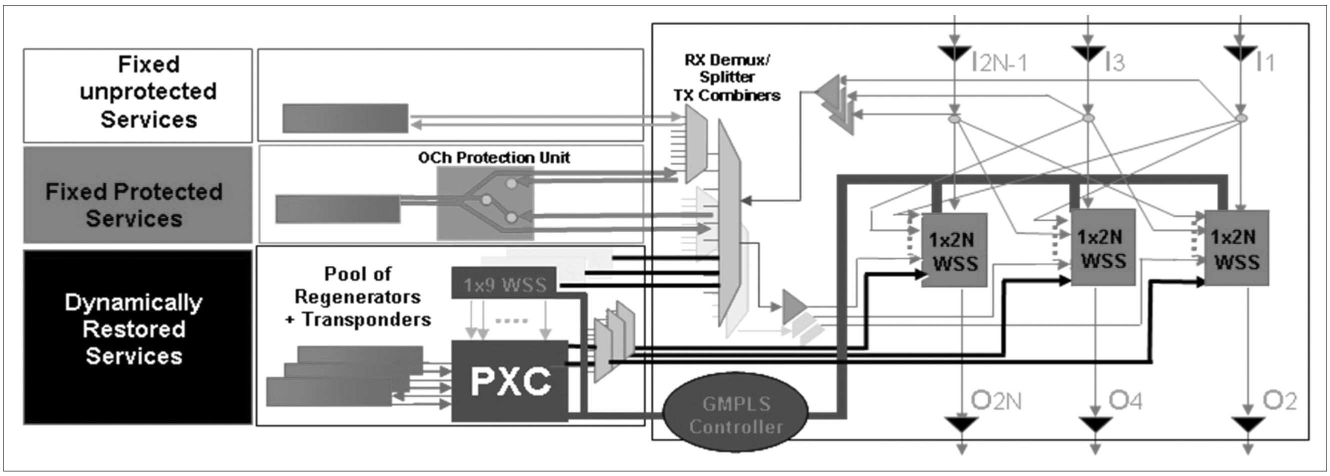
It seems then advisable to investigate alternative options to build more cost effective architectures considering that not all the services require the same treatment.

Figure 8 shows an example of a configuration that tries to minimize the number of costly units. It is based on a broadcast and select model.

With this architecture, unprotected services and 1+1 protected services are supported on a traditional configuration where fixed multiplexers/demultiplexers provides the add/drop capability. 1+1 transponder protection is performed through regular unit including a splitter and a switch, widely available through the different manufacturers. WSS units are used to provide the ROADM feature, as soon as the node degree exceeds 2. For dynamic services which shall represent a limited percentage of the whole services, splitters and WSS at RX side and combiners at the TX side are used to guarantee the colorless feature. A PXC connects a pool of transponders to the RX WSS and the TX combiners. Some additional amplifiers may be required to compensate for the insertion losses introduced by the different modules. The PXC sees any direction, and any transponders. It allows for saving a number of cascaded WSS (the number of WSS corresponds here to a one stage configuration). As far as we consider that not all the services pass through the PXC, its size maybe rather limited: a 64x64 PXC would handle 10% of the traffic in a degree 4 node. Using external combiners and splitters, we can also consider that this PXC will be closer to an automated patch-panel than to a complex PXC. The pool of transponders may be split in a pool of client transponders and a pool of transponders connected back to back through their client interfaces, acting as regenerators.

This does not prevent from partitioning the network in order to limit the number of regenerations and wavelength conversions used to face contention. The insertion loss of some of the components may challenge the engineering but all the components used in this configuration already exist and can be provided by optical component manufacturers. Thus providing this kind of configuration is more a matter of integration. A control plane shall be added in order to provide provisioning automation, and handle the potential dynamicity of the resulting configuration.

Figure 8. A possible option to provide partially tunable and directionless architecture



5. Conclusions

The implementation of transparent network at the scale of a European backbone remains utopian and NG WDM network architecture will be closer to hybrid transparent solutions, where regenerations are distributed in some predefined locations. In this framework, network partitioning will help solving wavelength contention, while releasing engineering constraints.

WSS used to switch channels from one link to another will efficiently complement degree 2 ROADM to add flexibility in the network. They could also help in reducing the total cost of ownership.

However, building a fully agile network which benefits from a control plane introduction will probably imply to consider new site architectures where the number of WSS is limited, to take up the challenge of keeping a reasonable total cost of ownership.

Acknowledgment

Authors thank Hubert Poignant, Frederic Neddam and Patrice Robert for their contribution in this study.

Authors

ESTHER LE ROUZIC received the telecommunications degree from Ecole Nationale Supérieure des Télécommunications de Bretagne, France, the M.S. degree from the University College London, U.K., in 1996, and the Ph.D. degree in electronics and communications from Ecole Nationale Supèrieure des Télécommunications, France, in 1999. She joined France Telecom R&D in Lannion in 2000, where she has been working on wavelength division multiplexing (WDM) networks, optical functions, optical transmission, and translucent networks. She is involved in national and European projects dealing with optics and authoredor co-authored about 25 national or international papers or communications.

GAËLLE YVEN received the degree in electronics engineering from ISEB, Institut Supérieur d'Electronique de Bretagne, France, in 2001. She began her career in the R&D Labs of Itis in Rennes, and worked on as a software Engineer on DAB (radio) and DVB (video) devices. She joined France Telecom in 2003 has a telecommunications engineer. She first participated to acces devices validation and then worked on defining the architecture of the Core and Backhaul network. She is now involved in defining the architecture of the network in order to support new audiovisual services.

OLIVIER RENAIS received the degree in optronics engineering from ENSSAT, Ecole Nationale Supérieure de Science Appliquées et de Technologies of Lannion, France, in 1992. He began his career in the optical Labs of SVFO Pirelli, and worked with different optical equipment manufacturers (Pirelli, Nortel and CISCO) as a system Engineer. He also collaborated to the network architecture definition of different NSP and Cable operators acting as a consultant. He joined France Telecom in 2005 has a Transport network architect. He has participated to studies defining the architecture of the Core transport network including NG-SDH OXC and WDM equipments.

References

[1] Annalisa Morea, Julien Poirrier, "Strategies for the Migration from Opaque to Hybrid Networks", ECOC2007.

[2] Annalisa Morea, Julien Poirrier, " A critical analysis of the possible cost savings of translucent networks", DRCN 2005.

[3] Lei Zong et al., "Study on wavelength cross-connect realized with wavelength selective switches", OFC-NFOEC 2006.

[4] Serge Melle et al., "Network planning and Economic Analysis of an Innovative New Optical Transport Architecture: the Digital Optical Network", OFC-NFOEC 2005.

[5] Doshi et al., " A systematic Study Of Next Generation Optical Core Network: Are flatter Network Better?", NFOEFC 2001.

[6] Daniel R Jeske, Ashwin Sampath, Restoration Strategies In Mesh Optical Networks: Cost Vs. Service Availability", PRDC06.

[7] H. Büloz et al., "40 Gb/s Distortion Mitigation and DSP based Equalisation, ECOC 2007.

All-optical networks and switching technologies for a 3D videoconference system with the feeling of presence

ANDREA BLANCO REDONDO, ENRIQUE AREIZAGA SÁNCHEZ
Tecnalia-Telecom, Broadband Networks Group
{ablanco, enrique}@robotiker.es
JOSEBA ZUBÍA ZABALLA
University of the Basque Country, Department of Electronics and Telecommunications
joseba.zubia@ehu.es

Keywords: all-optical network; ultra-low latency; optical burst switching; optical packet switching; videoconference

New ultra-low latency optical transmission and switching technologies are discussed in this paper. This work, carried out within the VISION project framework, sets the basis for building an ultra-low latency network for a high-quality videoconference system capable to provide a real feeling of presence, including 3D representation. The need of a full-optical transport network to achieve this aim is justified along the paper. The appropriate evolution of the switching techniques from optical circuit switching to optical packet switching is reasoned and the challenges hindering this evolution are stated. At a lower level, the technologies and materials available to perform optical switching, fulfilling the specific requirements, are reviewed. We present initial scalability simulations of electro-optic switches and future research in this field is posed. Finally, we propose and explain a design of a large optical packet switch employing optical labels.

1. Introduction

The need for high bandwidth is inherent to video data transmission. However, new generation videoconference systems will be far more demanding than current real-time video transmissions. A high-quality videoconference system capable of providing a real feeling of presence, including 3D representation, requires much more than a high capacity network. Briefly, such a system works as follows: a significant number of cameras capture the video images from different angles. Next, all the fluxes of data generated from the capture of video must be multiplexed with the capture of audio, also carried out from different positions. These audiovisual data have to be transported to the other end of the communication together with block matching and processing information [1].

It is not only the large volumes of data that must be transported which makes the design of the transport network challenging; videoconference applications, and specially such a quality demanding type, are extremely sensible to latency and jitter issues. Note, for instance, that video and data are coded and sent separately and thus a loss of synchronism at the receiver results in a low quality reproduction and spoils the feeling of presence. Some actual videoconferencing test results have proved that, to achieve a good performance, the end-to-end delay must be kept under 150 ms and the jitter, or variations in the delay, must be less than 20 ms [2]. Therefore, the goal when designing such a transport network is to achieve ultra-low latency, i.e. end-to-end latency practically equal to the propagation time, and to make jitter as small as possible.

In Section 2, the need for an all-optical network to fulfil these requirements is justified. Section 3 introduces the optical switching paradigms that mark the phases of VISION project. The switching mechanisms and technologies enabling those paradigms are reviewed in Section 4, and finally, conclusions are presented in Section 5.

2. The need for an optical network

Most of the currently deployed transport networks use optical technology just as a transmission medium, while the switching functions are performed in the electrical domain [3]. This fact underlies some limitations derived from the conversion of the optical signal to the electrical domain: a bottleneck arises in the intermediate nodes due to the practical limitation of electronics to 40 Gbps; on the other hand, conversions to the electrical domain themselves increase the complexity and cost of the nodes, and entail delays which penalize the end-to-end latency to an intolerable extent [4].

In the last decade, there has been a tendency to develop and implement the optical layer beyond its original transmission function. Ultra-low latency and high bandwidth features required for next-generation video communications can only be achieved by full-optical transport networks in which switching is performed in the optical domain [5]. Optical switching enables, in turn, the elimination of certain layers from the traditional model, IP/ATM/SDH/WDM. This model overloads networks with redundant information in the headers inserted at the different layers. It also leads to a higher complexity in equipment and control functions. The increasing penetration of optical technologies into network functions reduces costs and power consumption. On the other hand, it makes management and control plane implementation lighter: it enables models such as IP/Ethernet/WDM or IP (GMPLS)/WDM in which it is possible to conceive a unique distributed control plane [6].

3. Optical switching paradigms

3.1 Optical circuit switching – Phase A

Optical circuit switching (OCS) relies on the reservation of a fixed optical bandwidth, i.e. one or more wavelengths, to establish dedicated optical circuits over the network. The process consists of three phases: circuit set-up, data transmission, and circuit tear-down. One of the main characteristics of OCS is the two-way reservation process when setting up the circuit. This fact makes OCS suitable only for the cases in which the connection duration is long relative to the path set-up time. On the other hand, circuit switching results in low bandwidth utilization if the traffic to be supported is bursty [7]. In videoconference systems connections remain established for a relatively long time and the nature of the traffic is continuous, and hence this switching paradigm results a suitable choice. Phase A of our photonic network is based on the OCS paradigm.

In most of OCS current commercial implementations, optical circuits are permanent or semi-permanent. Nevertheless, bandwidth required in such a videoconference system is high and the number of wavelengths available in the network is limited. Consequently, our network should be provided with more intelligence and flexibility to reduce the overprovision of bandwidth and enable some QoS capabilities.

In VISION, an OCS network with the capability to provide bandwidth dynamically has been developed [8]. It is a meshed network with a distributed control plane based on GMPLS (Generalized Multi-protocol Label Switching) [9,10], as schematically represented in *Figure 1/a*, whose edge nodes have the building blocks depicted in *Figure 1/b*. This implies not only a better use of the available wavelengths, but also efficient and quick restoration of the services, the possibility of offering new services, as bandwidth on demand, and a range of traffic engineering mechanisms.

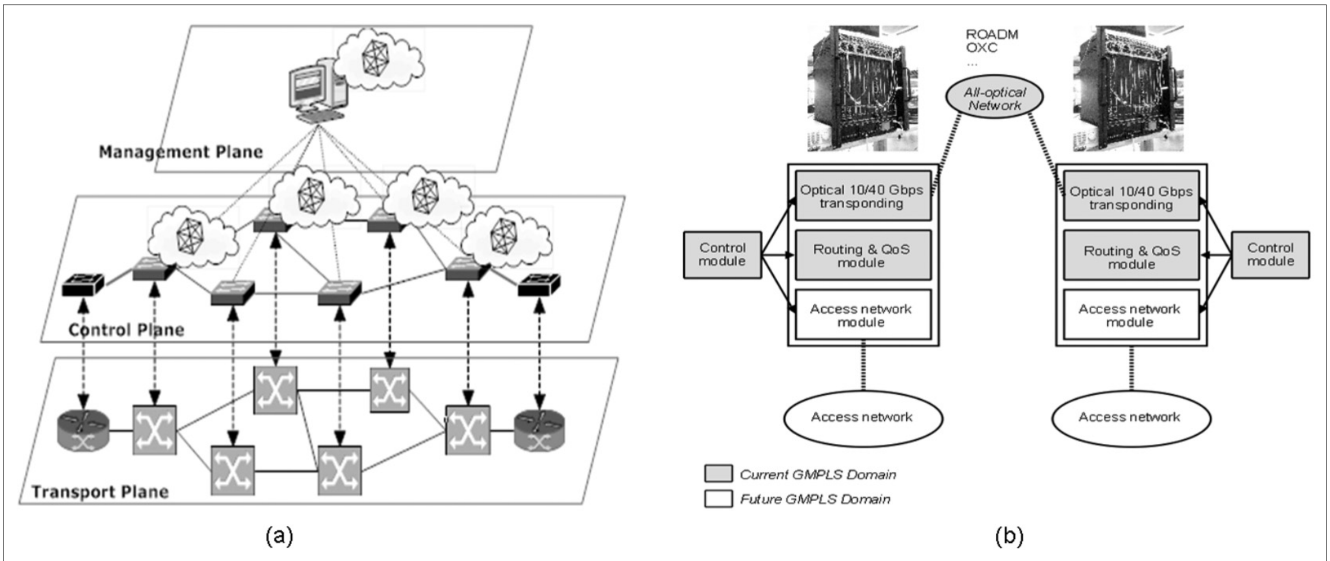
3.2 Optical burst switching – Phase B

Although OCS seems to be appropriate for providing enough bandwidth and low-latency for the targeted application, it is not cost-effective. A flexibility enhancement is required to adapt transport networks to the necessities of today's traffic [11]. Optical packet switching (OPS) is the paradigm we pursue. Nevertheless, significant advances in optical technology, both in switching mechanisms and buffering techniques, must be achieved before efficient OPS networks are made feasible. A medium-term approach is optical burst switching (OBS).

In the edge nodes of OBS networks, several optical packets are grouped into a burst. Once the burst is ready, a control packet is sent to "reserve" bandwidth and set switches along a path for transporting the data contained in the burst. The term reservation is not exact in this case, since the optical path that will be used it is not uniquely associated to a certain burst. There are several protocols to carry out the set-up process but all of them have common features: they use a one-way process, and more importantly, bursts can cut-through switches instead of being stored and then forwarded [7].

The one-way set-up process reduces vastly the set-up time relative to OCS. Burst lengths are generally of the order of tens of microseconds; hence switching times are required to be less than 1 μs, which is fast but realizable nowadays. However, the issues of how to deal with conflict situations and reduce burst drop-

Figure 1. Distributed control-plane based network: (a) network architecture and (b) edge nodes building blocks.



ping must be addressed. Optical buffering constitutes the ideal solution to retain the burst until the appropriate output is empty [12]. Nevertheless, this is not a requirement for OBS.

The design modularity of the network architecture developed in Phase A enables the gradual evolution to OBS. For instance, migration starts by introducing a fast optical cross-connect (OXC) instead of a millisecond range OXC.

However, the physical plane is not the only issue to consider in this evolution. The migration of the control plane from OCS to OBS is not trivial. Complexity grows tremendously, since optical packets are not sent anymore through predetermined dedicated circuits. Aspects like routing, signalling, or QoS, must be considered now for every burst and GMPLS is clearly dedicated to circuit management.

Many studies have dealt with the operation, architecture and signalling of OBS nodes, as in [13] or [14]. Nevertheless, OBS networks are still in an experimental stage. We are working on fast high-capacity switching matrices. Concerning the control plane, an effort is being done on adapting GMPLS to statistical multiplexing.

3.3 Optical packet switching – Phase C

Optical packet switching provides the highest degree of flexibility and efficiency to our network. However, there are some serious issues that complicate the deployment of an all-optical packet network and these are, as stated above, the practical realization of photonic buffers and the availability of ultra-fast optical switching matrices.

In OPS every optical packet is routed independently of the others. The packet has to be processed in every intermediate node in order to configure the switch. The packet header has to be analyzed while the payload needs to be temporarily stored in the optical domain. This store-and-forward way of switching entails the major problem of OPS, since it is not possible to buffer the light as it is done with electrical signals [15]. An approach for "buffering" light packets relays on employing coils of fibre to take advantage of the propagation delay, referred to as fibre delay lines. This solution is thorny and inflexible, since the fibre delay lines take up a lot of space and the propagation delay, and thus the storing time, is fixed. We focus our research in an exciting alternative: slow light devices. So far, some studies have claimed that intrinsic limitations will impede their use as buffers, as in [16]. However, some of the most recent works in this field [17-19] show promising results that indicate that it is just a matter of time these devices become a practical solution for optical buffering.

We have also considered the possibility of a "bufferless" optical packet network, e.g. employing wavelength conversion or the hot-potato or deflection algorithm, like in [20-22]. Regarding the reduced applicability of these techniques, we center our research upon slow light as a promising buffering approach to enable the evolution to OPS [23,24].

Photonic "buffering" is not the only issue in OPS. Even if storing light was made feasible, the optical switches must be incredibly fast. The switching times have to be small compared to the packet length. The transmission time of an IP packet at 40 Gbps can be between 10 ns and 300 ns depending on the length. Therefore, in order to avoid inefficiencies in utilization of the transmission channel, the switching time of the optical switch fabric needs to be around 1 ns or less [25]. The number of technologies capable to achieve this goal is even more limited than in OBS.

4. Optical switching mechanisms and materials

Optical switching operation typically occurs after a stimulus is applied. The stimulus is generated by a control module based on the analysis of signalling information, in the case of OCS and OBS, or based on the packet header, in the case of OPS [26]. Consequently, the design and optimization of an optical switch entails two separate blocks equally important: the optical switch fabric and the control processing part.

Several physical effects can unleash the switching operation in certain materials. In any case, the switching process influences the optical signal quality inserting losses, and other effects, such as polarization dependence and crosstalk, may occur. It also entails a switching delay that must be taken into account. Depending on the switching paradigm and on the size of the switch fabric, the upper limits regarding switching time and losses will vary. Therefore, the switching mechanism and the materials used to build a switching matrix must be chosen considering those particular restrictions.

4.1 Mechanical effect – MEMS

MEMS optical switches consist on a 2D or 3D matrix of micro-mirrors which connects the input fibres with the outputs.

They are a mature technology which presents some brilliant features, such as low insertion losses, low polarization dependent losses, and high cross-talk suppression. It is noticeable too that they enable cost-effective large switching matrices. However, taking into account our pursued evolution to OBS and OPS, we come across MEMS' biggest limitation: the switching time. The switching speeds reported for MEMS have been in the range of a few microseconds to milliseconds [27-30]. Research has been done to reduce the angle through which light is bent and switching times of hundreds of nanoseconds have been achieved very recently [31]. These times are suitable for applications using permanent or semi-permanent circuits, and even burst switching in a future, according to the latest results. Nonetheless, faster alternatives must be considered for implementing packet switches.

4.2 Electro-optic effect

An electro-optic (EO) effect is a change in the optical properties of the material, generally the absorption or the refractive index (*n*), in response to an electrical field that varies slowly compared with the frequency of light. Pockels effect, particularly strong in ferroelectric materials, is the most relevant to build EO switches. It relies on the variation of n by applied electric fields.

The EO effect is very fast and therefore high-speed optical communication devices are expected to be realized exploiting it. As a consequence of the change of *n*, the signal phase is modified. Coupled waveguides or Mach-Zender interferometers (MZI) are simple ways to build a switch module based on the EO effect.

Amongst crystals available to develop these devices, lithium niobate ($LiNbO_3$) presents many winning points: it is a low-loss ferroelectric crystal with very fast response to the EO effect, 1 ns, and also a stable technology; fabrication techniques have been well established for this material, as single growth techniques and waveguide fabrication using Ti diffusion or proton exchange technologies [26]. Nevertheless, $LiNbO_3$ presents some drawbacks to take into account, as a high-driving voltage, polarization dependence and DC drift.

Other alternatives have been studied to build electro-optic switches based on different materials than lithium niobate. Lead zirconate titanate (PZT) is a ceramic material with a higher EO coefficient than $LiNbO_3$, which leads to a reduction of the driving voltage [32]. PZT is nearly polarization insensitive and it presents also low loss values. Unfortunately, it counts with some important limitations: it is incapable to match the fast response of $LiNbO_3$; the fabrication processes are complicated; and it is an expensive technology.

Polymers are a low-cost option for building electro-optic switches. Devices made out of polymers achieve the higher operation speeds. However, so far their losses are too high to build efficiently operative switches and they suffer from instability which can result in serious problems when performing the switching function [26].

We consider switches based on the EO effect as a promising alternative for the development of small switch fabrics in OBS and OPS networks, particularly due to their fast response. Our efforts are focused in solving their drawbacks, especially those related to their low scalability. *Figure 2* shows some illustrative results of the scalability simulations we have performed. An 8x8 EO switch was built from 2x2 lithium niobate MZI modules, following Spanke-Benes architecture. One can see how the amplitude of the optical field decreases as a consequence of the insertion losses introduced by every 2x2 module in the path from the input to the desired output waveguide. Losses of current commercial 2x2 $LiNbO_3$ switches are on the order of 5 dB, which certainly entails a problem when concatenating many of them. On the other hand, the residual amplitude observed in the rest of output waveguides is increased, leading to a degradation of the extinction relations.

It is noticeable, looking at the differences between *Figure 2/b* and *2/c*, how neither the losses nor the extinction relations are homogeneous. They depend on the path followed by the optical field. This fact complicates

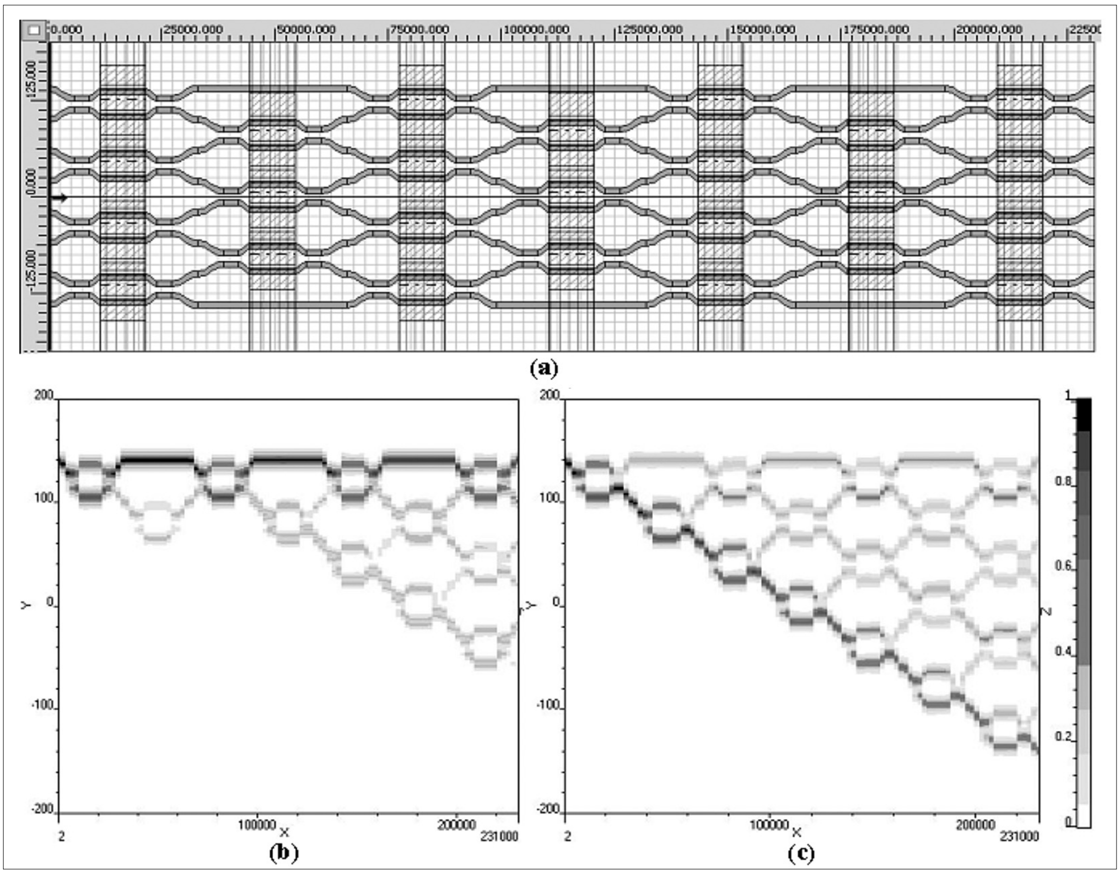


*Figure 2.
8x8 lithium niobate electro-optic switch:
(a) layout, in units of microns
(b) XZ slice of the optical field amplitude for the shortest path
(c) for the longest path.*

network design in aspects such as the gain equalization of optical amplifiers or the dynamic range of the receivers. Other architectures, such as Benes or Spanke, homogenize these parameters. However, they require waveguide crossovers, making it difficult to fabricate in integrated optics and leads to serious crosstalk issues.

The focus of our future research in this realm will be set on diminishing losses and improving extinction relation.

4.3 Semiconductor optical amplifier gain modulation

Semiconductor optical amplifiers (SOAs) are well-known and commercially available. They are essentially laser diodes without end mirrors, which have fibre attached to their ends. They are attractive to build all-optical logic gates. Instead of inserting losses, SOAs are able to provide gain to the propagating signals if they are correctly polarized. Moreover, they allow photonic integration due to their small size and inner structure.

A SOA can be used as an on-off switch by varying the bias voltage applied.

The gain of the SOA can also be optically modulated taking advantage of several nonlinear effects. For instance interferometric switches based on cross-phase modulation in combination with XGM in SOAs lead to an improvement of the extinction ratio [33].

The switching times of current switched SOAs are of the order of 100 ps. Much faster switching times can be achieved placing the SOA in a nonlinear loop mirror, and also employing two-photon absorption and free carrier absorption in combination with ultra-fast carrier cooling for the SOA recovery. These techniques achieve switching time faster than 1 ps, but they entail complications so far unsolvable [34].

The SOA-based switching techniques explained above constitute optical gates, i.e. 1x1 switches. To build larger optical switches, structures composed of couplers and SOAs are required, as shown in *Figure 3*. The fabrication of these structures results very expensive and hence just very small switch fabrics are expected to be built on the basis of SOA technologies.

4.4 Arrayed waveguide gratings

So far only MEMS-based switches were easily scalable to large switching matrices. Arrayed waveguide gratings (AWGs), also known as phased-array waveguide gratings, enable the realization of large optical switches with a single compact module.

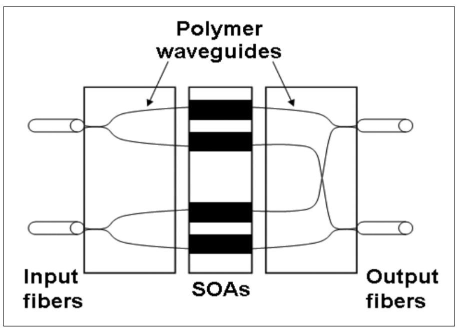


Figure 3. 2x2 hybrid SOA switch module

AWGs consist of two couplers connected by an array of waveguides whose different lengths are carefully chosen. They always route a certain incoming wavelength from a determined input to a determined output. To provide some control functionalities over the switch, reconfiguration properties must be added to the device.

Changing the input wavelength to the AWG in response to a control stimulus makes it possible to guide the signal to the desired output. Several research works have studied this type of solutions, e.g. the WASPNET (Wavelength Switched Packet NETworks) proposes an AWG-based optical packet switch using tuneable wavelength converters [35]. Tuneable lasers have also been considered for this purpose, as in [36], where wavelength accuracy and time stability is of outmost importance.

5. A proposed design for an optical packet switch

In *Figure 4* we propose an AWG-based optical packet switch. Control information will travel with the optical packet in the form of an optical label. Labels are processed and according to the stimulus generated by the control subsystem, buffers and label swappers are configured. No additional delay is introduced by the AWG and thence the switching operation time depends only on the label processing speed and on the switching time of the buffers and label swappers.

The electronic circuitry must perform the control functions within a range of time valid for OPS networks, of the order of 1 ns as stated in Section 2. This involves a big challenge for current electronics. The introduction of optical labels simplifies the electronic processing and thence contributes to the feasibility of OPS routers. Labels may also be processed optically in a future, which would reduce the requirements over optical buffering.

6. Conclusions

A 3D videoconference system with feeling of presence has strong requirements regarding latency, jitter, and bandwidth. These requirements can only be satisfied by an all-optical network.

An OCS based network with a GMPLS control plane has been implemented and current efforts are focused on enabling its evolution to OBS and OPS, both in the physical and control plane. Switching mechanisms have been discussed in this paper. Electro-optic switches represent a valid technology for OBS and OPS small switches. Other schemes, such as those including AWGs are required when talking of large fabrics.

We have presented a design for a large optical packet switch based on AWGs and label switching. Labels are capable of reducing control processing time and thus reduce the buffering requirements.

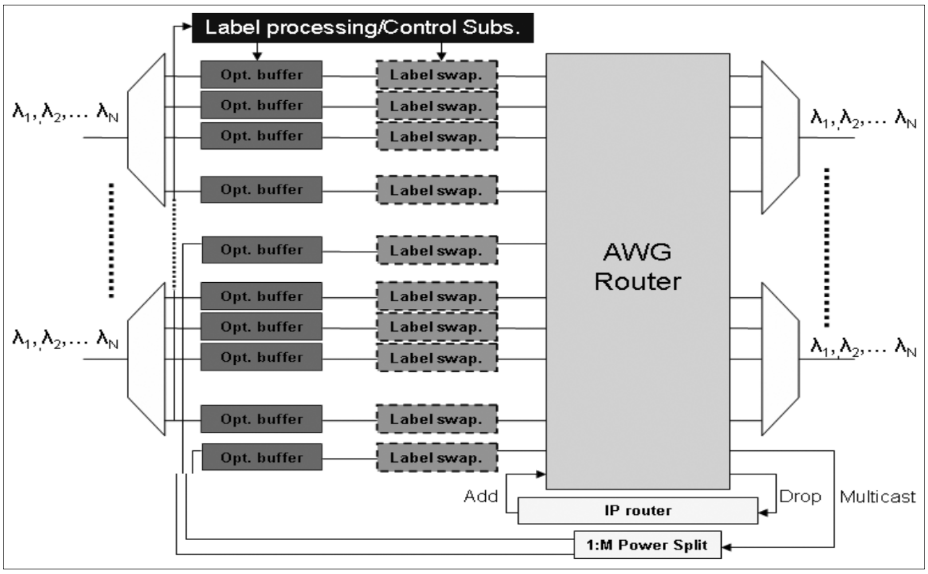


Figure 4. AWG-based optical packet switch

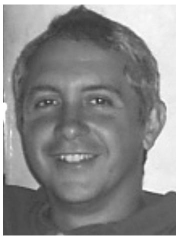
Along the VISION project we perform exhaustive research in two specific realms: electro-optic switches and photonic buffers. Hopefully, the results of our work will enable the realization of a practical optical packet switch and will contribute to the transition to OPS networks.

Acknowledgements

This work has been partially supported by the Spanish Administration organisation CDTI, under project CENIT-VISION 2007-1007.

Authors

ANDREA BLANCO received her M.Sc degree in Telecommunications Engineering in 2007 from the University of Valladolid, Spain. She spent a year, as a grant holder, performing research in nonlinear optics at Aston University, Birmingham, United Kingdom. She is currently the responsible for the photonic group of Tecnalia-Telecom and leads several projects involving disruptive photonic technologies and all-optical networks. She is also working on her Ph. D. thesis on optical memories at the Univesity of the Basque Country, Bilbao, Spain. Her research interests include all-optical switching, optical memories, photonic sensors, and optical packet switching.



ENRIQUE AREIZAGA is the Broadband Networks manager at Tecnalia-Telecom. He is responsible for directing and leading the company's business in the broadband markets. He supervises the Tecnalia-Telecom's strategic cooperation with carriers around the world and leads several projects with leading carriers defining future broadband access solutions. He received his M.Sc. degree in Physics in 1989 from the Basque Country University, Spain. He spent the year 1989 as grant holder at the Dr. Neher Laboratories (currently TNO) in The Netherlands. He has more than 20 years of professional experience in the telecommunications industry. He has collaborated as reviewer and evaluator in several national and international programs, and has coauthored more than 20 conference papers.

JOSEBA ZUBIA received the M. Sc. degree in solid-state physics and the Ph. D. degree in physics from the University of the Basque Country, Bilbao, Spain, in 1988 and 1993, respectively. His Ph. D. work focused on the optical properties of ferroelectric liquid crystals. He is currently a Full Professor with the Department of Electronics and Telecommunications, School of Engineering of Bilbao, University of the Basque Country. He has more than 12 years of experience doing basic research in the field of polymer optical fibres and is currently involved in research projects in collaboration with universities and companies from Spain and other countries in the field of polymer optical fibres, fibre-optic sensors, and liquid crystals. Prof. Zubia was a recipient of a special award for Best Thesis in 1995.

References

[1] Kauff, P. and Schreer, O., "An inmersive 3D video-conferencing system based on a shared virtual environment," Proc. Int. Conf. on Media Futures, Florence, May 2001.

[2] Calyam, P., Sridharan, M., Mandrawa, W., Shopis, P., "Performance measurement and analysis of H.323 traffic," PAM Workshop, April 2004.

[3] Agrawal, G.P., Fibre-Optic Communication Systems, 3rd ed., John Wiley & Sons, pp.1–18., 2002.

[4] Goff, D.R., Fibre Optic Reference Guide, 3rd ed., Woburn Massachusetts: Focal Press, 2002.

[5] Durresi, A., Jain, R., Chandhok, N., Jagannathan, R., Seetharaman, S., Vinoddkrishnan, K., "IP over all-optical networks – Issues," Proc. IEEE Glob. Telecom. Conf., San Antonio, Vol. 4, November 2001.

[6] Banerjee, A., Drake, J. P. Lang, G., Turner, B., Kompella, K., Rekhter, Y., "Generalized multiprotocol label switching: an overview of routing and management enhancements," IEEE Com. Mag., Vol. 39, pp.144–150., Januar 2001.

[7] Qiao, C. and Yoo, M., "Choices, features and issues in optical burst switching," Opt. Net. Mag., Vol. 1, No. 2, pp.36–44., 2000.

[8] Beltrán, P., Blanco Redondo, A., Cortés, F.J., Pozo, A., "All-optical transport layer and GPON access for inmersive communications," Proc. International Conference in UltraModern Communications (ICUMT 2009), St. Petersburg, Russia, October 2009.

[9] Yoo S.J.B., "Optical-label switching, MPLS, MPLambdaS and GMPLS," Special Issue on: Prospects and Challenges: Next Generation Optical Network Architectures, Opt. Net. Mag., 2002 (invited).

[10] Iselt, A., Kirstadter, A., Chahine, R., "Bandwidth trading- a business case for ASON?," Telecom. Net. Strateg. and Plann. Symp., pp.63–68., June 2004.

[11] Xu, L., Perros, H.G., Rouskas, G., Techniques for optical packet switching and optical burst switching," IEEE Com. Mag., Vol. 39, pp. 36–142., Januar 2001.

[12] Qiao, C. and Yoo, M., "Optical burst switching – a new paradigm for an optical internet," J. on High-Speed Networks, Vol. 8, pp.69–84., 1999.

[13] Mehrotra, P., Baldine, I., Stevenson, D., Franzon, P., "Network processor design for use in optical burst switched networks", In Proc. of the International ASIC/SOC Conference, Washington, September 2001.

[14] Chang, J. and Park, C., "Efficient channel-scheduling algorithm in optical burst switching architecture", In Workshop on High Performance Switching and Routing, Kobe, 2002.

[15] Ramaswami, R. and Sivarajan, K.N., Optical Networks, A Practical Perspective, 2nd ed., Morgan Kaufmann Publishers, 2001.

[16] Burmeister, E.F., et al., "A comparison of optical buffering technologies," Optical Switching and Networking (in press), 1999.

[17] Boyd, R.W., Gauthier, D.J., Gaeta, A.L., Willner, A.E., "Maximum time delay achievable on propagation through a slow-light medium," PRA, Vol. 71, 023801, 2005.

[18] Khurgin, J.B., "Optical buffers based on slow light in EIT media and coupled resonator structures – comparative analysis," J. Opt. Soc. Am., B 22, pp.1062–1074., 2005.

[19] Tsakmakidis, K.L. and Hess, O., "Slow light in metamaterial heterostructures," Proc. of SPIE, Vol. 6904 690405-1, 2008 (invited).

[20] Busch, C., Magdon-Ismail, M., Mavronicolas, M., "Efficient bufferless packet switching on trees and leveled networks," J. of Parallel and Distributed Computing, Vol. 67, Issue 11, pp.1168–1186., November 2007.

[21] Øverby, H., "Traffic modelling of asynchronous bufferless optical packet switched networks," Computer Com., Vol. 30, Issue 6, pp.1229–1243., March 2007.

[22] Chang, C.S., Chen, W.J., Huang, H.Y., "Birkhoff – von Neumann input buffered crossbar switches," IEEE INFOCOM2000, Tel Aviv, Israel, pp.1614–1623., 2000.

[23] Blanco Redondo, A., Beltrán Pellicer, P., Zubía Zaballa, J.A., "Slow light for future all-optical packet switched networks," Proc. ICUMT 2009, St. Petersburg, Russia, October 2009.

[24] Blanco Redondo, A., Areizaga Sánchez, E., Zubía Zaballa, J.A., "Slow light for microwave photonics applications," Proc. of Mediterranean Microwave Symp. (MMS 09), Tangiers, Morocco, November 2009 (invited).

[25] Tucker, R.S., "Optical packet switching: A reality check," Optical Switching and Networking (in press), 2007.

[26] Martin, Nord, Optical Switching Technologies for Optical Line-, Burst, and Packet switching, Telenor R&D Report, 2002.

[27] Husain, A., "MEMS-based photonic switching in communication networks," OFC, [OSA Technical Digest, WX1], Baltimore, USA, March 2000.

[28] Rai-Choudhury, P., (ed.), MEMS and MOEMS: Technology and Applications, SPIE Press Monograph, 2000.

[29] Rebeiz, G.M., RF MEMS Theory, Design, and Technology. Hoboken NJ: Willey, 2003.

[30] Castañer, L.M. and Senturia, S.D., "Speed-energy optimization of electrostatic actuators based on pull-in," IEEE Journal Microelectromech. Syst.., Vol. 8, No. 3, pp.290–298., September 1999.

[31] Nielson, G.N., et al., "High-Speed, sub-pull-in voltage MEMS switching," Sandia Reports, January 2008.

[32] Kurihara, K., Kondo, M., Sato, K., Ishii, M., Wakiya, N., Shinozaki, K., "Electro-optic properties of epitaxial lead zirconate titanate films on silicon substrates," JJAP, Vol. 46, No. 10B, pp.6929–6932., 2007.

[33] Ju, H., et al., "SOA-based all-optical switch with sub-picosecond full recovery," Opt. Express, Vol.13, No. 3, Februar 2005.

[34] Yang, X., et al., "Sub-picosecond all-optical switch using a multi-quantum-well semiconductor optical amplifier," Opt. Com., Vol. 236, pp.392–334., 2004.

[35] O'Mahoney, M.J. et al., "WASPNET: A Wavelength Switched Packet Network," IEEE Com. Mag., Vol. 37, pp.120–129., March 1999.

[36] Chan, C.K., Sherman, K.L., Zirngibl, M.A., "Fast 100-channel wavelength-tuneable transmitter for optical packet switching," Electron. Lett., Vol. 13, No. 7, pp.729–731., 2001.

Reducing total call-blocking rates by flow admission control based on equality of heterogeneous traffic

SUMIKO MIYATA, KATSUNORI YAMAOKA

Tokyo Institute of Technology, Japan
sumiko@net.ss.titech.ac.jp, yamaoka@ss.titech.ac.jp

Keywords: CAC, call blocking, multimedia traffic, VoIP

Multimedia applications such as video and audio have recently come into wide use. Because this heterogeneous traffic occupies most networks, call admission control (CAC) is required to maintain high-quality service. User satisfaction depends on the CAC's success in accommodating application flows. However, conventional CACs do not consider user satisfaction because their main purpose is to improve resource utilization. In this paper, we propose a novel CAC to maximize total accommodated flows based on a new philosophy that heterogeneous traffic is fair in networks. Theoretical analysis was used to derive optimal parameters for various traffic configurations. We also carried out numerical analyses to show the effectiveness of our proposed CAC.

1. Introduction

Multimedia applications such as video and audio have recently become widely used, and the various bandwidth flows associated with them now occupy most networks. Under the resulting heterogeneous traffic conditions, the character of the traffic flows is different from that of best-effort data traffic. Therefore, quality-of-service (QoS) controls are important. One of these QoS controls is call admission control (CAC), which judges whether new flows that arrive can be accommodated in a network. CAC will be increasingly required to maintain QoS as streaming flows occupy more and more network bandwidth.

In a heterogeneous traffic environment, conventional CAC results in some broadband flows, such as video, etc., not being accommodated because of blocking by narrowband flows such as audio, etc. This degrades network utilization and is known as a fraction effect problem [1]. Several studies have proposed the use of reservation controls to solve these problems, which means that unused bandwidths are reserved for subsequent broadband flows [2-6]. These controls have improved resource utilization in networks [7].

It should be noted, however, that while network carriers benefit from improved utilization of network resources, it is important to improve the satisfaction of individual users. The level of individual satisfaction depends on the success of CACs in accommodating flows, but users' satisfaction is not always proportional to users' assigned bandwidth because the values of various applications are not always proportional to the bandwidth.

Therefore, conventional CACs, which aim to increase resource utilization under the condition that users' satisfaction is proportional to the bandwidth available to them, are unlikely to maximize total users' satisfaction.

As stated above, a focus on user satisfaction means that the correlation between each users' satisfaction and the users' own bandwidth must be taken into account. The basic study reported in this paper is that broadband and narrowband flows should be treated equally. Let us assume that each flow receives equal satisfaction as a result of being accommodated in a network, even if each flow has a different bandwidth. Operating under this condition should improve total user satisfaction. This means that as many flows as possible should be accommodated in networks, under the assumption that flows that have different bandwidths are equal from the viewpoint of user satisfaction.

In a conventional research approach to this problem, a CAC that accommodates as many VoIP flows as possible has been proposed [8]. However, the only form of control provided by this CAC is merely to give priority to admission of VoIP flows. Moreover, this research does not consider that users' satisfaction is equal even if there are differences in the bandwidth used for various flows. Therefore, the method does not maximize the total number of flows accommodated in a network.

In this paper, we propose a novel CAC strategy for maximizing total accommodated flows based on a new philosophy that heterogeneous flows should be treated equally in networks. We also propose a CAC algorithm that limits flows to two types, narrowband and broadband.

We first describe the concept on which our proposed CAC is based and then present a theoretical analysis of the model. We also present numerical analyses that show the effectiveness of our CAC.

2. Proposed CAC

A. Concept of new flow admission control

In typical CACs, arriving flows are evaluated whether or not they can be accommodated in a network. When a bandwidth is available to accommodate a flow, the flow is accommodated, but when no bandwidth is available, the flow is rejected. As stated in Section 1, the purpose of the CAC proposed in this paper is to maximize total accommodated flows by treating heterogeneous flows equally. In conventional CAC studies, unused bandwidth is often reserved for broadband flows to increase resource utilization. In contrast, in this paper, bandwidth is reserved for narrowband flows based on the fact that multiple narrowband flows can be accommodated if one broadband flow is rejected.

In fact, maximizing the number of total accommodated flows is equal to minimizing the total call blocking rate of broadband and narrowband flows in a network. Therefore, each arriving flow is evaluated, whether it is accommodated or not, using the procedure proposed below.

B. Proposal procedure of CAC

1) Let b_f be the bandwidth of an arriving new flow.
2) Let b_{now} be the total bandwidth of the flows currently accommodated in a network when a new flow arrives.
3) b_{now} is evaluated to assess whether b_{now} is greater than or equal to the threshold *Th* or not. If b_{now} is less than *Th*, the arriving flow is accommodated.
4) If b_{now} is greater than or equal to *Th* and $b_{now}+b_f$ is less than or equal to *B* which means bandwidth of the link, narrowband flow is accommodated, but a broadband flow is rejected. If $b_{now}+b_f$ is greater than *B*, every arriving flow is rejected.

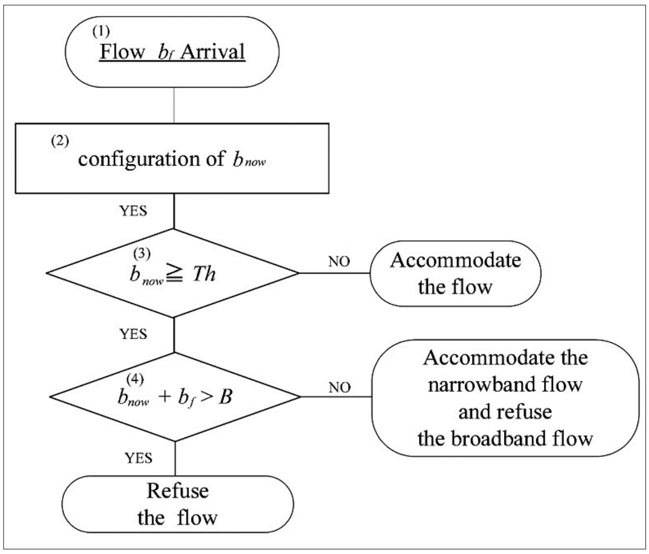


Figure 1. Flow chart

Th is the threshold that decides whether the bandwidth for broadband flows should be reserved for a narrowband flow. This *Th* should be appropriately configured to maximize the total number of accommodated flows. Below, we show the performance of our proposed CAC using a theoretical analysis based on a queuing system.

3. Theoretical analysis

A. Construction of model

In our proposed CAC, each arriving broadband and narrowband flow is evaluated to assess whether or not it can be accommodated in a network. Therefore, our proposed CAC is modeled as a $M_1M_2/M_1M_2/S/S$ loss system, which means that both broadband and narrowband flows arrive independently in a network. We derive the total flow blocking probability (total call blocking rate) from the state transition probability by solving the state transition equations of this queuing model. Thus, an optimal *Th* is calculated that minimizes the total call blocking rate. As a result, we can maximize the total number of flows accommodated in a network by applying the optimal *Th*. Let every bandwidth in this $M_1M_2/M_1M_2/S/S$ system be normalized by the bandwidth of a narrowband flow.

Then, as shown in *Figure 2,* the bandwidth of the link is normalized to s servers; one narrowband flow occupies one server, and one broadband flow occupies *m* servers. When one or more servers are idle, a narrowband flow can be accommodated in the network. When *m* or more servers are idle, a broadband flow can be accommodated in the network. In other cases, neither flow can be accommodated in the network.

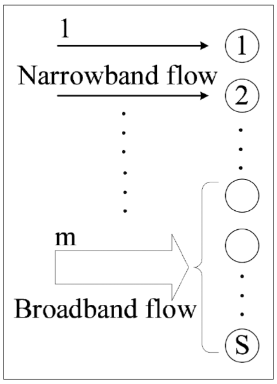


Figure 2. Model of heterogeneous traffic

Table 1 shows the definitions of each parameter. The bandwidth of the link is *kB* [Mbps]. The arrival rates of both the narrowband and broadband flows comply with Poisson distributions independently.

ρ_1 and ρ_2 denote the arrival rates, respectively. The holding times of the two types of flows comply with exponential distributions, with rates of b_1 for narrowband flows and b_2 for broadband flows. The state probability when the number of narrowband flows is n_1 and the number of broadband flows is n_2 in the networks is $P(n_1, n_2)$. In addition, the narrowband traffic intensity ρ_1 and broadband traffic intensity ρ_2 for each flow is $\rho_1 = \frac{\lambda_1}{\mu_1}$ and $\rho_2 = \frac{\lambda_2}{\mu_2}$, respectively.

Table 1. Parameters

	Parameter	Definition
Narrow band	λ_1[flows/s]	Flow Arrival Rate
	μ_1[flows/s]	Flow Processing Rate
	n_1[number]	Flow Number in Networks
	kb_1[Mbps]	Bit-rate of Narrowband flow
Broad band	λ_2[flows/s]	Flow Arrival Rate
	μ_2[flows/s]	Flow Processing Rate
	n_2[number]	Flow Number in Networks
	N_2[number]	Maximum Number of Broadband flows
	kb_2[Mbps]	Bit-rate of Broadband flow

B. Derivation of state transition equations

The state probability and call blocking rate when there is no bandwidth reservation control in a heterogeneous traffic network were generally derived in a previous study [9]. The narrowband and broadband call blocking rates are shown as follows, respectively:

$$r_1 = \frac{\displaystyle\sum_{n_2=0}^{N_2} \frac{\rho_1^{B-b_2n_2}}{(B-b_2n_2)!}\frac{\rho_2^{n_2}}{n_2!}}{\displaystyle\sum_{n_2=0}^{N_2}\sum_{n_1=0}^{B-b_2n_2}\frac{\rho_1^{n_1}\rho_2^{n_2}}{n_1!n_2!}} \tag{1}$$

$$r_2 = \frac{\displaystyle\sum_{n_2=0}^{N_2-1}\sum_{n_1=B-b_2(n_2+1)+1}^{B-b_2n_2}\frac{\rho_1^{n_1}}{n_1!}\frac{\rho_2^{n_2}}{n_2!}+\frac{\rho_2^{N_2}}{N_2!}}{\displaystyle\sum_{n_2=0}^{N_2}\sum_{n_1=0}^{B-b_2n_2}\frac{\rho_1^{n_1}\rho_2^{n_2}}{n_1!n_2!}} \tag{2}$$

General state transition equations in heterogeneous traffic networks are shown in [7] for situations when the bandwidth for both narrowband and broadband flows is reserved.

In conventional research on improving resource utilization, only the bandwidth for narrowband flows has been reserved. There has been no reservation of bandwidth for broadband flows. General state transition equations in heterogeneous traffic networks are also shown [3] for situations when only the bandwidth for narrowband flows is reserved. Because it is difficult to solve both these general equations, these methods have only been evaluated using numerical analysis or an approximate solution in conventional studies [10].

In contrast, we investigated performances when the bandwidth for broadband flows was reserved for narrowband flows. The purpose of this approach, which has not been used in previous research, is to maximize the total number of accommodated flows.

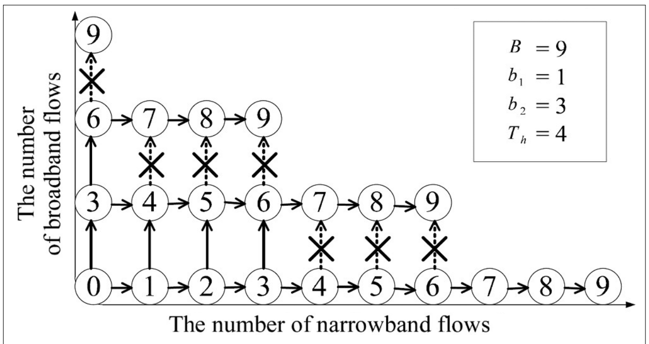


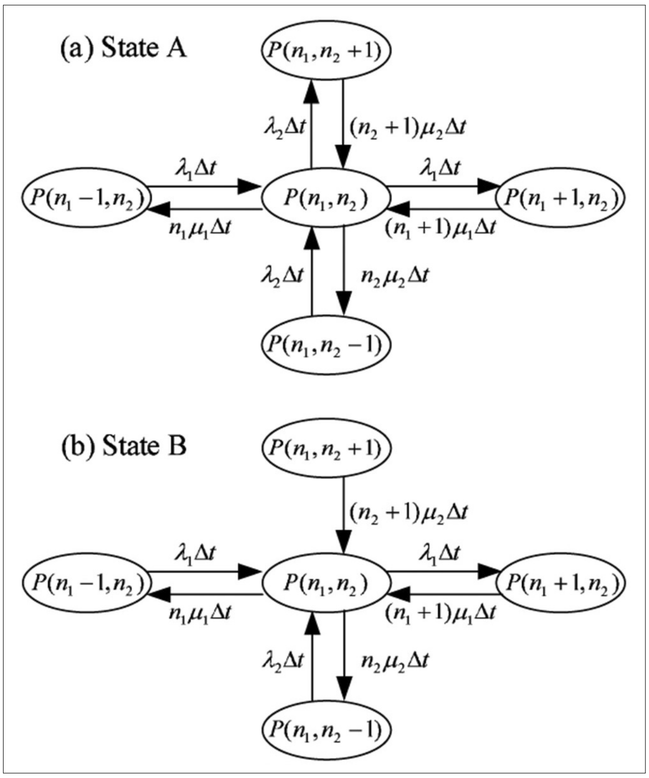
Figure 3.
Transition diagram for broadband and narrowband flows (reservation control)

Figure 3 shows a state transition diagram. In this figure, the vertical axis gives the number of broadband flows, and the horizontal axis gives the number of narrowband flows. In figure, the parameters are $B=9$, $b_1=1$, and $b_2=3$. The number in each circle indicates the total bandwidth of narrowband and broadband flows currently accommodated in a network at the state. For simplicity, arrows that show a flow leaving (in a downward or left direction) are omitted. From here on, bandwidths kB [Mbps], kb_1 [Mbps] and kb_2 [Mbps] are normalized to kb_1, that is, $kb_1=1$, $kB=B$, and $kb_2=b_2$, respectively. In addition, let B be b_2N_2 for simplicity in this study.

In our proposed CAC, as shown in Section 2.B, if the total accommodated bandwidth b_{now} is greater than or equal to the threshold Th when a new flow arrives, the bandwidth for broadband flows is reserved for narrowband flows, and an arriving broadband flow is rejected. N_2' is the maximum number of broadband flows that can be accommodated with reservation control. The range of the threshold Th is $0 \le Th \le B-b_2$.

Figure 3 shows an example when $Th=4$. When the total accommodated bandwidth b_{now} is greater than or equal to 4 ($b_{now} \ge 4$), reservation control begins, and arriving broadband flows are rejected. When our proposed system of reservation control is not applied, the number of maximum accommodated broadband flows N_2 is 3 ($N_2=3$). When our reservation system is applied with a threshold Th, the number of maximum accommodated broadband flows decreases to 2 ($N_2=2$). Instead of using this reservation, more narrowband flows could be accommodated.

To set up static state transition equations from the transition diagram which shows our proposed CAC system, let the total accommodated flows b_{now} be considered as $b_1n_1+b_2n_2$ ($b_{now}=b_1n_1+b_2n_2$). Here, state transition events in infinitesimal time (Δt) are limited to neighboring states because this state transition diagram assumes a Poisson distribution for arrival and an exponential distribution for processing. Therefore, each state is grouped into six state groups from state A to state F:



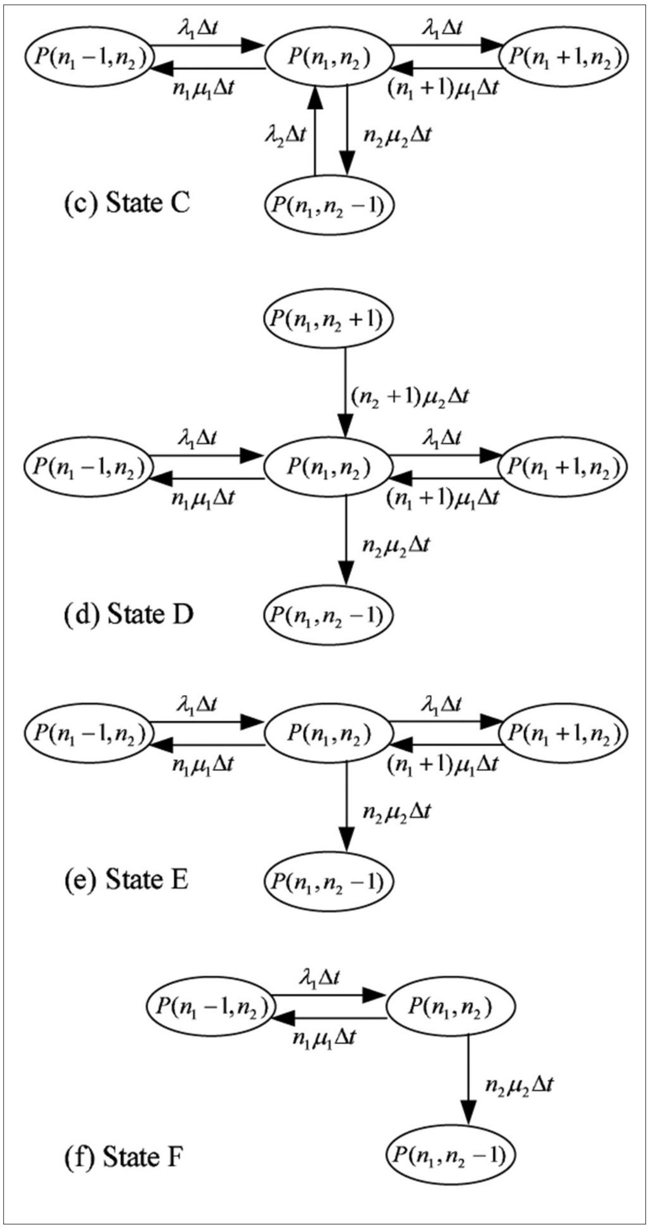


Figure 4. State transition diagram

In addition, because $n_1 \geq 0$ and $n_2 \geq 0$ are required, when $n_1 < 0$ or $n_2 < 0$, let $P(n_1,n_2) = 0$.

[1] When $n_2 < N_2'$

- When $0 \leq b_{now} < Th$ (StateA)

$$(\lambda_1 + \lambda_2 + n_1\mu_1 + n_2\mu_2)P(n_1, n_2) = \lambda_1 P(n_1 - 1, n_2) + \lambda_2 P(n_1, n_2 - 1) + \mu_1(n_1 + 1)P(n_1 + 1, n_2) + \mu_2(n_2 + 1)P(n_1, n_2 + 1)$$

- When$\{ (Th \leq b_{now} < Th + b_2) \wedge (0 \leq Th \leq B - 2b_2) \} \vee \{(Th \leq b_{now} < B - b_2) \wedge (B - 2b_2 < Th \leq B - b_2)\}$ (StateB)

$$(\lambda_1 + n_1\mu_1 + n_2\mu_2)P(n_1, n_2) = \lambda_1 P(n_1 - 1, n_2) + \lambda_2 P(n_1, n_2 - 1) + \mu_1(n_1 + 1)P(n_1 + 1, n_2) + \mu_2(n_2 + 1)P(n_1, n_2 + 1)$$

- When$\{(B - b_2 < b_{now} < Th + b_2) \wedge (B - 2b_2 + 1 < Th < B - b_2) \vee \{(B - b_2 < b_{now} < B) \wedge (Th = B - b_2)\}$ (StateC)

$$(\lambda_1 + n_1\mu_1 + n_2\mu_2)P(n_1, n_2) = \lambda_1 P(n_1 - 1, n_2) + \lambda_2 P(n_1, n_2 - 1) + \mu_1(n_1 + 1)P(n_1 + 1, n_2)$$

- When $(Th + b_2 \leq b_{now} < B - b_2) \wedge (0 \leq Th \leq B - 2b_2)$ (StateD)

$$(\lambda_1 + n_1\mu_1 + n_2\mu_2)P(n_1, n_2) = \lambda_1 P(n_1 - 1, n_2) + \mu_1(n_1 + 1)P(n_1 + 1, n_2) + \mu_2(n_2 + 1)P(n_1, n_2 + 1)$$

- When $\{(B - b_2 < b_{now} < B) \wedge (0 \leq Th \leq B - 2b_2 + 1)\} \vee \{ (Th + b_2 \leq b_{now} < B) \wedge (B - 2b_2 + 1 < Th < B - b_2) \}$ (StateE)

$$(\lambda_1 + n_1\mu_1 + n_2\mu_2)P(n_1, n_2) = \lambda_1 P(n_1 - 1, n_2) + \mu_1(n_1 + 1)P(n_1 + 1, n_2)$$

- When $b_{now} = B$ (StateF)

$$(n_1\mu_1 + n_2\mu_2)P(n_1, n_2) = \lambda_1 P(n_1 - 1, n_2)$$

[2] When $n_2 = N_2'$

- When $(N_2' b_2 \leq b_{now} < Th + b_2)$ (StateC)

$$(\lambda_1 + n_1\mu_1 + n_2\mu_2)P(n_1, n_2) = \lambda_1 P(n_1 - 1, n_2) + \lambda_2 P(n_1, n_2 - 1) + \mu_1(n_1 + 1)P(n_1 + 1, n_2)$$

- When $\{(Th + b_2 \leq b_{now} < B) \wedge (Th \neq B - b_2)\}$ (StateE)

$$(\lambda_1 + n_1\mu_1 + n_2\mu_2)P(n_1, n_2) = \lambda_1 P(n_1 - 1, n_2) + \mu_1(n_1 + 1)P(n_1 + 1, n_2)$$

- When $b_{now} = B$ (StateF)

$$(n_1\mu_1 + n_2\mu_2)P(n_1, n_2) = \lambda_1 P(n_1 - 1, n_2)$$

The sum of all state probabilities is equal to 1. This summation is given by the following equation:

$$\sum_{(n_1,n_2) \in \{A.B.C.D.E.F\}} P(n_1, n_2) = 1 \tag{3}$$

In this equation, {A,B,C,D,E,F} mean a set of each state, A, B, C, D, E, and F, respectively.

Above these equations, which relate to all states, are simultaneous linear equations with variable $P(n_1, n_2)$.

By using $P(n_1, n_2)$, which is derived from these state transition equations, the narrowband call-blocking rate r_1 and broadband call-blocking rate r_2 can be given by the following equation, respectively:

$$r_1 = \sum_{n_2=0}^{N_2'} \mathrm{P}(B - b_2 n_2, n_2) \quad (4)$$

$$r_2 = \sum_{n_2=0}^{N_2'-1} \sum_{n_1=T_h-b_2n_2}^{B-b_2n_2} \mathrm{P}(n_1, n_2) + \sum_{n_1=0}^{B-b_2n_2} \mathrm{P}(n_1, N_2') \quad (5)$$

Both r_1 and r_2 mean the summation of state probabilities which cannot transit to neighboring states when a new flow arrives.

C. Derivation of Total Call Blocking Rate

As shown in Section 3.B, when both the narrowband and broadband call blocking rates are calculated, the total call blocking rate, which reflects user satisfaction, is given by the following equation:

$$r_{total} = \frac{\rho_1 r_1 + \alpha \rho_2 r_2}{\rho_1 + \rho_2} \quad (6)$$

In this equation, α is the weight of user satisfaction when a broadband flow is accommodated in a network. For example, when user satisfaction is proportional to the user's own bandwidth, user satisfaction with the broadband flow is equal to its own bandwidth for broadband flow (α=b_2) [7]. In this paper, the weight of user satisfaction is set to 1 (α=1) as a basic condition of the study. Under this condition, Eq. 6 shows the total call-blocking rate when every flow is equal.

Thus, to minimize the total blocking rate, we need to find the optimal threshold Th_{opt} under the constraint of state transition equations. However, because it is difficult to solve these state equations [3], P(n_1, n_2) and r_{total} are calculated using numerical calculation for numerical analysis. In the following section, we describe the performance relations between the optimal threshold Th_{opt} and the optimal total call-blocking rate r_{opt}.

4. Evaluation of the characteristics of our proposed CAC using numerical analysis

To examine the effectiveness of our proposed CAC using numerical analysis, optimal total call-blocking rates r_{opt} were calculated by changing the traffic intensity ρ_1 and ρ_2 and by changing the threshold *Th*. This optimal call-blocking rate r_{opt} is given by the optimal threshold Th_{opt}, which minimizes the total call-blocking rate for each traffic condition.

Figures 5, 6, 7, and 8 show the optimal total call-blocking rates r_{opt} and the total call-blocking rates without the proposed reservation control by changing the traffic intensity ρ_1 and ρ_2 and by changing the broadband bandwidth b_2.

Fig. 5. Total call-blocking rate ($b_2 = 3$)

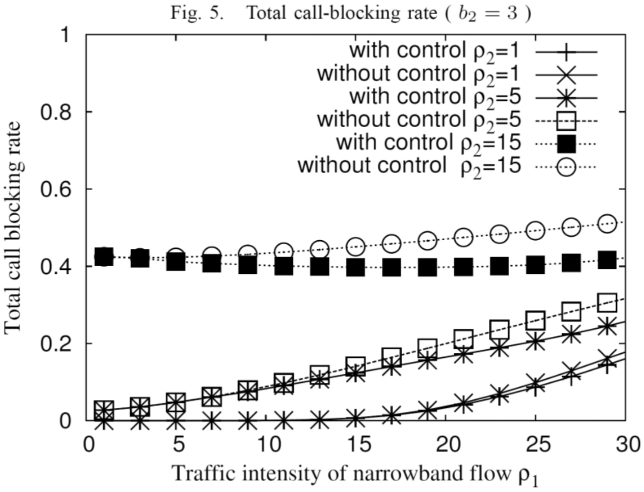


Fig. 7. Total call-blocking rate ($b_2 = 10$)

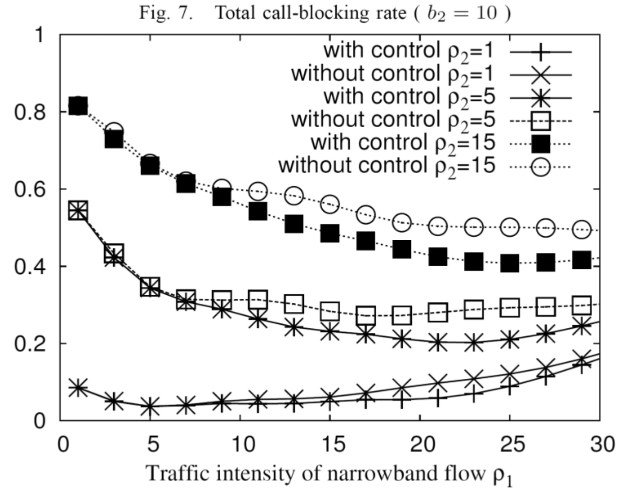


Fig. 6. Total call-blocking rate ($b_2 = 6$)

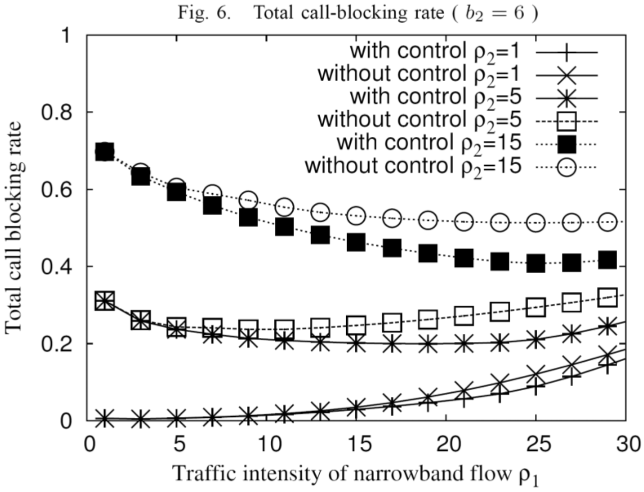
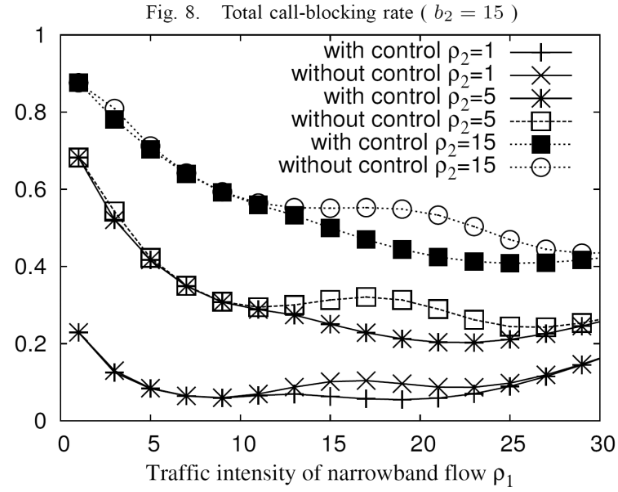


Fig. 8. Total call-blocking rate ($b_2 = 15$)



These total call-blocking rates without the proposed reservation control are calculated by substituting equations (1) and (2) into equation (6). The horizontal axis is the traffic intensity of narrowband flows ρ_1. Each parameter is shown as in *Table 2*.

Bandwidth of Link (B)	30
Flow Arrival Rate of Narrowband Flow(λ_1)	0.02[flows/s]
Flow Arrival Rate of Broadband Flow(λ_2)	0.02[flows/s]
Sending Rate of Narrowband Flow(b_1)	1
Sending Rate of Broadband Flow(b_2)	3, 6, 10, 15

Table 2.
Parameters for mathematical analysis

We can describe the characteristic relations between traffic intensity and optimal total call-blocking rate r_{opt} as follows.

1) When ρ_2 is fixed, the larger ρ_1 is, and the larger the decrease is by the proposed reservation control.
2) The larger ρ_2 is, the larger the decrease is by the proposed reservation control.
3) The larger b_2 is, the larger the wave oscillation is of r_{total} without the proposed reservation control.
This is known as a fraction effect problem [9].
4) The larger ρ_1 and b_2 are, the smaller r_{opt} is.

Regarding 1), in Figure 5 when $\rho_2 = 15$ is fixed, while the difference between r_{opt} with our reservation control and r_{total} without our reservation control is about 1% when $\rho_1 = 5$, the difference is about 10% when $\rho_1 = 30$.

Regarding 2), in Figure 5 when $\rho_1 = 30$, while the difference between r_{opt} with our control and r_{total} without our control is about 1% when $\rho_2 = 1$, the difference is about 10% when $\rho_2 = 15$. Therefore, the larger ρ_1 and ρ_2 are, the larger the decrease in the total call-blocking rate is.

Regarding 3), in Figure 8, if the total call-blocking rate without our reservation control waves by the fraction effect problem, our reservation control can prevent the total call-blocking rate from increasing. For example, in Figure 8, when $\rho_1 = 21$ and $\rho_2 = 15$, the optimal total call-blocking rate r_{opt} decreases about 15% compared with the total call-blocking rate without our control.

Regarding 4), in our proposed reservation control, r_{total} includes the weight of each traffic intensity. Under this condition, the larger ρ_1 is, the more narrowband flows can be accommodated in the networks. Therefore, r_{total} decreases. In the following figure, we show characteristics of the optimal threshold Th_{opt}.

Figures 9, 10, 11, and 12 show characteristics of the optimal threshold Th_{opt} by changing the traffic intensity ρ_1 and ρ_2 and by changing the broadband bandwidth b_2. The optimal threshold Th_{opt} is the threshold that mi-

Fig. 9. Optimal threshold ($b_2 = 3$)

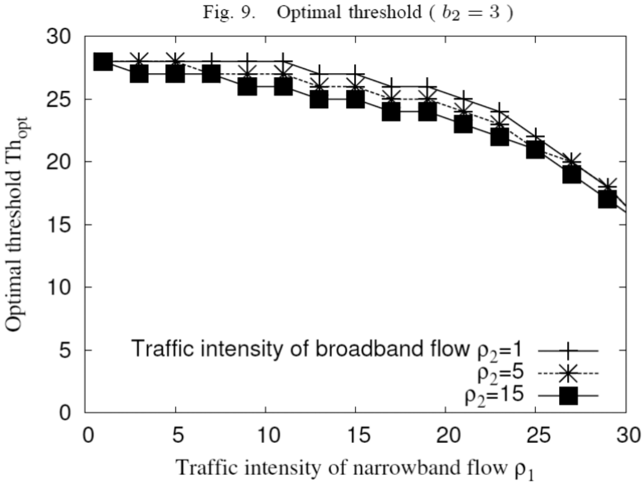


Fig. 11. Optimal threshold ($b_2 = 10$)

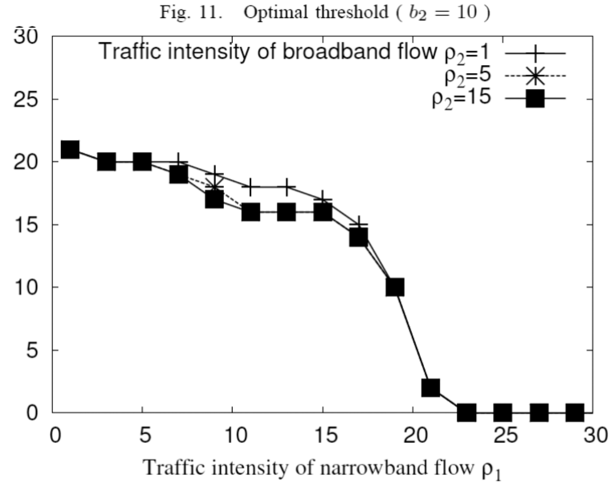


Fig. 10. Optimal threshold ($b_2 = 6$)

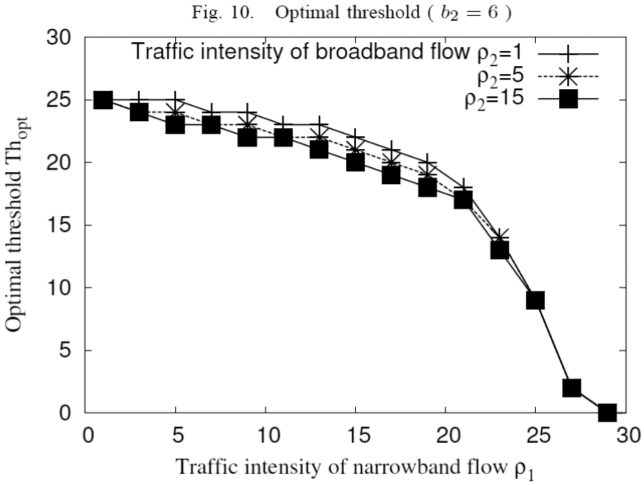
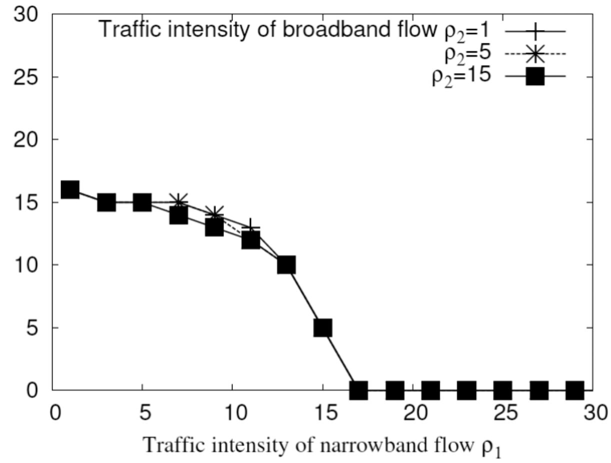


Fig. 12. Optimal threshold ($b_2 = 15$)



nimizes the total call-blocking rate for each traffic condition. The horizontal axis is the traffic intensity of narrowband flows ρ_1. When $Th_{opt} = B - b_2 + 1$, this r_{opt} is the total call-blocking rate when the proposed reservation control system is not applied. In addition, when $Th_{opt} = 0$, no broadband flows are accepted.

These figures show the results when the traffic intensity of broadband flows ρ_2 is fixed; the larger ρ_1 is, the smaller Th_{opt} is. This result indicates that the more narrowband flows there are, the more broadband flows should be rejected with a smaller Th_{opt}. This enables more narrowband flows to be accommodated, thus reducing the total call-blocking rate. Moreover, these characteristics do not depend on ρ_2. In Figures 9-12, the larger b_2 is, the smaller ρ_1 making $Th_{opt} = 0$. When $Th_{opt} = 0$, arriving new broadband flows cannot be accommodated. This result indicates that the larger broadband bandwidth b_2 is, the more narrowband flows can be accommodated by our reservation control.

Figures 13 and 14 show the total call-blocking rates when Th is changed. Figure 13 shows the results when ρ_1 is $\rho_1 = 30$ in Figure 9. In this figure, when ρ_1 is as large as $\rho_1 = 30$, the change in r_{total} becomes flatter. For example, in Figure 13 when $\rho_1 = 30$ and $\rho_2 = 15$, the optimal threshold is $Th_{opt} = 15$. However, the total call-blocking rates r_{total} are kept almost unchanged when the optimal threshold Th_{opt} is set in the range from $Th = 0$ to $Th = 21$.

In Figure 14, the difference in the decrease in the total call-blocking rates r_{total} varies according to traffic conditions when Th is nearly Th_{opt}. For example, when $\rho_1 = 10$, the optimal threshold is $Th_{opt} = 22$. However, the total call-blocking rate r_{total} increases to about 3% when the threshold Th is set to nearly Th_{opt}. Meanwhile, when $\rho_1 = 30$, the optimal threshold is $Th_{opt} = 0$. Under this condition, the total call-blocking rates r_{total} are kept almost unchanged. Therefore, under this condition, a near-optimal total call-blocking rate r_{total} is given by using reservation control when the threshold Th is nearly Th_{opt}. In other words, a near optimal total-call blocking rate r_{total} is given when Th is set to nearly $Th = 0$. In contrast, when $\rho_1 = 1$, the total call-blocking rate decreases, or cascades, suddenly every b_2, which is the bandwidth for broadband flow; the smaller the Th is, the larger the total call-blocking rate r_{total} is. Therefore, under this condition, the proposed reservation control is not expected to be effective.

Here, we focus attention on the cascade in Figure 14 when $\rho_1 = 1$. *Figure 15* shows the narrowband call-blocking rates, broadband call-blocking rates, and total call-blocking rates with this configuration. As shown in the figure, this cascade is caused by changing the difficulty in accommodating broadband flows every b_2.

Overall, we can describe the characteristic relations between traffic intensity and threshold as follows.

- When b_2 is fixed, the larger ρ_1 and ρ_2 are, the smaller Th_{opt} is, which minimizes the total call-blocking rate.
- The larger b_2 is, the larger the decrease by our proposed reservation control.
- The larger ρ_1 is, the flatter the differences in r_{total} become. Therefore, under this condition, a near optimal total call-blocking rate r_{total} is given by using reservation control when the threshold Th is nearly Th_{opt}.
- When ρ_1 is small, r_{total} steps down with every b_2. Under this condition, the proposed reservation control is not expected to be effective.

Figures 13-15.

Fig. 13. total call-blocking rate ($b_2 = 3$, $\rho_1 = 30$)

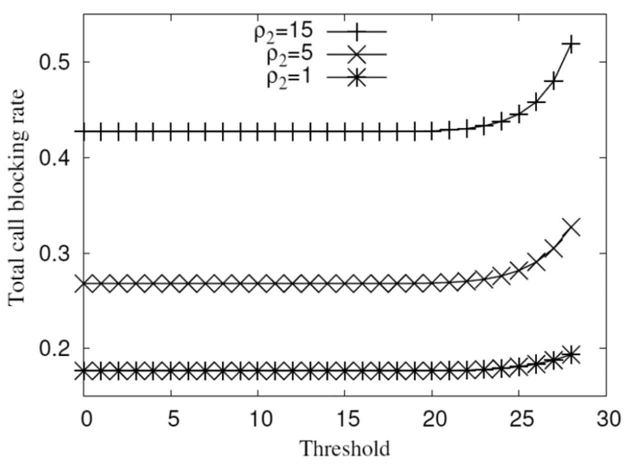


Fig. 14. total call-blocking rate ($b_2 = 6$, $\rho_2 = 5$)

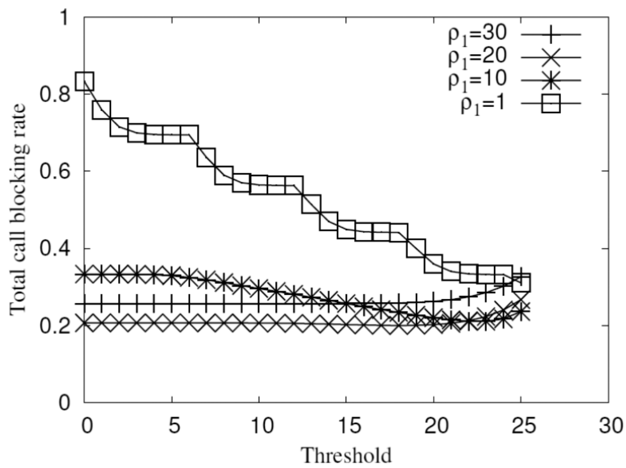
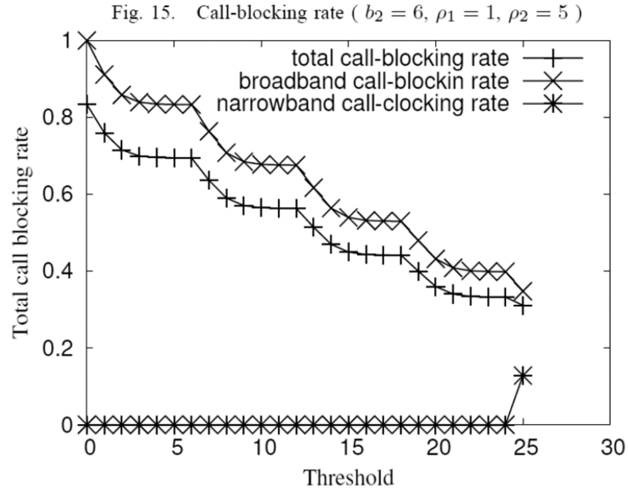


Fig. 15. Call-blocking rate ($b_2 = 6$, $\rho_1 = 1$, $\rho_2 = 5$)



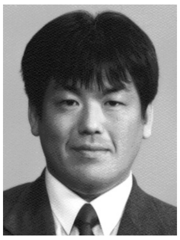
Therefore, when Th_{opt} is set appropriately according to traffic intensities, r_{total} is minimized. In addition, as shown in Figure 13, r_{total} approaches the optimal total call-blocking rate under certain traffic conditions, such as when $\rho_1 = 30$, even if the threshold Th is not equal to Th_{opt}.

5. Conclusions

In this paper, we proposed a novel CAC strategy for maximizing total accommodated flows based on the new philosophy that heterogeneous flows should be treated equally in networks. Our proposed CAC is modeled on a $M_1M_2/M_1M_2/S/S$ system, and theoretical numerical analyses show its effectiveness. In future work, we will evaluate the performance of our proposed CAC under various traffic configurations because the number of accommodated flows becomes close to the maximum number of total accommodated flows under some traffic configurations, even if the optimal threshold Th_{opt} is not applied. We will also derive the optimal threshold Th_{opt} for minimizing the total call-blocking rate when the parameters B, b_1, and b_2 change, and examine how to establish the most practical Th_{opt}.

Authors

SUMIKO MIYATA received the B.E. degrees from the Shibaura Institute of Technology in 2007, and the M.E. degrees from the Tokyo Institute of Technology in 2009. She is currently undertaking a Ph.d. candidate of the Department of Communications and Technology in the Tokyo Institute of Technology. Her research interests include QoS management techniques, multimedia communications.



KATSUNORI YAMAOKA received the B.E., M.E., and Ph.D. degrees from the Tokyo Institute of Technology in 1991, 1993 and 2000, respectively. He left Ph.D program in 1994 and joined the Tokyo Institute of Technology as an assistant professor at that time. In 2000, he joined the National Institute of Multimedia Education (NIME) in Japan as an associate professor. Since 2001, he has been an associate professor at the Tokyo Institute of Technology. He has also been a visiting associate professor of the National Institute of Informatics (NII) in Japan since 2004. His research interests are network QoS control for multimedia communicatons.

References

[1] L. Gimpelson, "Analysis of Mixtures of Wide- and Narrow-Band Traffic," IEEE Transactions on Communication Technology, Vol. 13, No. 3, September 1965.

[2] T. Oda, Y. Watanabe, "Optimal trunk reservation for a group with multislot traffic streams," IEEE Transactions on Communications, Vol. 38, No. 7, pp.1078–1084, July 1990.

[3] T. Yamaguchi, M. Akiyama, "An Integrated-Hybird-Traffic Switching System (Using Adaptive Trunk Reservation)," IECE Transactions Japan (in Japanese), Vol. 53, No. 9, pp.496–498, September 1970 .

[4] K. Lindberger, "Blocking for multi-slot heterogeneous traffic streams offered to a trunk group with reservation," Presented at Proc. ITC-12, Torino, Italy, 1988.

[5] Kai-Wei Ke, Chin-Tau Lea, "State reduction and call admission in multirate circuit-switched networks," IEEE International Conference on ICC 96, Vol. 2, pp.777–781, June 1996.

[6] K. Kawashima, "Trunk reservation models in telecommunications systems," The international seminar on Teletraffic analysis and computer performance evaluation, pp.29–36, 1986.

[7] K. Miyake, "Optimal Trunk Reservation Control for Multi-Slot Connection Circuit" IEICE Transactions on Communications, Vol. J71-B, No. 12, pp.1419–1424, November 1988.

[8] Y. Nasuno, S. Konno, Y. Iwaya, T. Kinoshita, K. Arai, "A proposal of VoIP quality assurance method by admission control," IEICE Transactions on Communications, Vol. 90, pp.12–22, October 2006.

[9] M. Akiyama, "Modern communication switch engineering," Denkisyoin (1972), pp.205–211, (in Japanese)

[10] M.I. Reiman, "Asymptotically Optimal Trunk Reservation for Large Trunk Groups," IEEE International Conference on CDC, Vol. 3, pp.2536–2541, December 1989.

Flow optimization in IP networks with fast proactive recovery

MATEUSZ DZIDA, MICHAL ZAGOZDZON, MATEUSZ ZOTKIEWICZ, MICHAL PIÓRO

Institute of Telecommunications, Warsaw University of Technology, Poland
{mdzida, mzagozdz, mzotkiew, mpp}@tele.pw.edu.pl

Keywords: protection, rerouting, fast reroute, LFA, MHRP

The post-failure convergence of shortest path routing (SPR) protocols used in IP networks can be too slow to meet the restrictive requirements (i.e. maximum allowable delay, jitter, etc.) of the multimedia services and therefore new protection/restoration mechanisms combined with IP routing are of interest. The article addresses the optimization of three potential rerouting mechanisms based on the IP fast reroute mechanism proposed by Shand and Bryant [1].

The first mechanism takes advantage of multiple equal-cost (shortest) paths (ECMP) where two or more ECMP paths outgoing from one router can be used to protect one another in the IP fast reroute mechanism. However, due to a limited number of the ECMP paths, ECMP protection cannot be used as a stand-alone protection mechanism assuring protection against link failures. Therefore, two other mechanisms, called loop-free alternate (LFA) and multi-hop repair path (MHRP) are considered. LFA protection consists of determining an alternative next-hop address used in the case of a link failure. MHRP is a generalization of LFA using multi-hop tunnels to redirect packets from the failing link to a router that is able to send them to the destination based on a shortest path based forwarding. For each mechanism we formulate an optimization problem as a mixed integer programme (MIP).

We also consider a combined approach where protection is assured through ECMP paths, LFA next-hop addresses, or MHRP paths. Thanks to the variety of such a protection mechanisms, the IP fast reroute technique is able to provide protection for any single link failure. The associated optimization problem (consisting in a simultaneous optimization of a weight system, LFA alternative next-hop addresses and MHRP paths) is difficult and is thus approached with a heuristic method. In our numerical experiments we evaluate the effectiveness of this heuristic method.

1. Introduction

Shortest-path routing (SPR) is a widely deployed intra-domain routing mechanism in today's Internet. Interior gateway protocols (IGP) like Open Shortest Path First (OSPF) or Intermediate System-to-Intermediate System (IS-IS) determine routing paths as shortest paths with respect to predefined link weight systems. Hence, selection of proper link weights is a traffic engineering task in the networks using SPR.

An SPR weight setting problem is NP-hard (cf. [2]) and is typically approached with heuristics. Although the problem can be resolved with exact methods (e.g., formulated as a MIP problem and solved with a branch-and-cut algorithm) such an approach allows for resolving the problem only for small instances with up to, say, 10 nodes. An SPR routing mechanism accompanied with periodic broadcast of the link weights can be used to restore traffic flows on paths affected by link failures. Once a failure of a link is detected in its adjacent router, the router broadcasts unavailability of the link, assigning a large weight to it, e.g., the maximum possible value, which in the case of OSPF is equal to 2^{16}-1. Each time a router obtains a weight update message, it recalculates the routing paths and in this way all failed links are eliminated from the routing paths. However, link weight updates and recalculation of the shortest paths introduces delay in traffic restoration.

According to [1] typical delays in OSPF post-failure convergence are as follows:

1) **Detection time**

Delay of failure detection is of the order of a few milliseconds (in the case of the detection on the level of a physical layer) to tens of seconds (using hello packets of routing protocols).

2) **Failure local reaction**

Delay related to generation and broadcasting of link state update messages that depends on some internal hold-down delays.

3) **Broadcasting**

Delay resulting from the link state update messages propagation through the network. Typically of the order of tens of milliseconds per each hop.

4) **Shortest paths update**

Delay of the order of a few milliseconds due to calculations of the shortest paths with respect to the updated weight system.

5) **Forwarding table update**

Delay introduced by an update of the forwarding table in forwarding hardware. For some routers it is up to several hundred milliseconds.

It follows that the total path-update delay is not negligible. Although recent developments in tuning the timeouts of the IGP protocols allowed to reduce the delay of the SPR re-convergence to second(s) [3], new multimedia services based on streaming transmission, e.g., VoIP or VoD, are typically delay (jitter) sensitive and still may not accept this relatively slow SPR re-convergence after a failure. Therefore, many current research activities focus on developing other restoration mechanisms applicable in SPR networks and satisfying the requirements of the streaming services (see [4,5]). A promising idea of an alternative recovery mechanism is a proactive IP Fast Rerouting (IFR) mechanism, based on the concept of MPLS Fast Rerouting. The IFR recovery mechanism is described in IETF draft [1].

Typically, network operators set SPR link weights proportional to the inverse of link capacities or use uniform link weights equal to 1 (in the latter case routing paths are the shortest paths with respect to the number of hops). In fact, these two weight systems do not take into account volume of the traffic demands (actual or approximated). Therefore, the resulting routing patterns may lead to otherwise avoidable link overloads or to unbalanced network utilization. To overcome these drawbacks, the SPR weight setting problem can instead be approached with network optimization techniques. An appropriate routing optimization problem is formulated (as an instance of multi-commodity flow optimization taking into account network topology and traffic demands) and resolved with proper optimization methods. Moreover, while formulating and solving the problem, network resilience to failures can be explicitly taken into account.

Assuming a set of significant failures that can occur in the network, we can formulate an optimization problem which allows for calculating the SPR routing paths in the normal (failure-free) state as well as in any considered failure state. Resolving such a formulation, we may try to find a weight system that does not cause link overloads in any failure state (of course if such a solution exists) and optimizes certain traffic engineering criteria, e.g., maximal residual capacity. Hence, approaching the problem with network optimization techniques is a suitable approach to tackle the problem of weight selection.

In the article we discuss three proactive IFR mechanisms:

- equal-cost multiple paths,
- loop-free alternates,
- multi-hop repair paths,

that are taken into account in the recovery scenario assuming single link failures.

Let *e* be the failing link and w^o be the current weight system. If there exists at least one surviving shortest path from the starting router of link *e* to a particular destination, affected packets are redirected to the surviving SPR paths (with appropriate ECMP split in the case of multiple surviving paths). Otherwise (all shortest paths from the starting node of *e* to the destination are using link *e* in the normal state), the router checks if the failing link *e* is assigned an alternate next-hop address (LFA) for the required destination. If this is the case, the packets are redirected to the router associated with this address and then are forwarded on the shortest paths from this router to the considered destination. As both recovery mechanisms do not guarantee protection in all failure states (all single link failures), a third mechanism seems unavoidable. This is the case when for some router neither ECMP split is used nor there exits an LFA next-hop address such that when used does not result in flow loops. One such mechanism is to use multi-hop repair paths (MHRPs). A repair path in this case is a pre-established MPLS tunnel from the starting router of link *e* to a remote router. Similarly, as in the case of LFA, end-router of the repair path forwards redirected packets to the destination.

The considered IFR mechanisms allow for the compensation of the negative impact of relatively slow convergence of the routing protocol after a failure is detected. In fact, IFR framework, if deployed in the network, eliminates the need for the post-failure routing protocol re-convergence. The detecting routers simply locally reroute the affected flows using either the surviving ECMP paths, alternative LFA next-hop addresses, or multi-hop repair paths (MHRPs). This can be done very quickly as no exchange of information with other routers is needed. Moreover, it guarantees that the rerouted flows will reach the destination and will not be lost, e.g., due to transient flow loops that are most likely to occur while the re-convergence of the conventional IP routing is concerned.

An important aspect related to the considered IFR restoration mechanisms is the fact that they should be taken into account while designing the distribution of network flows, e.g., by the appropriate selection of link weights, LFA addresses, and routes for MHRP paths. The goal might for example be to optimize the utilization of available resources. The best way to approach this problem is to use network optimization techniques, which is the main goal of this paper. In this regard, the main contributions of the article are as follows. For each mechanism we formulate an optimization problem as a mixed integer programming (MIP) and consider a combined approach where protection is assured either through ECMP paths, LFA next-hop addresses, or MHRP paths.

Thanks to the variety of the protection mechanisms used, IP fast reroute technique is able to provide protection for any single link failure. The associated optimization problem consisting of a simultaneous optimization of a weight system, LFA alternative next-hop addresses and MHRP paths, is difficult and is thus approached with a heuristic method. In our numerical experiments we evaluate effectiveness of this heuristic method.

The article is organized as follows. In Sections 2, 3, and 4 we discuss in more detail three considered recovery mechanisms and present corresponding MIP optimization problems. Due to the difficulty of the combined approach involving all these three mechanisms, Section 5 is devoted to the presentation of an efficient heuristic method that deals with this kind of problem. It is then evaluated by means of the numerical study presented in Section 6.

2. Equal-cost multiple paths

In this section we present an idea of using equal-cost multiple paths (ECMP) as recovery paths in the case of a link failure (cf. [1]). Typically, when a network operator chooses a link weight system to be used in a network, it may appear that there are multiple shortest paths connecting some routers.

Basic implementations of OSPF or IS-IS routing protocols do not clearly state which shortest path should be used in such a case. Therefore, the deployed implementations of the SPR routing arbitrarily choose routing paths among the shortest paths, e.g., the protocols can choose a next-hop address of the router with the smallest id among the neighboring routers associated with the shortest paths. Due to ambiguity in routing path selection, many works concentrate on selection of the weight systems generating single unique shortest paths. Such systems allow for exactly determining the routing paths without any ambiguity.

However, in the case of a failure it may happen that instead of one shortest path multiple shortest paths appear. Thus, in general, this approach does not overcome the ambiguity. Another approach to deal with multiple shortest paths is to use all available shortest paths to carry traffic. With the help of hashing functions, the flow is equally split in the router among each outgoing shortest-path link (this is the main principle of the equal-cost multiple path (ECMP) flow split). Ideally, the hashing function should allow for equal traffic splitting such that each outgoing shortest-path link carries flow of the same (aggregated) volume.

Notice that multiple shortest paths that do not share a common outgoing link in a particular router may be used as IFR backup paths, protecting one each other. The router detecting a failure immediately erases a next-hop address related to the unavailable neighbor, and distributes evenly the flow destined to a particular destination to all remaining next-hop addresses, related to the non-affected shortest paths. To protect as many links as possible the split must be done at each intermediate router. Unfortunately, a weight system satisfying this condition does not exist. This is because for a particular destination there is always at least one router that, due to the properties of the shortest paths, must not apply ECMP.

However, the protection based on ECMP paths is among the fastest restoration IP mechanisms (we have valid alternative routes in hand). Hence, the optimization problem that arises is how to choose a weight system that generates ECMP shortest-path flows such that the link capacities are not exceeded and a number of routers with multiple outgoing paths is maximal. Below we state the related flow optimization problem. Its solution is an ECMP flow pattern (and the corresponding weight system) such that the link loads do not exceed the link capacities and demands are realized on the shortest paths. Our objective is to maximize the number of routers in which the ECMP splits is done. This problem is a combinatorial optimization problem, and it may be formulated as a MIP.

A. Problem formulation

To formulate the problem as a MIP we use the formulation presented in [2] (see also [6-9]) that searches for shortest-path routing patterns (and associated link weight systems) optimizing simple traffic engineering criteria for the nominal state of the network. Below we re-write this formulation introducing a different objective function. We use the following variables:

- $x = (x_{et} \geq 0 : e \in E, t \in V)$:
  vector of continuous flow variables;
  x_{et} denotes the total flow destined to node t
  realized on link e
- $u = (u_{et} \in \{0,1\} : e \in E, t \in V)$:
  binary vector of routing variables;
  $u_{et} = 1$ if, and only if, link e is on a shortest path
  to destination t
- $y = (y_{vt} \geq 0 : v, t \in V)$:
  vector of continuous flow variables;
  y_{vt} denotes volume of total flow assigned to each link e
  used for sending the traffic outgoing from node v
  to destination t
- $w = (1 \leq w_e \leq W : e \in E)$:
  vector of continuous variables representing link weights
  (W – maximum weight)
- $r = (r_{vt} \geq 0 : v, t \in V)$:
  vector of variables representing lengths of the shortest paths;
  r_{vt} denotes the length of the shortest path from node v
  to destination t, calculated according to link weights w
- $q = (0 \leq q_{vt} \leq 1 : v, t \in V)$:
  vector of binary variables indicating whether router v
  applies ECMP split for the destination t;
  $q_{vt} = 1$ if, and only if, there are at least two links outgoing
  from v that belong to the shortest paths to destination t
  (ECMP is used); 0 otherwise.

According to the above definitions the variables are supposed to fulfill the following relations:

- $u_{et} = 0$ implies that $x_{et} = 0$
- $r_{vt} = \sum_{e \in P} w_e$, where P is a shortest path
  from v to t ($u_{et} = 1$ for all $e \in P$); $r_{tt} = 0$ for each $t \in V$.

The formulation is as follows:

$$\textbf{maximize} \ \sum_{v \in \mathcal{V}} \sum_{t \in \mathcal{V}} q_{vt} \tag{1a}$$

subject to

$$\sum_{e \in \delta^+(v)} x_{et} - \sum_{e \in \delta^-(v)} x_{et} = h_{vt} \ \ v, t \in \mathcal{V} \tag{1b}$$

$$\textstyle\sum_{e\in\delta^-(t)} x_{et} = \sum_{v\in\mathcal{V}\setminus\{t\}} h_{vt} \qquad t\in\mathcal{V} \tag{1c}$$
$$\textstyle\sum_{t\in\mathcal{V}} x_{et} \le c_e \qquad e\in\mathcal{E} \tag{1d}$$
$$r_{b(e)t} + w_e - r_{a(e)t} \ge 1 - u_{et} \qquad t\in\mathcal{V},\ e\in\mathcal{E} \tag{1e}$$
$$r_{b(e)t} + w_e - r_{a(e)t} \le M(1 - u_{et}) \qquad t\in\mathcal{V},\ e\in\mathcal{E} \tag{1f}$$
$$x_{et} - y_{vt} \ge 0 \qquad v,t\in\mathcal{V},\ e\in\delta^+(v) \tag{1g}$$
$$x_{et} - y_{vt} \le M(1 - u_{et}) \qquad v,t\in\mathcal{V},\ e\in\delta^+(v) \tag{1h}$$
$$x_{et} \le M u_{et} \qquad t\in\mathcal{V},\ e\in\mathcal{E} \tag{1i}$$
$$\textstyle\sum_{e\in\delta^+(v)} u_{et} \ge 1 + q_{vt} \qquad v,t\in\mathcal{V},\ v\ne t \tag{1j}$$
$$r_{tt} = 0 \qquad t\in\mathcal{V} \tag{1k}$$
$$q_{vt}\in[0,1] \qquad t\in\mathcal{V},\ e\in\mathcal{E} \tag{1l}$$
$$u_{et}\in\{0,1\} \qquad t\in\mathcal{V},\ e\in\mathcal{E}. \tag{1m}$$

Formulation (1) specifies a multi-commodity flow optimization problem in the aggregated node-link notation (cf. [2]). Constraints (1b)–(1c) express the aggregated flow conservation conditions for the flow variables x. Constant h_{vt} denotes the volume of the requested demand from node v to node t.

Constraint (1d) does not allow the total link flow to exceed link capacity c_e.

The quantity $r_{b(e)t} + w_e - r_{a(e)t}$ in constraints (1e)–(1f) measures the difference between the length of the shortest path from $a(e)$ to t (given by $r_{a(e)t}$) and the length of the shortest path from $a(e)$ to t necessarily traversing link e (for the latter, the length of the sub-path from $b(e)$ to t is determined by $r_{b(e)t}$). Note that link e is on a shortest path to node t if, and only if, $r_{b(e)t} + w_e = r_{a(e)t}$ – this condition is enforced by constraints (1e)–(1f). Hence, the routing vector u determines the shortest paths according to the weight vector w.

Constraints (1g)–(1h) enforce that traffic outgoing from each node is distributed evenly on all shortest paths to a destination, and constraint (1i) assures that traffic is not routed on the links that do not belong to the shortest paths, i.e., constraint (1j) enforces traffic destined to node t to use only the links $e \in E$ allowed by the routing configuration specified by vector u (i.e., the links with $u_{et} = 1$). Constraint (1j) allows maximized variables q to be ones if only the shortest paths split in the considered node.

Formulations analogous to (1) were published in the literature (see [10-13]). In [14], a non-linear formulation of the considered problem is described. The basic formulation presented in [2] is able to provide good quality lower bounds calculated using its linear relaxation and uses less constraints as compared to other models. However, a direct use of standard MIP solvers to IFR/ECMP can fail already for rather small networks with, say, 10 nodes.

Notice that if a router uses just one outgoing link to transit traffic to a specific destination and this particular link fails, the router is unable to transfer packets until the standard SPR re-convergence is performed. Since restoration based on the ECMP paths cannot be used in such a case, another fast restoration mechanism should be used instead. Authors of [15] (see also [16]) proposed one such mechanism, called loop-free alternatives described in the next section.

3. Loop-free alternates

The idea of using loop-free alternates (LFA) is to define, for each destination, an alternative next-hop address used in the case of a failure of any adjacent link. Once such a failure occurs, the detecting router transmits all packets originally sent through the failed link to the predefined alternative next-hop address of a neighboring router. Because the described mechanism concerns only the local protection, the router associated with the new next-hop address is not aware of the failure. Hence, selection of the alternate next-hop address must be done carefully enough to avoid traffic cycling, i.e., the router associated with the new next-hop address cannot send back the redirected packets to the router performing the recovery action. Inappropriate application of the IPR/LFA recovery mechanism may result in flow loops. Therefore, we can only choose the next-hop addresses of the routers that can transmit traffic to the destinations omitting the failing link. Clearly, to meet this requirement, the link weight system must satisfy the triangle condition [15].

Assuming destination t, the condition determining if link e, starting in node v, can be used as an alternate next-hop is as follows:

$$r_{b(e)t} \le r_{b(e)v} + r_{vt} : v,t \in V, e \in \delta^+(v). \tag{2}$$

As already mentioned, protection on the ECMP paths cannot guarantee 100% protection against single link failures. For this reason, a combined IFR/ECMP+LFA protection mechanism can be considered. Below, we study an optimization problem related to selection of LFA router addresses used to redirect packets from the links not protected through ECMP paths for the given weight system w^o. Considering a fixed vector w^o we can easily calculate the corresponding solution of formulation (1). Let us denote this solution by (u^o, r^o, x^o). Vector w^o induces a routing pattern. Thus, at this stage we know which links can be protected using ECMP and which can not. The latter one corresponds to the links that are the sole shortest path links outgoing from some router to a given destination. Clearly all paths using such links are affected by the failures and therefore for all such links we validate which of them can be protected using the LFA mechanism.

To select the next-hop address for a specific destination we determine all links starting at the same router as the failing link e and choose the addresses of the routers satisfying the triangle condition for the considered destination. Then, for each selected addresses we choose another valid next-hop address such that the resulting flow distribution does not exceed given link capacities. Although this problem is likely NP-hard, due to a small number of admissible (with respect to condition (2)) links outgoing from one router (typically several links) it seems reasonable to formulate the problem as a MIP and resolve it with the use of a standard MIP solver.

For a given link e to be protected the MIP formulation is as follows:

$$\textbf{minimize } z \tag{3a}$$

subject to

$$\sum_{t\in\mathcal{V}}\sum_{f'\in\delta^+(a(e))\setminus\{e\}} l^o_{f'tf} g_{f't} \le c_f z \qquad f \in \mathcal{E} \tag{3b}$$
$$\sum_{f\in\delta^+(a(e))\setminus\{e\}} g_{ft} = G^o_t \qquad t \in \mathcal{V} \tag{3c}$$
$$g_{ft} \le r^o_{a(e)t} + r^o_{b(e)a(e)} - r^o_{b(e)t} \qquad t \in \mathcal{V},\ e \in \mathcal{E} \tag{3d}$$
$$g_{et} = 0 \qquad t \in \mathcal{V} \tag{3e}$$
$$g_{ft} \in \{0,1\} \qquad t \in \mathcal{V},\ f \in \mathcal{E}, \tag{3f}$$

where $l^o_{f'tf}$ is a constant representing the load of link f resulting from the selection of link f' as a LFA address to redirect traffic originally flowing through e with destination in t. This is indicated by a binary variable $g_{f't}$ equal to 1 if link f' is chosen as an LFA next hop address for a destination t; and 0 otherwise. Clearly, the weight system w^o exactly determines the vector of link loads $l^o = (l^o_{f'tf} : f', f \in E, t \in V)$.

G^o_t is a binary constant determining if traffic to destination t can be protected in the considered router by the LFA mechanism, i.e., if there exists at least one neighboring router satisfying the triangle condition.

The formulation (3) is an instance of multi-knapsack problem with constrained selection of items (which in this case are redirected flows), where capacity constraint (3b) is mentioned knapsack constraint. Constraint (3c) is used to enforce selection of either 0 or 1 alternate next-hop address for the considered destination, accordingly to value of the constant G^o_t. Admissibility of a specific next-hop address (to protect link e) depends on obeying the triangle condition by the candidate protection links f', that is assured by constraint (3d).

As an objective function we consider minimization of the maximal link utilization or overload (when protection requires more capacity than available).

Similarly, as in the case of the ECMP protection, the delay related to the IPR/LFA restoration is typically much smaller than the delay of the standard SPR re-convergence. However, applicability of the ECMP or LFA mechanism can be still questionable from the operator's point of view. It is so because these mechanisms may require link overprovisioning, but as in the case of any local protection there is a tradeoff between short restoration time and large capacity requirement. Authors of [1] anticipate that the two discussed IFR mechanisms, would be able to provide protection in around 80% cases of single link failures. To assure 100% protection, another mechanism is needed. One such mechanism, called multi-hop repair paths (MHRP), is discussed in the next section.

4. Multi-hop repair paths

MHRP is a protection mechanism applied when a disrupted flow cannot be rerouted on a protection path using ECMP or LFA, i.e., in a case when there is exactly one outgoing link, say e, (used by the disrupted flow) belonging to the shortest path and none of the links different than e and outgoing from the originating node of link e obey the triangle condition (2). Clearly, it may happen that none of both mechanisms can be applied. In order to assure 100% reliability, we need to consider multi-hop reroute paths that do not necessarily terminate in the routers adjacent to the originating node of the failing link. In general, following [1], in most of the cases it can be achieved when using two-hops reroute paths only. Although, the ECMP and LFA mechanisms can be adopted in IP quite easily, e.g., as simple extensions to the router's software, the third protection mechanism requires more sophisticated solutions, like MPLS tunnels (cf. [17]), multiple FIBs (cf. [18]), or the mechanism of the alternative shortest paths (cf. [19]).

Below, we investigate an application of MHRP as a mechanism complementary to the ECMP+LFA protection mechanism. Namely, a router detecting a failure performs the following restoration actions for each destination:

1) The router verifies if there exists at least one surviving ECMP path to the destination.
   If this is the case, the flow from the broken link is evenly distributed over all the surviving ECMP path links.
2) If no ECMP path exists, the router checks if an alternative next-hop is defined for this destination.
   If this is the case, the flow is redirected to the defined next-hop address.
3) If neither ECMP nor LFA can be used, the router establishes a tunnel to a remote router used as a multi-hop repair path.

Clearly, using such a three-fold protection mechanism 100% protection can be achieved.

Now, we consider an optimization problem computing the multi-hop repair paths for all links that are protected neither through ECMP nor LFA for the given link weight system w^o. Again, the considered problem may be formulated as a MIP.

Assuming that link e is failing, the formulation is as follows:

$$\textbf{minimize } (a+b)\alpha + \sum_{f\in\mathcal{E}}\sum_{t\in\mathcal{V}} s^e_{ft} + z \tag{4a}$$

subject to

$$\sum_{f\in\delta^+(v)} s^e_{ft} - \sum_{f\in\delta^-(v)} s^e_{ft} = -S_{vt} \qquad v,t \in \mathcal{V} \tag{4b}$$
$$\sum_{f\in\delta^+(a(e))\setminus\{e\}} s^e_{ft} = 1 \qquad t \in \mathcal{V},\ f \in \mathcal{E} \tag{4c}$$
$$\sum_{t\in\mathcal{V}} x^o_{et} s^e_{ft} + \sum_{t\in\mathcal{V}}\sum_{v\in\mathcal{V}} n^o_{vtf} S_{vt} \le c_f z \quad f \in \mathcal{E} \tag{4d}$$
$$s^e_{et} = 0 \qquad t \in \mathcal{V} \tag{4e}$$
$$S^e_{a(e)t} = 0 \qquad t \in \mathcal{V} \tag{4f}$$
$$z \le 1 + aM \tag{4g}$$
$$z \le b + 1 \tag{4h}$$
$$b \ge 0 \tag{4i}$$
$$s^e_{ft} \in \{0,1\} \qquad t \in \mathcal{V},\ f \in \mathcal{E} \tag{4j}$$
$$S^e_{vt} \in \{0,1\} \qquad v,t \in \mathcal{V}. \tag{4k}$$
$$a \in \{0,1\} \tag{4l}$$

The above problem is solved for each link e that for a current weight system w^o cannot be protected either by ECMP or LFA mechanisms. As a result of solving Problem (4) we obtain multi-hop routes that can be used to protect these links.

The resulting tunnels are described by means of binary variables $s^e = (s^e_{ft} : f \in E, t \in V)$ specifying, for a given link e to be protected, whenever link f is a part of a multi-hop repair path to destination router t. Using these variables we write down the flow conservation constraints (4b)–(4c) assuring the continuity of the flow carried through the tunnel. Since the end-nodes of the multi-hop repair paths are not predetermined (they are optimized), they are expressed by the values of the binary variables S^e_{vt}. A positive value of variable S^e_{vt} states that router v is the terminating node of the protection tunnel used for destination t. Calculating the link loads one must take into account two elements: the load resulting from the selected multi-hop repair paths, and the load resulting from the ECMP paths from the selected end-nodes of the multi-hop repair paths to the traffic destinations. Constraint (4d) is a capacity constraint taking into account link loads resulting from the multi-hop repair paths, determined by the first term in the sum, where x^o_{et} is a constant representing an amount of traffic sent in a normal state to destination t on link e (link e is the link to be protected). The second term of the sum (4d) represents the link load components induced by the flows leaving MHRP paths and sent to the destination using shortest paths. This is achieved using constant n^o_{vtf} which is equal to the amount of traffic sent through link f to destination t if we decide to use an MHRP path that terminates in v. Constraints (4e) and (4f) assure that protected link e cannot be used in any MHRP path and that MHRP path cannot terminate in the originating node of link e respectively.

When solving Problem (4) our primary goal is to find a tunnel configuration that does not cause link overloads, or when it is unavoidable, tries to minimize the overload given as z. This is reflected in the objective function by the component $(a+b)\alpha$. α is a big-M constant equal to $|E|\cdot|V|$, and a is a binary variable equal to 1 if at least one link becomes overloaded $z>1$ (cf. constraint (4g)). b was used to linearize (cf. constraint (4h)) the original version of $(a+b)\alpha$ given as a non-linear expression αaz focusing optimization on the minimization of z whenever $z>1$. This is our primary goal. Whenever z goes below 1 (there are no overloaded links) the optimization switches to minimization of the total length of MHRP paths (due to $(a+b)\alpha = 0$. The second goal is expressed in the objective function using component $\sum_{f\in E}\sum_{t\in V} s_{ft}$. If it happens that for two different solutions their resulting total lengths of MHRP paths are equivalent, the solution with smaller value of z if preferable (thanks to sole z that is also minimized in the objective function).

5. Resolving methods

In this section we present an efficient heuristic method for resolving the general SPR routing optimization problem with combined ECMP+LFA+MHRP protection. The problem consists in determining a weight system (inducing the ECMP protection), valid LFA next-hop addresses, and the MHRP paths, such that all links in the network are protected through one of the considered IFR mechanisms. We assume that the link capacities and the traffic demands are given. To resolve the problem we adopt a heuristic approach that consists in decomposing the resolution process into three phases.

Algorithm 1 Resolving method

P1: Find a weight system w (e.g., resolving the formulation (1)). Denote the resulting solution by w^o. For w^o calculate SPR flow distribution pattern in the normal state. Notice that the given weight system already determines where the ECMP protection can be used.

P2: Denote by S the set of all unprotected links (associated with single shortest paths outgoing from a specific router). For each unprotected link $e \in S$ resolve formulation (3) and remove links that can be protected by the LFA alternate next-hop addresses from the set S.

P3: For each of the remaining links $e \in S$ resolve formulation (4). Evaluate the obtained nominal and protection routing patterns with respect to objective function (5). If the obtained value is unsatisfactory go back to phase **P1** in order to find another candidate link weight system different from the previously found; otherwise, stop the procedure.

The resulting procedure (cf. Algorithm 1) searches for a weight system that minimizes the value of the penalty function given below:

$$\alpha az + \beta \sum_{e\in E}\sum_{t\in V} m^e_t + \sum_{e\in E}\sum_{f\in E}\sum_{t\in V} s^e_{ft} + z, \quad (5)$$

The role of the first component is to minimize the value of z regardless to the number of MHRP paths used and their total length if $z>1$. Again, a is a binary variable reflecting the existence of link overloads. More precisely $a=1$ if $z>1$ – overloaded links are identified, and $a=0$ otherwise. Since the coefficient of the first term α is set to $|E|\cdot|V|\cdot(\beta+1)$, i.e., the value bigger than any possible value of the other remaining components, it dominates them all. Therefore, the resolution procedure in its early stage tries to find the solutions that do not contain overloaded links with a little regard to the number of MHRP paths used.

When the solution with $z \leq 1$ is found it is evaluated with respect to the number of used MHRP paths in the first place. For this purpose we introduce a binary variable $m^e_t, e \in E, t \in V$ equal to 1 if traffic sent on link e to the destination t is protected by the MHRP path, and 0 otherwise. Constant $\beta=|E|^2\cdot|V|+1$ is set big enough in order to dominate the remaining components.

The third term in (5) is responsible for minimization of the total length of MHRP paths used. As a result, a weight system with the smallest total length of MHRP paths is evaluated as most profitable. Assuming that the value of z does not exceed 1 in the currently considered weight system (no overloaded links are present), the total length of the MHRP paths is expected to dominate the last term, that minimizes the maximal link utilization.

Notice that the general form of Algorithm 1 admits to use any method for searching the global space of feasible link weight systems. For this purpose any metaheuristic or exact branch-and-bound-like procedure can be applied. In the second phase the method recalculates a flow distribution pattern with the ECMP protection applied in all routers where at least two shortest paths exist to a given destination. Next, the procedure verifies if non-protected links may utilize the LFA next-hop mechanism. If it is the case they are selected so to minimize the value of *z*.

The method we used for the selection of w^o used in **P1** is based on the Local Search meta-heuristic. It starts from a base weight system with all link weights equal to 5. In each iteration we modify the link weights of 30% randomly chosen links. Each time a better weight system is found it replaces the base weight system. The weights are modified by adding random numbers chosen uniformly from a range $[-R;R]$.

More precisely, $w_e = \max\{1, w_e + rand(-R,R)\}$. Initially R is set to 5 and every 200 iterations is decreased by 1. When R reaches the value of 0 the process is stopped and the whole method is restarted. Additionally, before each iteration weights that correspond to most heavily loaded links are increased by $\frac{R}{2}$, if $z > 1$ or the number of MHRP connections in the current best solution is equal to zero. The same changes apply to weights which correspond to links that are not protected either by ECMP or by LFA, when $z \leq 1$ and the current best solution requires the MHRP connections to be established in order to provide the full protection of affected flows.

6. Numerical results

In our numerical experiments we tested the efficiency of the method presented in Section 5. In the implementation the subproblems of phases **P2** and **P3** were solved using the branch-and-cut algorithm of CPLEX 10.1 (cf. [20]). We tested the method using 17 different network examples. More details about the instances used together with numerical results are presented in *Table 1* (down).

For each considered example network the table contains the information about the number of nodes and links in the network. Moreover, it contains results obtained for three different methods for the weight system selection. The first strategy (called "Heuristic") refers to the heuristic method described in Section 5. The second and the third method are simple approaches to the weight setting problem, i.e., all link weights are uniformly set to 1 ($w_e = 1, e \in E$) or proportionally to the inverse of link capacity ($w_e = \frac{K}{c_e}, e \in E, K = \max_e c_e$).

For each of those strategies three different numerical values were collected: *z* – fraction of the capacity used on the most heavily loaded link, MHRP – number of the MHRP connections that have to be established in order to assure a full protection of all demands, MHRP length – sum of the lengths of the established MHRP connections.

The table also contains the information about the number of iterations the heuristic method managed to perform for each problem instance during a 3-hour period of time.

Table 1. Network data and results for the three methods of choosing weight system W

	Number of		**Heuristic**				$w_e = 1, e \in E$			$w_e = \frac{K}{c_e}, e \in E$		
	nodes	links	z	MHRP	MHRP length	iterations	z	MHRP	MHRP length	z	MHRP	MHRP length
network_6	6	18	2.26	0	0	69900	3.01	2	5	3.18	4	10
network_8	8	24	2.61	3	6	21400	4.02	12	24	3.23	9	26
network_10	10	30	2.13	7	14	22600	2.83	10	21	4.51	12	26
network_12	12	36	2.05	27	72	12000	2.23	29	64	3.60	32	79
network_14	14	42	2.48	54	120	7900	3.52	69	144	3.13	38	91
network_16	16	48	3.04	68	164	5800	3.90	66	154	5.43	43	127
artificial_6n	6	28	1.25	0	0	6700	1.25	0	0	2.18	1	2
polska_12n	12	36	2.49	14	43	18400	2.74	26	80	3.37	29	81
cost239 defaultTM	11	52	0.71	0	0	23300	1.08	2	4	1.08	2	4
cost239 ecmpScaledTM	11	52	0.83	0	0	10500	1.14	2	4	1.14	2	4
cost239 sprScaledTM	11	52	0.50	0	0	14200	0.68	2	4	0.68	2	4
geant defaultTM	19	60	0.93	46	98	13700	0.82	124	248	0.92	75	152
labnet ecmpScaledTM	21	106	0.99	1	2	5700	0.95	10	20	0.95	10	20
labnet sprScaledTM	21	106	0.69	1	2	3200	0.65	10	20	0.65	10	20
nobel ecmpScaledTM	27	82	0.99	216	453	2700	0.97	293	590	0.97	293	590
nobel sprScaledTM	27	82	0.99	183	381	6400	0.98	293	590	0.98	293	590

The experiments were performed on a machine running Windows XP on an Intel Pentium 4 (3.0 GHz) processor with 1.96 GB of RAM. The CPU usage limit was set to 50% and the execution time of the procedure was limited to 3 hours for each particular run.

Analyzing the results presented in *Table 1* obtained for simple weight system selections with the results provided by our method we can conclude that optimization of weights is highly recommended. Neither the equal weights nor the weights proportional to the inverses of link capacities provide results comparable to those achieved using the system calculated by our method.

The greatest differences, from an operator's point of view, can be seen as far as *cost239_defaultTM, cost239*

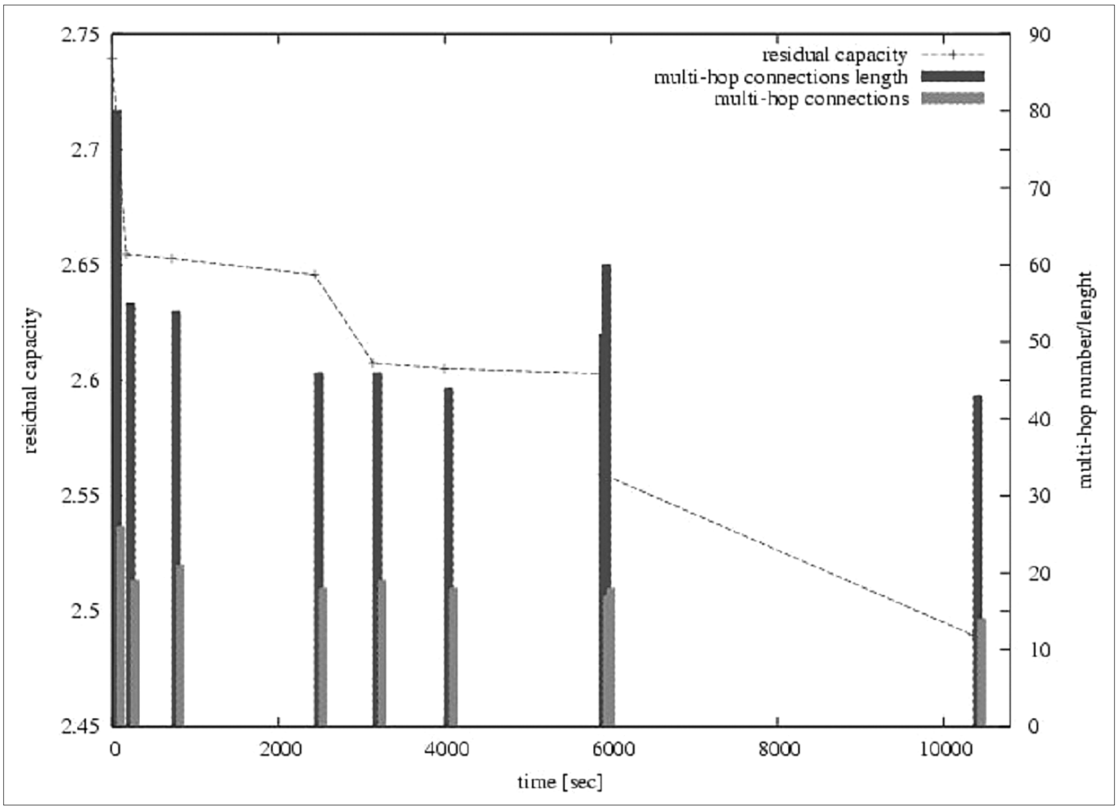


Figure 1. Procedure performance for network12

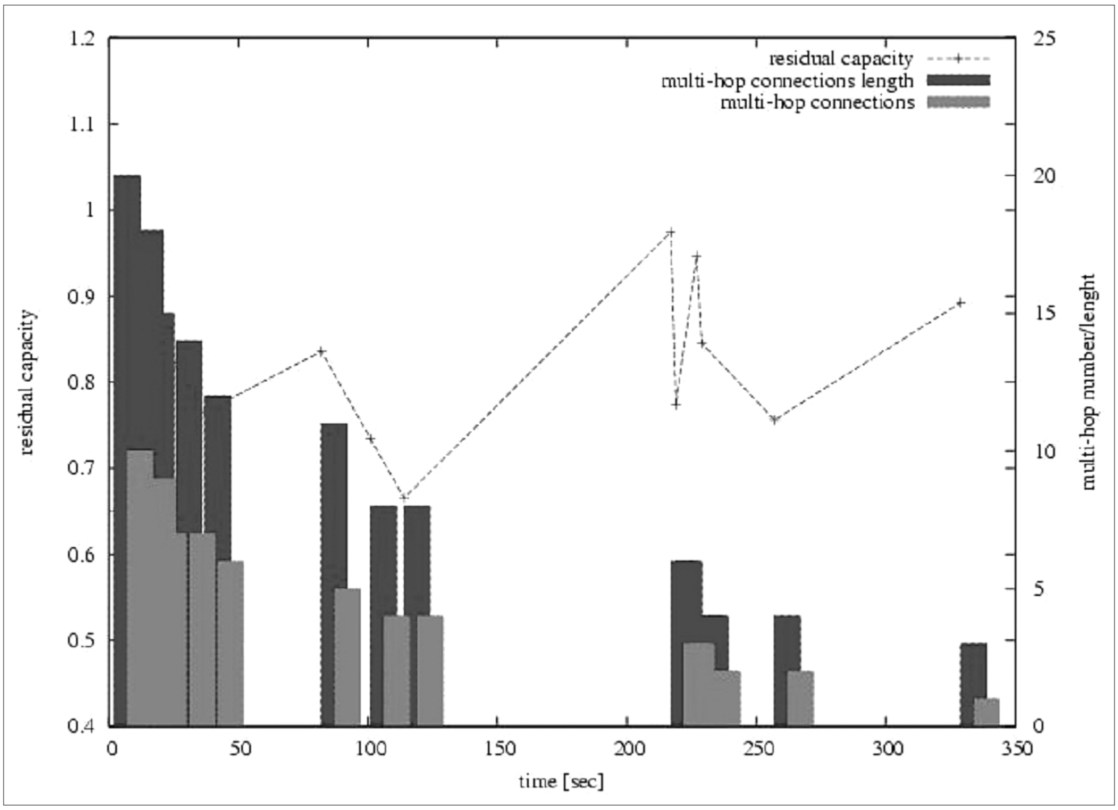


Figure 2. Procedure performance for labnet_sprscaled

_ecmpScaledTM, and cost239_sprScaledTM networks are concerned. In the first two cases the proposed heuristic approach provides a solution that does not cause network overloads. Note that, this is not the case with the naive approaches where provided solutions are characterized by $z>1$. In the third network, the proposed method finds a solution that does not require the MHRP connections to be implemented, and this fact significantly limits the financial burden imposed on the operator.

Concerning the behavior of our algorithm, the conducted experiments allow for concluding that the method works well for both overloaded ($z>1$) and non-overloaded ($z\leq1$) networks. This is illustrated in two figures (on the previous page) that present the behavior of the algorithm in the function of time.

Figure 1 contains the performance data of the method while solving *network12* problem instance, while *Figure 2* deals with problem *labnet_sprScaledTM*. Each column in the diagrams corresponds to a moment when a better solution was found by the heuristic method. Each solution is characterized by the number of the required MHRP connections (grey column), sum of their lengths (black column) and the value of z (solid line). In Figure 1, when $z>1$, the algorithm optimizes z regardless of the number and lengths of the MHRP connections. On the other hand, Figure 2 shows that the procedure focuses on minimization of the number of the MHRP connections when $z\leq1$. Note that, the latter figure does not present the whole optimization process. The reason is because the following improvements were much less frequent, and extending the x-axis to 3 hours would blur the presented best solutions found in the early iterations of the algorithm.

The final results for the problem instance can be found in *Table 1*.

7. Conclusion

In this article we have considered three different protection/restoration IP mechanisms that aim at improving the reactivity of IP routing in the case of failures. These are Equal Cost Multiple Path (ECMP) based flow restoration, Loop-Free Alternates (LFA) and Multi-Hop Repair Paths (MHRP). The considered IP fast reroute (IFR) mechanisms compensate the negative impact of relatively slow convergence of the routing protocol after a failure is detected. In fact, IFR framework, if deployed in the network, eliminates the need for the post-failure routing protocol re-convergence. The detecting routers simply locally reroute affected flows with the guarantee that the rerouted flows will reach the destination and will not be lost (e.g., due to transient flow loops). This is in general very important for different types of streaming services requiring low packet loss ratio, low jitter, etc.

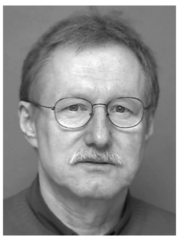
The considered restoration mechanisms, when deployed in the network, could be taken into account while designing the network flows distribution in order to optimize the utilization of available resources. In this context the article focuses on optimization techniques that can be used to solve related network design problems. First, each presented mechanism is considered separately and exact resolution approach based on solving the MIP problem formulations using a Branch-and-Cut algorithm is discussed. Then all IFR mechanisms are assumed to coexist in order to assure 100% reliability. Due to the complexity of the resulting problem, exact optimization methods may not be applicable. Therefore, an efficient heuristic method based on a three phase problem decomposition is proposed. It focuses on determination of a link weight system for which the network flows do not cause link overloads, and such that the number of flows protected either by ECMP or LFA is maximal.

Analyzing the results of the numerical experiments we conclude that our heuristic algorithm allows to obtain solutions that are much better as compared to the solutions with non-optimized systems of link weights, both in terms of network resources utilization and the number of flows protected using ECMP and LFA mechanisms.

Authors

MATEUSZ DZIDA received his MSc and PhD, both from Warsaw University of Technology, in year 2003 and 2009, respectively. His scientific interests focus on designing telecommunication networks and architectures of the post-IP Internet networks. He is an author and co-author of over 30 conference and journal publications in these domains.

MATEUSZ ZOTKIEWICZ is a PhD. student at the Computer Networks and Switching Division at the Institute of Telecommunications, Warsaw University of Technology, Poland, and at Institut TELECOM, TELECOM SudParis, France. He received the M.Sc. degree in telecommunications in 2007 from Warsaw University of Technology. His research interests concentrate on modeling and design of telecommunications networks.



MICHAL PIÓRO is a professor and head of the Computer Networks and Switching Division at the Institute of Telecommunications, Warsaw University of Technology, Poland, and a professor at Lund University, Sweden. He received the PhD degree in telecommunications in 1979 and the DSc degree in 1990, both from the Warsaw University of Technology. In 2002 he received a Polish State Professorship. His research interests concentrate on modeling, design and performance evaluation of telecommunication systems. He is an author of four books and more than 150 technical papers presented in the telecommunication journals and conference proceedings. He has lead many research projects for telecom industry in the field of network modeling, design, and performance analysis. He is deeply involved in international research projects including FP7, Celtic and COST projects.

References

[1] M. Shand, S. Bryant,
"IP Fast Reroute framework,"
Tech. Rep.,
Internet Draft draft-ietf-rtgwg-ipfrr-framework-13, 2009.

[2] M. Pióro, D. Medhi,
Routing, Flow, and Capacity Design
in Communication and Computer Networks.
Morgan Kaufman, 2004.

[3] P. Francois, C. Filsfils, J. Evans, O. Bonaventure,
"Achieving subsecond IGP convergence
in large IP networks,"
In: ACM SIGCOMM, 2005.

[4] G. Iannaccone, C.-N. Chuah, S. Bhattacharrya, C. Diot,
"Feasibility of IP restoration in a tier-1 backbone,"
Tech. Rep., sprint ATL Research Report,
Nr. RR03-ATL-030666, Sprint ATL, March 2003.

[5] A. Wielosz, K. Islam,
"Achieving fast restoration times in IP networks
for IPTV video transport,"
In: NAB, 2007.

[6] M. Dzida,
"Optimization models for designing resilient
intra-domain IP routing,"
Ph.D. dissertation, 2008.

[7] A. Tomaszwski, M. Pióro, M. Dzida,
M. Mycek, M. Zagozdzon,
"Valid inequalities for a shortest-path routing
optimization problem,"
In: Int. Network Optimization Conf. (INOC),
Spa, Belgium, 2007.

[8] A. Tomaszewski, M. Pióro, M. Dzida, M. Zagozdzon,
"Optimization of administrative weights in
IP networks using the branch-and-cut approach,"
In: The 2nd Int. Network Optimization Conf. (INOC),
Lisbon, Portugal, 2005.

[9] M. Dzida, M. Zagozdzon, M. Pióro,
"Optimization of resilient IP networks with
shortest path routing,"
In: 6th Int. Workshop on Design and
Reliable Communication Networks (DRCN),
La Rochelle, France, 2007.

[10] K. Holmberg, D. Yuan,
"Optimization of Internet protocol network design
and routing,"
Networks, Vol. 43, No. 1, pp.39–53., 2004.

[11] L.D. Giovanni, B. Fortz, M. Labbe,
"A lower bound for
the internet protocol network design,"
In: Proc. of the Int. Network Optimization Conf. (INOC),
Lisbon, Portugal, 2005.
pp. B.2. 401–408.

[12] A. Bley, M. Grötschel, R. Wessäly,
"Design of broadband virtual private networks:
Model and heuristics for the b-win,"
In: Proc. DIMACS Workshop on Robust Communication Networks and Survivability,
AMSDIMACS Series, 53, 1998, pp.1–16.

[13] P. Broström, K. Holmberg,
"Multiobjective design of survivable IP networks,"
In: P. Broström – PhD Dissertation,
Optimization Models and Methods for
Telecommunication Networks Using OSPF.
Linkoeping University, 2006.

[14] N. Bourquia, W. Ben-Ameur, E. Gourdin, P. Tolla,
"Optimal shortest path routing for Internet networks,"
In: Proc. of the 1st Int. Network Optimiz. Conf. (INOC),
Evry-Paris, France, pp.119–125., 2003.

[15] A. Atlas, A. Zinin,
"Basic specification for IP Fast-Reroute:
Loop-free Alternates,"
Tech. Rep., Proposed standard, 2008.

[16] S. Nelakuditi, S. Lee, Y. Yu, Z. Zhang, C. Chuah,
"Fast local rerouting for
handling transient link failures,"
IEEE/ACM Trans. Netw., Vol. 15, No. 2,
pp.359–372., 2007.

[17] S. Bryant, C. Filsfils, S. Previdi, M. Shand,
"IP fast reroute using tunnels,"
Tech. Rep., 2004, Internet Draft (work in progrees).

[18] A. Kvalbein, A.F. Hansen, T. Cicic,
S. Gjessing, O. Lysne,
"Fast IP network recovery using multiple routing
configurations,"
In: 25th IEEE International Conference on Computer
Communications (INFOCOM), 2006.

[19] A. Tian, N. Shen,
"Fast Reroute using alternative shortest paths,"
Tech. Rep., 2004, Internet Draft (work in progrees).

[20] CPLEX, CPLEX User's Manual.
ILOG, 1999.

Network resilience requirements and algorithms for multicasting and broadcasting digital TV

JUAN PEDRO FERNANDEZ-PALACIOS GIMENEZ, RAUL DUQUE
Telefónica Investigación y Desarrollo, Telefónica, Spain
{jpfpg, e.rt1}@tid.es

TIBOR CINKLER, PÉTER FODOR, PÉTER SOPRONI, MARCELL PERÉNYI, JÁNOS TAPOLCAI, ANDRÁS GULYÁS, GYULA SALLAI
Department of Telecommunications and Media Informatics, Budapest University of Technology and Economics, Hungary
{cinkler, fodorp, soproni, perenyim, tapolcai, gulyas, sallai}@tmit.bme.hu

JAVIER ARACIL RICO
Departamento de Ingeniera Informática, Universidad Autonoma de Madrid, Spain
javier.aracil@uam.es

RUPERT GRUENZINGER
Alcatel Lucent, Nürnberg, Germany
rgruenzinger@alcatel-lucent.com

Keywords: digital TV, metro and core networks, network resilience

In this paper we give an overview of the network architecture and of the resilience requirements for both metro and core networks. We propose and evaluate various protection mechanisms for the metro network and restoration mechanisms for the core network, and evaluate them by extensive simulations, showing that the quality of the tree obtained after the failure is much less important than the restoration time, that strongly depends on the algorithm that determines the new tree. We evaluate how the interrupts of protection switching and restoration affect the experienced quality of service for different video formats and resolutions.

1. Introduction and motivation

Broadcast TV is one of the key services offered by most telecom operators nowadays. This service presents strict requirements in terms of survivability since a very high number of users would be affected by a failure in its distribution. Therefore, broadcast TV transport connections should be able to face multiple simultaneous link failures in order to achieve 99,999% availability.

Currently, there exist multiple transport alternatives for broadcast video distribution in metro and core networks. Operators can choose among layer 1 (NG-SDH, OTH), layer 2 (PBB, PBB-TE, T-MPLS, RPR) and layer 3 (IP/MPLS) transport solutions. While most of these technologies already include protection mechanisms for multicast connections, the development of restoration mechanisms for multicast traffic is still an open issue. In this context, this paper provides a performance ana-

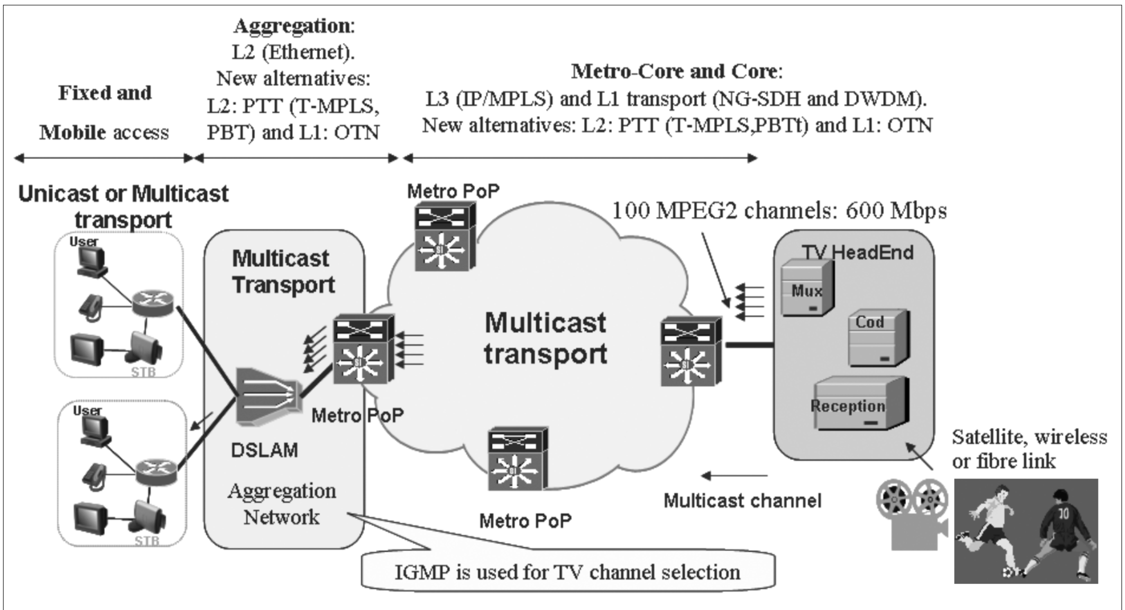


Figure 1. Multicast distribution over metro and core networks [1]

	Bandwidth [Mbps]			
Type of service	**Peak down**	**Peak up**	**Mean down**	**Mean up**
Video Broadcast 0 (Mobility TV)	0.384	0	0.256	0
Video Broadcast 1 (SDTV mpeg2)	6	0	6	0
Video Broadcast 2 (SDTV mpeg4)	3	0	3	0
Video Broadcast 3 (HDTV mpeg2)	20	0	20	0
Video Broadcast 4 (HDTV mpeg4)	10	0	10	0

Table 1. The bandwidth requirements of the most common Broadcast TC services

lysis of resilience solutions for broadcast TV services in metro and core networks based on a combination of both multicast transport and restoration.

2. Distribution architecture and requirements for broadcast TV services

Broadcast TV service is based on the digitalization, compression and transmission and decoding of video signals over IP networks.

The distribution architecture of this type of services is explained as follows *(Figure 1)*:

- A regional TV head-end receives all the national and international TV channels by means of satellite, fibre or wireless connections.
- Then TV channels are broadcasted within the regional network following two steps: a multicast transport over IP in the backbone and a multicast transport over Layer 2 in the metro area.
- Therefore, each DSLAM can receive all TV channels. However, due to the bandwidth limitations of the access segment, the end user only receives the selected channels.

Table 1 shows the bandwidth requirements of most common Broadcast TV services:

Broadcast TV traffic volume does not depend on the number of customers but on the number, definition and coding of TV channels. So, TV traffic volume would be similar in the metro access and metro core segments. For example, 100 HDTV channels, with MPEG 4 coding, will need 1 Gbps from the TV head-end to the rest of Service PoPs in the metro core, and from the Service PoP to the access nodes in the metro access.

Considering resilience issues of IP-TV services, we can mention the following topics:

- This service requires resilience mechanisms for high capacity multicast traffic.
  Total capacity depends on the number of TV channels, and the capacity per channel (definition and coding technique).
- A very high number of users would be affected by a total service cut. Therefore, broadcast TV transport connections should be able to face multiple simultaneous link failures.
- Recovery speed should be lower than 50 ms in case of a single failure and lower than 1 s in case of multiple failures [1].
- Neither retransmission nor FEC techniques are used due to strict jitter requirements.
  Therefore very low packet loss rates are required.

Table 2 shows a summary of the main resilience requirements for Broadcast TV services.

3. Transport and resilience alternatives for Broadcast TV distribution

A possible alternative for Broadcast TV distribution is Multi-Protocol Label Switching over IP (IP/MPLS) whose Fast Reroute (FRR) mechanism supports both unicast and multicast traffic. FRR is based on pre-planned protection schemes which are specially adapted to single failures situations.

Virtual Private LAN Service (VPLS) is a layer 2 multipoint VPN that allows multiple sites to be connected in a single bridged domain over a provider managed IP/MPLS network. So, MPLS-based Fast Reroute mechanism can be used to ensure sub-50 ms protection [2].

Provider Backbone Bridges (PBB) support multicast over Carrier Ethernet networks. Packets are forwarded and replicated according to their Backbone MAC. Service is still connectionless, flooding is used when destination MAC addresses are not recognized, and spanning tree protocol (STP) is used to prevent loops.

Provider Backbone Bridges – Traffic Engineering (PBB-TE) is solving the survivability problems of PBB by disabling STP and implementing 50 ms recovery with fast 802.1ag CFM OAM. However 802.1ag only implements protection mechanisms for unicast traffic. Restoration mechanisms are not available yet due to the lack of a distributed control plane. Current PBB-TE standards work does not address P2MP architectures.

Table 2. The main resilience requirements for Broadcast TV services

	Multicast	Multiple failure survivability	Max. Jitter	Recovery Time	
				Single failure	Multiple failures
IP-TV	YES	YES	20 ms	50 ms	1 s

Provider Link State Bridging (PLSB) is an extension of PBB-TE which supports multicast transport. It could include either protection or restoration mechanisms. However, there are no standardization initiatives for PLSB yet.

Resilient Packet Ring (RPR) that is based on a dual counter-rotating ring can be used for video multicasting.

Transport-Multiprotocol Label Switching (T-MPLS) is a connection-oriented packet switched transport layer technology. T-MPLS supports almost all the protection mechanisms of typical transport networks with sub-50 ms protection switching time (relying on hardware-based OAM implementation). Some open issues of this technology are the standardization of OAM and resilience mechanisms for P2MP connections and the T-MPLS control plane definition.

Next Generation-Synchronous Digital Hierarchy (NG-SDH) is one of the most extended technologies in metro-core and core networks, and is mainly aimed at providing a bridging point between the legacy TDM architectures and new IP and Ethernet transport networks. Currently, NG-SDH networks can also include a GMPLS control plane. According to it, both legacy protection and new restoration mechanisms can be implemented.

There are two more technologies to provide Broadcast TV service. *Optical Circuit Switching* (OCS) implements both dedicated and shared protection mechanisms in order to allow sub-50 ms lambda recovery. If optical nodes include add&drop functionalities, then such mechanisms could be easily used for multicast protection. It is important to highlight that the implementation of restoration mechanisms in all optical networks is much more complex. *Optical Burst Switching* (OBS) is another possibility. It can be seen as a long-term alternative for video transport and it could be a good mechanism for fast video downloads. From a resilience-based perspective, OBS is more fault-tolerant than OCS.

Table 3 summarizes some characteristics of the previous technical transport alternatives for delivering IP-TV traffic.

Although T-MPLS and PBB-TE do not support multicast distribution solution by now, there are a lot of technical alternatives providing it, as it can be observed in the table. Then, we will focus in multicast mechanisms for delivering Broadcast TV services. In addition, it is checked that it achieves bandwidth savings in comparison with unicast distribution solution.

In this paper, we propose and evaluate different resilience mechanisms for multicast connections.

4. Resilience mechanisms for metro networks

Once we have chosen multicast transport as the most efficient one in terms of resource consumption, we aim to quantify the possible differences between two resilience strategies, such as 1+1 protection and restoration, measuring the provided service availability.

4.1 Case studies

In the metropolitan area, we have developed the simulations over the Madrid's Metro-Core reference scenario, presented in *Figure 2*.

As we mentioned previously, we will focus on multicast transport solution. So, a point-to-multipoint connection is established from the TV head node to the metro access nodes to transport all the TV channels.

In this part of the study, we have analyzed the following case studies:

- **Multicast distribution combined with protection:** the use of global 1+1 protection for the multicast tree is analyzed in a multiple failure scenario.

	Multicast	Recovery speed		Resilience Standards
		1+1 Protection	Restoration	
IP/MPLS	YES	< 50 ms	IP rerouting process (sec or min)	Unicast: RFC 4090; Multicast: work in progress
PLSB	YES	TBD	TBD	None
RPR	YES	< 50 ms	NO	802.17
VPLS	YES	< 50 ms	IP rerouting process (sec or min)	Just like MPLS
PBB	YES	NO	STP convergence (sec or min)	802.1D (STP)
PBB-TE	Under definition	< 50 ms	NO	802.1ag
T-MPLS	Under definition	< 50 ms	NO	Y.1720/G.8131
OBS	YES	TBD	TBD	None
NG-SDH (GMPLS)	YES	< 50 ms	< 100 ms	Y.1720/G.8131
OCS (GMPLS)	YES	< 50 ms	seconds	None

Table 3. Some characteristics of the previous technical transport alternatives for delivering IP-TV traffic

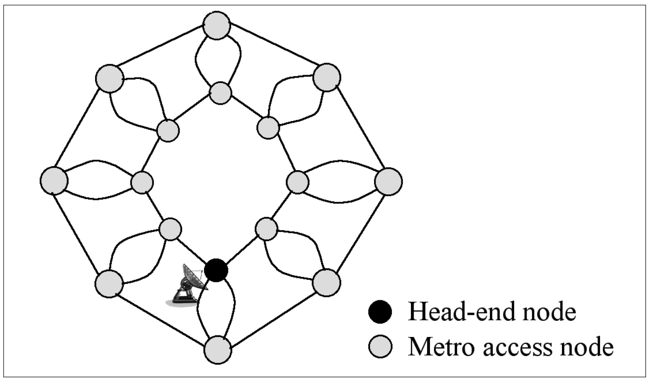


Figure 2. Madrid's Metro-Core reference network

- **Multicast distribution combined with restoration:** in this case, we analyze a global restoration mechanism to recover the branches of the multicast tree in a situation with multiple failures.

4.2 Recovery procedures

In this section we are going to explain the operation of the two resilience mechanisms that we expect to compare in terms of service availability:

- **1+1 Protection**: it consists of pre-calculating a working multicast tree and a backup multicast tree, both according to shortest path algorithm. There is another condition to compute the backup tree: it must be link-disjoint respect to the working one *(Figure 3)*.

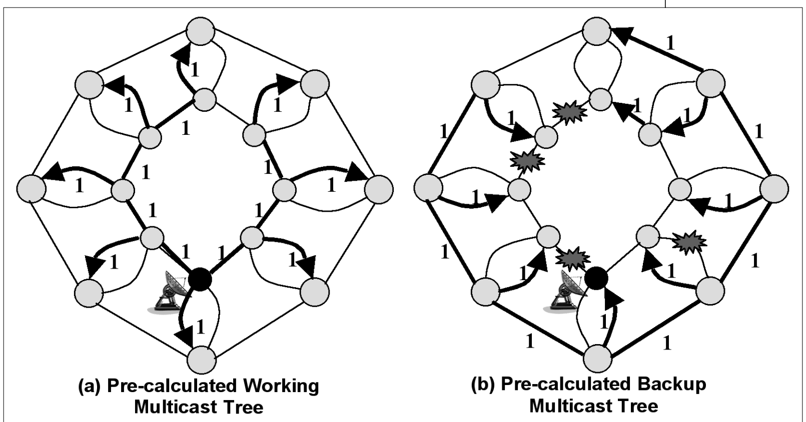


Figure 3. Multicast with 1+1 protection.

At the initial situation, we have established the pre-calculated working multicast tree (Fig. 3a). After a multiple link failure event, if there is an error affecting any link of the primary multicast tree, the resilience mechanism (1+1 protection) switches to the pre-calculated backup multicast tree, if possible (Fig. 3b).

In this situation, if a link failure affects any link of the backup multicast tree, the resilience mechanism tries to re-establish the pre-computed working multicast tree (Figure 3a).

- **Restoration**: as soon as the network starts to work, a multicast tree is computed and established according to shortest path algorithm (Fig. 4a). After a multiple link failure event, if there is an error affecting any link of the established multicast tree, the resilience mechanism (restoration) searches another possible multicast tree according to the available network resources (Fig. 4b), avoiding the broken links. In this manner, it is tried to maintain the IP-TV service in all the access nodes.

4.3 Study parameters

The simulations have been carried out using OMNET++ [3]. OMNET++ is an object-oriented modular discrete event network simulator. The most common application area of OMNET++ is the simulation of telecommunications networks.

The simulation model includes, as well as the network nodes and bidirectional links, a central module that computes the corresponding multicast trees depending on both the considered resilience mechanism and the available network resources (i.e., non-cut links) at every moment. Also, it calculates the number of metro access nodes that can not receive the IP-TV traffic and the period of time being out of service.

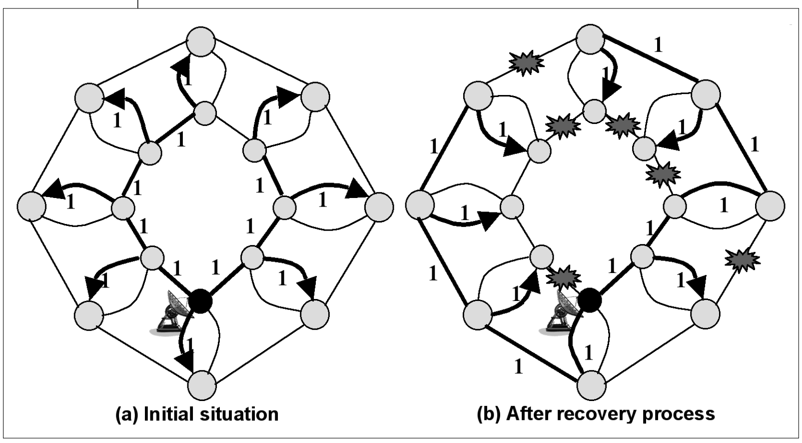
For all the simulations, we have chosen two input parameters: *mean time between link failures* (MTBF) and *mean time to repair the link failures* (MTTR).

The chosen values have been the following:

– MTBF: 30 days, 45 days, 2 months and 6 months.
– MTTR: 6 hours, 12 hours, 1 day, 2 days, 5 days, 1 week and 2 weeks.
– Simulation time: 5 years.

We have carried out simulations combining all the considered MTTR values for every MTBF one separately. The random number generator has been fed with different seeds in order to obtain statistically reliable results for each pair of MTTR and MTBF values. Specifically, we have used

Figure 4. Multicast with restoration. Recovery procedure example.



the L'Ecuyer random number generator [4] with a generation period of 2^{191} that guarantees a great amount of independent streams.

In this study, a subset of the obtained results is presented in the next section.

The output parameter has been the service unavailability percentage obtained with each considered resilience mechanism. To evaluate this parameter, every time when a subset of metro access nodes can not be reached, we measure the number of access nodes that are out of service and the period of time when they are not receiving the IP-TV traffic, and finally, we carry out the next operation:

$$SU(\%) = \frac{\sum_{j \in P_{out}} AN_{out}^{j} \cdot T_{out}^{j}}{AN_{total} \cdot T_{sim}}$$

where SU is the total service unavailability percentage, AN_{total} is the total number of metro access nodes in the network, T_{sim} is the total simulation time, AN_{out} is the number of access nodes being out of service during a certain period of time T_{out}. Each of these pairs of values (AN_{out}, T_{out}) represents an element j of the group P_{out}.

4.4 Results

For these input parameters, MTBF = 6 months, MTTR varies from 6 hours to 14 days, we have measured the service unavailability percentage for both IP-TV distribution mechanisms. This MTBF value is evaluated as a worst case and also it can be considered as a typical value for an air network deployment. We have obtained the following graphics in *Figure 5* and *Figure 6*.

Figure 5. Service unavailability percentage with Multicast with 1+1 Protection

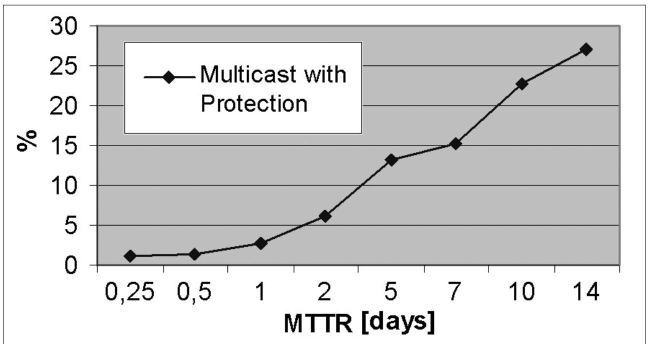
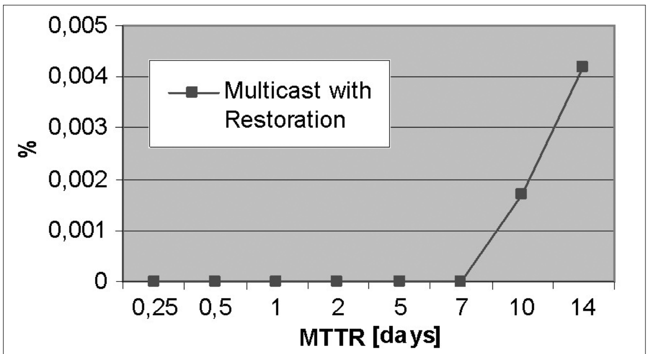


Figure 6. Service unavailability percentage with Multicast with Restoration



Thanks to the adaptability capacity of the restoration mechanism to the available network resources in case of multiple failure, the service unavailability time is zero or negligible for multicast with restoration (Fig. 6), for all the considered MTTR range. In the case of the 1+1 protection resilience mechanism, as we have only two possible multicast trees, the probability that a failure affects a branch of any of the two possible trees is many greater, increasing the access nodes that can not be reached (Fig. 5).

Summing up, according to our simulation results, only restoration-based solutions can achieve 99.999% figures for providing Broadcast TV services, regardless of the MTTR (mean time to repair) parameter.

4.5 Description of assumptions

Our analysis is based on the following assumptions:

1. There is enough bandwidth to carry the TV broadcast channel.
2. However, failures may happen at any link, with Poisson rate λ, which is equal to 1/MTBF and recovery rate μ, which is equal to 1/MTTR.
3. The possible transmission strategies are either multicast with restoration or protection.
4. The maximum number of failures is equal to four. Namely, recovery will happen immediately if four failures occur. This implies that the operator will make every effort not to have four failures whatsoever.

Based on these assumptions, we will evaluate:

1. Number of failures that lead to loss of connectivity in the multicast tree.
2. Distribution of the time to disconnection, namely, time it takes to loose connectivity in the multicast tree.

By multicast tree, we mean the (Hamiltonian) path in the network, such that every node is visited once. The goal is to assess the network performance in terms of fault tolerance of the multicast tree, in terms of MTBF and MTTR.

4.6 Number of failures to loose connectivity

For both the protection and restoration case we evaluate the number of failures that lead to loss of connectivity. Clearly, the chances of disconnection increase with increasing number of failures. Our aim is to evaluate the minimum number of failures are necessary for an eventual disconnection.

4.6.1 Multicast tree with restoration

Our preliminary results show that there are four possible link-disjoint paths originating at the head-end node to any other node *Figure 7*.

Being the node outdegree equal to four, a Hamiltonian path exists (Theorem 6.10 [5]). On the other hand, since the maximum number of link-disjoint paths is four then the minimum number of "cut-sets" that disconnects the subgraph is four (Menger's Theorem, [5]). As a result, the minimum number of failures is equal to four. However, the failure location is essential. Namely, four failures may no lead to loss of connectivity in the tree.

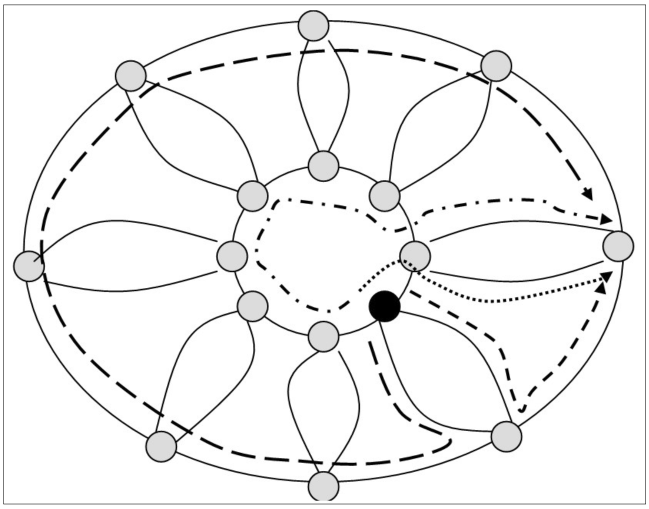


Figure 7. Link disjoint paths in a restoration scheme

4.6.2 Multicast tree with protection

In this case, there are two possible link-disjoint multicast trees, a primary and a backup one *(Figure 8* and *9)*.

Figure 8. Primary multicast tree

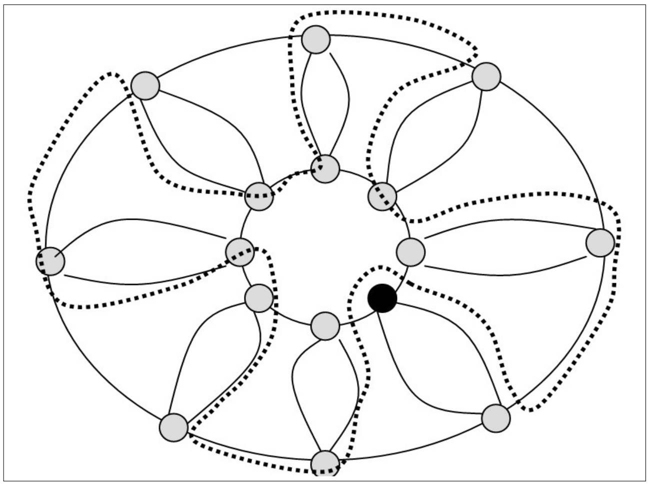
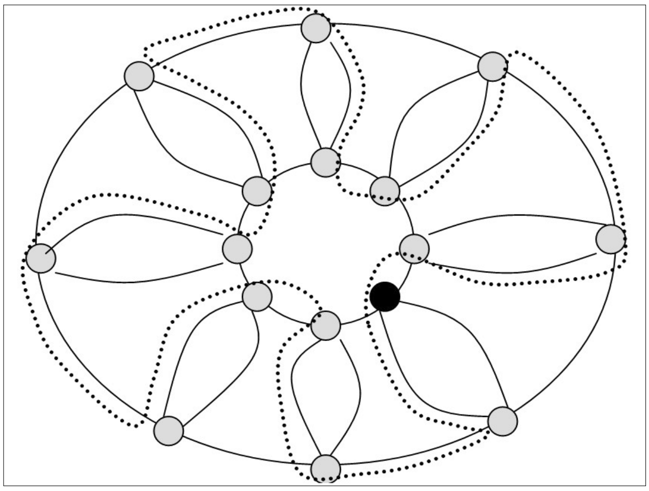


Figure 9. Backup multicast tree



Since the number of trees is equal to two the minimum number of failures is also equal to two, one per tree, in contrast to restoration where typically more than two failures can be "survived".

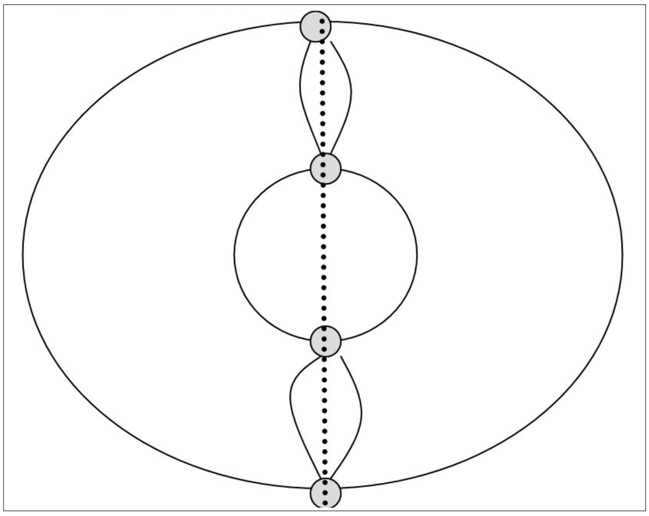
4.7 Distribution of the time to disconnection

In this section, we will characterize the time to disconnection, as a Continuous Time Markov Chain, with the hypothesis set forth in previous sections. First we will derive the set of states. Then, we will evaluate the hitting time to the state "disconnection of the multicast tree".

4.7.1 The case with four nodes

Let us examine the case with four nodes, in order to provide a lower bound for the time to disconnection *(Figure 10)*.

Figure 10. The case with 4 nodes



First note, that a minimum of four failures are necessary to produce a disconnection. Secondly, we note that the graph is symmetric. In what follows, we exploit the symmetry to derive the time to disconnection.

The following graph *(Figure 11)* is obtained by folding the previous graph along the axis of symmetry.

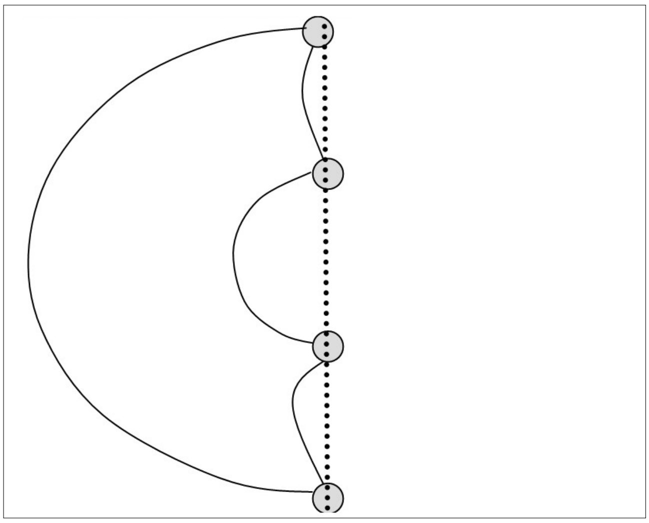


Figure 11. The folded graph with 4 nodes

Interestingly, let us consider the *reflected failures across the symmetry axis from right to left.* For example, a failure in the right outer edge can be reflected to the left outer edge as shown in the following figure *(Figure 12)*.

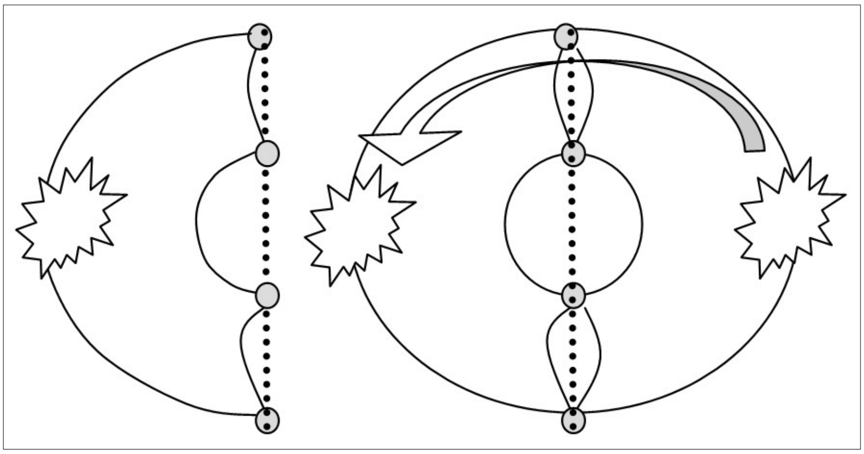


Figure 12. Reflected failures from left to right

Now, if we consider the folded graph only, it turns out that *disconnection happens if and only if there are four reflected failures, in two adjacent edges*.

4.7.2 Time to disconnection with four nodes

In this section we evaluate the time to disconnection, i. e. the distribution of the time elapsed until the multicast tree is disconnected. To do so, we consider the Continuous Time Markov Chain of the number of failures and *ratio of sample paths that lead to disconnection*.

The former can be easily derived using the M/M/4/4 model. The latter requires an explicit calculation of the number of scenarios that give raise to disconnection, with a grand total of four failures.

We distinguish the following cases:

- One reflected failure per link in four links: we only have one possibility.
- Two reflected failures per link, non-adjacent: we have two possibilities.
- Two reflected failures per link, adjacent: we have 4 possibilities.

Thus, the ratio of the number of scenarios leading to disconnection is 4/(4+2+1)=4/7=0.57.

On the other hand, the distribution function of the time to disconnection can be obtained from the M/M/4/4 as follows. First, we note that the infinitesimal generator is given by

$$P=\begin{pmatrix} -\lambda & \lambda & 0 & 0 & 0 \\ \mu & -\lambda-\mu & \lambda & 0 & 0 \\ 0 & 2\mu & -\lambda-2\mu & \lambda & 0 \\ 0 & 0 & 3\mu & -\lambda-3\mu & \lambda \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Let X be the time until four failures, let $M=e^{-Pt}$. Then $P(X\leq t)=M(1.5)$. Finally, note that the latter is the distribution function of the time elapsed until four failures occur, regardless of whether they bring disconnection or not. Consequently, this is a conservative analysis. *Figure 13* shows the results for MTTR=15 days and MTBF=60 days (worst case).

Our preliminary result shows that four failures will happen in a time interval of 160 days with probability 0.9975. Namely, we may expect four failures in a time interval of approximately five months. However, only 4/7 of them will lead to disconnection.

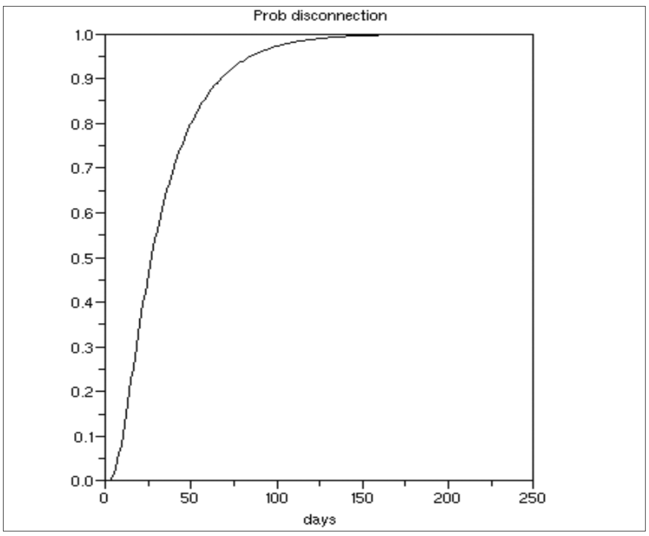
5. Multicast/Broadcast solutions for core networks

In core networks the video content is distributed (multi-casted) in bundles of tens to hundreds of programs to the metro networks. Depending on the resolution and encoding of certain program channels this requires a capacity from 100 Mbps to a few Gbps. Therefore, some multicasts having smaller bandwidth requirement can share a single wavelength path, while others that exhaust the capacity of a wavelength channel may even require multiple bundled wavelength channels.

We assume a two-layer network architecture, where the upper layer is an asynchronous time switched one, e.g., IP transport over Ethernet and/or MPLS while the lower layer is a circuit switched one, based on wavelength division multiplexing (typically DWDM eventually with OTN framing).

In such a two-layer architecture we assume multicasting capability at both layers. At the upper, IP/MPLS/ Ethernet layer multicast is supported, by sending the same packets to two or more outgoing ports. This increases the load of the backplane of the switch. At the lower, optical layer the multicast is done physically, i.e., the signal, as well as its power is divided among two or multiple outgoing ports. This approach requires splitters in the optical switches that, although not yet supported by many manufacturers, can be done by a simple and cheap splitter.

Figure 13. The folded graph with 4 nodes



We also assume grooming in our approach as follows. If there are two or more sub-lambda traffic streams that use the same path in a part of the network, they can be groomed together into a single wavelength channel by any grooming capable node as well as they can be separated (de-multiplexed) again by any other grooming capable node. This leads to much better resource utilization.

This two-layer network is represented as a single graph, with as many parallel edges between certain nodes as many wavelengths are supported over that link and using sub-graphs connecting these parallel edges in nodes to model different functionality including cross-connecting, grooming, etc. The model of this network is discussed in our earlier papers, including [6] and [7].

5.1 Methods for multicast routing

We have assumed that a single source (the root of the tree) supplies a few sinks, (destinations, leaves of the tree). This is a special Steiner tree, where the idea is to carry the information in a single exemplar (copy) as long as possible and to multiply at the farthest node to use as few capacity as possible for the whole multicast connection (tree). However, there are two constraints. Both upper (electronic) and lower (optical) layer multi-cast capabilities have breadth limitations, i.e., each node has limitation to how many output ports can it copy the same content. Furthermore, the depth of the tree, i.e., the largest source-destination distance has to be limited as well.

Here we have evaluated the following three multi-cast routing methods that we proposed earlier in [6] and [7]: ASP, MPH and ILP.

• **ASP**: Accumulative Shortest Path (Dijkstra)

This method is the fastest and simplest one however the results it provides are suboptimal. The root-to-leave demands are not routed at once simultaneously, but in a sequence one after the other using Dijkstra's algorithm. The idea is that the cost of elements (links in the wavelength graph) already used by a root-to-leaves demand of the same tree is set to zero, that means it can be used for free for all future root-to-leaves demands of the same tree. Of course the chosen sequence significantly influences the result.

• **MPH**: Minimal Path Heuristic

We have adapted [8] to our wavelength graph model. The idea is that we calculate the shortest path in our wavelength graph model between all leaves and between the leaves and the route. This results in a complete graph where the number of vertices equals to the number of leaves plus one for the root. In this simpler graph Prim's algorithm [9] is used to find the least-cost spanning tree. This minimum spanning tree is then traced back to the wavelength graph. Analogously to the ASP, where a new demand joins the tree, here while reconnecting the cut leaves the costs of all already used edges are set to zero.

• **ILP**: Integer Linear Programming

Since this method provides always the global optimum in terms of the objective function this was the reference method to compare other methods to. The time requirements for ILP were the largest among the three methods ranging from a few to a few hundred seconds in our case. The ILP formulation was proposed and explained in our earlier paper [6].

5.2 Methods for restoring multicast sessions

If a link or a node fails in the network it will affect all the multicast connections that use that element. However, if this element is just a leaf (a single user) its failure will affect only that user, however if an element close to the source (to the root of the tree) fails, than typically many leaves (end users) will be cut from the source. We propose methods for all the cases that reconnect the cut leaves (users) or whole branches (groups of users) to the healthy part of the tree or directly to the source.

Here we propose and discuss the different methods for restoring the trees upon failures. The four methods (ASP, ASP partial, ILP and ILP partial) for restoration that we propose here are based on methods for routing as follows.

• *ASP*

ASP restoration can be applied to any tree that was set up by any algorithm. Its idea is that if a link fails it can cut a single leaf or multiple (even all the) leaves from the root. We use here Dijkstra's algorithm to find a new path from each cut leave to the root, where the costs of already used links are set to zero as explained for the ASP routing.

• *ASP Partial*

ASP partial restoration is a kind of link restoration, i.e., if a branch of the tree is cut, then the whole branch as it is will be reconnected to the closest point of the tree.

• *ILP*

The whole tree is configured from scratch in optimal way. Instead of the original graph we use the graph without the elements that failed. This is the optimal new tree. However, it can be very different from the original one. This is a drawback, since many connections will have to be interrupted for reconfiguration purposes.

• *ILP Partial*

This is very similar to the ILP restoration approach with the difference, that the part of the tree that is not affected by the failure is kept, i.e., all unaffected links will have zero cost.

5.3 Simulation results

The simulations have been carried out on the COST 266 BT European reference network that consists of 28 nodes and 41 links. Each tree consisted of one 'root' and 5-27 'leaves' all randomly chosen with uniform distribution.

First, we have optimally configured the multi-cast trees using 'ASP', 'MPH' and 'ILP' as explained in Section 5.1 and shown as the leftmost triplet of bars in *Figure 14(a) – 14(f)*.

Then, we have simulated link failures one-by-one for all links used by the considered tree, and for each such failure scenario we have restored the tree using the four methods ‘ASP’, ‘ASP partial’, ‘ILP’, ‘ILP partial’ as explained in Section 5.2.

The evaluation criteria were as follows.

First we have evaluated the cost of the obtained tree as shown in *Figure 14(a)*. The failure-less tree was always the ‘cheapest’ particularly that obtained by ‘ILP’. After the failure, the ‘ILP’ has best restored the tree, regardless what was the initial tree set up method. For other restoration methods ‘MPH’ had roughly the same performance as ‘ILP’, while ‘ASP’ was the worst.

Second, the time required to calculate the multi-cast tree as well as to recalculate the restoration of the tree was evaluated as shown in *Figure 14(b).*

Here we see the drawback of the ‘ILP’ method for both routing and restoring the tree. However, it gives the global optimum in terms of its cost-based objective function. ‘ILP’ has the most significant time requirement, while ‘ASP’ and ‘ASP partial’ are the fastest.

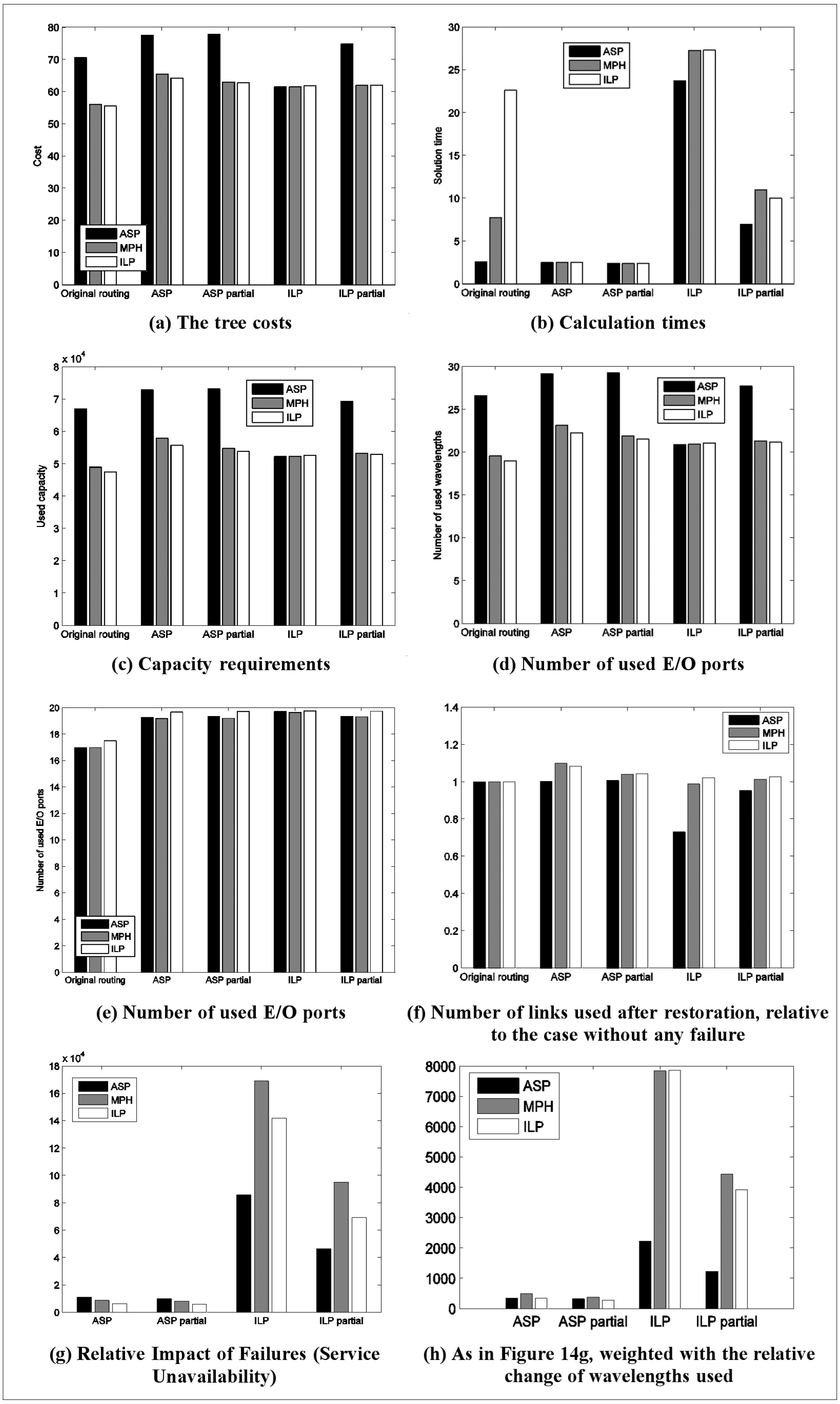
The amount of used capacity shown in *Figure 14(c)* has similar character to that of the cost (Figure 14(a)).

Figure 14(d) shows how many wavelengths are used by the different methods to set up and restore the trees. For both, ILP is followed by ASP. For restoration the partial methods have better performance than the simple full ASP.

Figure 14(e) shows how many E/O ports are required to perform multi-cast in the electronic (upper) layer. This is slightly related to the number of wavelengths used (Fig.14(d)).

If more wavelengths are used, slightly less E/O and O/E conversions are requested, since in some cases ‘E’ (electronic) multicasting can be substituted by the ‘O’ (optical) multicasting. Any failure will cause significant growth in using O/E and E/O ports.

In *Figure 14(f)* it is interesting to note that the size of the network relative to the failure-less case can be somewhat smaller, particularly for the ASP tree set-up with ILP tree-restoration! The explanation of this beha-



(a) The tree costs

(b) Calculation times

(c) Capacity requirements

(d) Number of used E/O ports

(e) Number of used E/O ports

(f) Number of links used after restoration, relative to the case without any failure

(g) Relative Impact of Failures (Service Unavailability)

(h) As in Figure 14g, weighted with the relative change of wavelengths used

Figure 14.
The results of simulating failures and recovering after them using four methods: ASP, ASP partial, ILP, ILP partial. The triple columns show the three methods ASP, MPH, ILP for setting up trees initially. The leftmost triplet of columns is the failureless reference case in Figure 14(a)–14(f).

Format	Resolution	Frames/sec	Declared bit-rate	Real average bit-rate
SDTV slow (Spiderman)	720x576	25	4.16 Mbps	~ 4,5 Mbps
SDTV fast (Polar express)	720x576	25	4.16 Mbps	~ 5 Mbps
HDTV (Magic of flight)	1920x1080	30	38.8104 Mbps	~ 9,6 Mbps

Table 4. The parameters of the 3 videos evaluated

viour is, that in the failure-less case ASP did not find a good tree, so relative to it ILP resets the whole tree from scratch, resulting a much better tree even if a link is unavailable due to its failure!

Finally, *Figure 14(g)* and *Figure 14(h)* show how the tree set-up method and the restoration strategy upon a failure impact the users. For this purpose we have defined two metrics, the Relative Impact (Figure 14(g)), and its variant (Figure 14(h)) weighted by the relative change of the number of wavelengths used, i.e., by the ratio of the number of wavelengths in the failure-less case to that in the case of failures.

We have defined the relative impact of failures as the average of the following products for all failure scenarios:
- The ratio of leaves cut from the root of the tree by the considered failure to all the leaves of the tree.
- The time of restoring the tree, i.e., calculating and setting up the new tree.
- The length of the link which failure is being considered (the longer the link is the more prone to failures is, i.e., has lower availability and will fail more often, therefore, it is taken with higher weight into the average).

In Figure 14(g) and Figure 14(h) it is to be noted that regardless of the tree set-up methods, the faster 'ASP' and 'ASP partial' methods should be used for restoration upon the failure, since although they provide slightly cheaper trees, their calculation times are not acceptable!

6. Quality of experience for video streaming in case of short interrupts caused by network failures

In this section we present our experimental study to evaluate how the different protection and restoration times upon failures affect the experienced quality of different video formats and contents.

In our experiments we have used the *ACR* (Absolute Category Rating) method [10]. The IETF has a similar framework, the MDI, Media Delivery Index [11], where the packet delay and packet loss are mostly considered.

However, since the failures happen very rarely, e.g., a few times per year that hardly affects the quality of the service in its classical sense we have condensed the failure events in the following way. We assume that there are exactly three failures of equal duration at random time instants with at least 10 seconds of difference between two failures within each 40-second video clip.

In all cases MPEG2 encoding has been used, with maximum packet size of 1310 bytes. The video frames have been carried by UDP over IP, around 400 packets per second, 16 packets per video frame on average. Three videos have been evaluated as shown in *Table 4*. The bit-rates of videos were analyzed by the Elecord Stream Application software. In all cases there were 15 evaluators. First, the video clip with no failures has been shown, then failures of duration of 30, 50, 200,

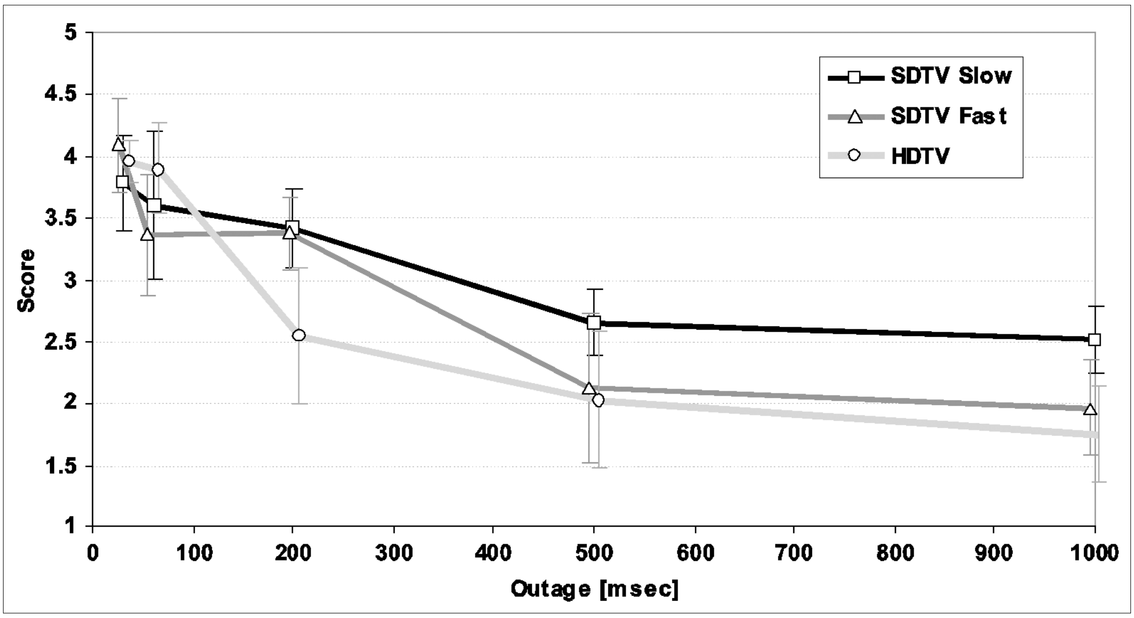


Figure 15. Mean opinion score with variance for the three videos for different durations of protection or restoration. Higher score means better quality.

500 and 1000 ms in random order and again these five failure durations in another random order with different failure instants.

Figure 15 shows the average and the variance of scores provided by the 15 evaluators. The effect of failures typically was that the motion stopped, the sound disappeared, and the picture was in part covered by squares off different sizes of colors mostly similar to the picture, however sometimes very different colors appeared as well.

It can be seen that slow motion SDTV is less affected by failures than the HDTV, while fast motion SDTV is between them for longer outages. For shorter outages the experienced quality is in general better. Two interesting properties can be noticed. First, that the quality of HDTV that was most seriously affected by longer outages is least impacted by shorter outages. Some short interrupts could be noticed, however, the squares were much smaller than for SDTV that explains the better experienced quality. Second, for shorter interrupts it does matter where the outages occur, i.e., what kinds of frames are lost.

In MPEG encoding each frame (GOP: Group of Pictures) consists of a series of frames, where the first one is the so-called I-frame which corresponds to a whole fixed image. It is followed by other frames that do not carry the whole fixed image, only differences relative to the picture carried in the I-frame. Therefore, if an I-frame is lost it is more critical than loosing only differences to this frame. Also, if there is sound or particularly speech when the outage happens it is more critical from the perspective of the user. This explains that for very short interrupts the subjective scoring of experienced quality can vary depending on the exact timing of failures.

Regarding the effects of failures onto the experienced quality of video streaming we can conclude, that interrupts of length from 30 to 1000 ms can be all noticed, they cause a minor disruption; however, considering that they happen a few times per year only, they are not critical at all. Although for SDTV interrupts over 100 ms, while for HDTV interrupts of over 50 ms can be annoying, if the service is restored within 50 ms the user will not loose any content that could hinder him understanding a sport event, a movie or news.

7. Conclusions

At present, IP-TV distribution in the metro and core networks is based on packet transport technologies such as IP/MPLS at level 3, including sub-50 ms 1+1 protection and restoration mechanisms (i.e., fast rerouting) and NG-SDH and DWDM technologies for transport at level 1.

At this point, it is important to mention the new existing alternatives for delivering IP-TV services, like PBT (T-MPLS or PBB-TE) at level 2, and OTN for level 1. However, these technologies only implement protection mechanisms for unicast traffic. In addition, restoration mechanisms are not available yet due to the lack of a distributed control plane in both technologies. So, in the future, these technologies can be planned to be used for the distribution of IP-TV service, as long as the standardization of OAM and resilience mechanisms for P2MP connections and the control plane definition for both technologies will be achieved.

In this paper we have analyzed the resilience requirements of IPTV based video streaming (multicast, broadcast) services, and also compared a wide range of resilience mechanisms and evaluated their capabilities and performance for both metro and core networks.

Firstly, in Section 4.4, it has been demonstrated by means of simulation that the utilization of restoration mechanism is the most appropriate one in order to support an acceptable quality of service when it is highly dependent on the service availability time, in a multiple failure scenario. This is the case of IP-TV broadcast service since a very high number of users would be affected by a total service cut. So, we recommend the application of restoration as the resilience mechanism in combination with multicast transport, since it provides total service availability in all the metro access nodes, regardless of the mean time to repair.

Finally, in Section 5.3 and 6, our results show that while there are few failures at a time the protection is fast enough not to affect the understandability and enjoyability of the video content. However, if there are multiple failures at a time, and instead of protection restoration has to be used that can last for seconds the users will not be satisfied with the quality. The probabilities of having such a failure pattern that will interrupt the streaming for more than half a second is very rare. In case of interrupts longer than a few tens of milliseconds the content should be cached and streamed again as soon as the network, or the cut branches of the tree have recovered.

Acknowledgement

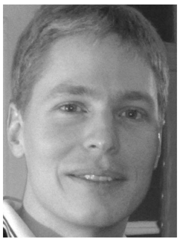
The authors are grateful to all colleagues in the FP6 IST IP NOBEL project, particularly within WP2.

Authors

TIBOR CINKLER has received M.Sc.(1994) and Ph.D. (1999) degrees from the Budapest University of Technology and Economics (BME), Hungary, where he is currently associate professor at the Department of Telecommunications and Media Informatics. His research interests focus on optimisation of routing, traffic engineering, design, configuration, dimensioning and resilience of IP, Ethernet, MPLS, ngSDH, OTN and particularly of heterogeneous GMPLS-controlled WDM-based multilayer networks. He is author of over 130 referreed scientific publications and of three patents. He has been involved in numerous related European and Hungarian projects including ACTS METON and DEMON; COST 266, 291, 293; IP NOBEL I and II and MUSE;NoE e-Photon/ONe andNoE e-Photon/ONe+; CELTICPROMISE; NKFP, GVOP, ETIK; and he is member of ONDM, DRCN, Broad-Nets, AccessNets, IEEE ICC and Globecom, EUNICE, CHINACOM, Networks, WynSys, ICTON, etc. Scientific and Programm Committees. He has been guest editor of a Feature Topic of the IEEE ComMag and reviewer for many journals.

PÉTER FODOR received his M.Sc. degree in electrical engineering from the Budapest University of Technology and Economics (BME), Budapest, Hungary, in 2001. He was consultant manager for a couple of years at the Department of Telecommunications and Media Informatics and a member of the High Speed Networks Laboratory. His research interests focus on network controlling, modeling and generalization. Currently he is working for Ericsson Hungary as Tool Expert.

PÉTER SOPRONI received his M.Sc. degree in Computer Science from the Budapest University of Technology and Economics (BUTE), Hungary, in 2008. He is currently a Ph.D. student at the Department of Telecommunication and Media Informatics in the same institute. He has participated in several research projects supported by the EU and the Hungarian government. His research interests include simulation, algorithmic optimization and planning of optical networks, as well as soft-computing especially bacterial algorithms. He has experience in .NET and Linux based software development.



MARCELL PERÉNYI received his M.Sc. degree in Computer Science from the Budapest University of Technology and Economics (BUTE), Hungary, in 2005. He awarded the Ph.D. title from the Doctoral School of Informatics of BUTE in 2009. He participated in several research projects, supported by the EU and the Hungarian government, at the Department of Telecommunication and Media Informatics. He is a member of IEEE and HTE. His research interests include simulation, algorithmic optimization and planning of optical networks, as well as identification and analysis of traffic of IP networks, especially P2P, VoIP and other multimedia applications. He has experience in planning and maintaining of database systems, web services and Microsoft infrastructures. Currently he is working for Ericsson Hungary as System Engineer.

JÁNOS TAPOLCAI received his M.Sc. (2000 in Technical Informatics), and Ph.D. (2005 in Computer Science) degrees in Technical Informatics from Budapest University of Technology and Economics (BME), Budapest, Hungary. Currently he is an associate professor at the High-Speed Networks Laboratory at the Department of Telecommunications and Media Informatics at BME. His research interests include applied mathematics, combinatorial optimization, linear programming, communication networks, routing and addressing, availability analysis and distributed computing. He has been involved in several related European and Canadian projects. He is an author of over 40 scientific publications, and is the recipient of the Best Paper Award in ICC'06.

ANDRÁS GULYÁS received M.Sc. and Ph.D. degree in Informatics at Budapest University of Technology and Economics, Budapest, Hungary in 2002 and 2008 respectively. Currently he is a research fellow at the Department of Telecommunications and Media Informatics. His research interests are complex and self-organizing networks, network calculus and traffic management.

GYULA SALLAI: see his cv after the Guest Editorial.

References

[1] NOBEL 2, "Deliverable 2.3: Report on Resilience Mechanism for NOBEL Solutions in Medium and Long Term Scenarios", IST NOBEL Phase 2, 2008. http://www.ist-nobel.org/Nobel2/imatges/D2[1].3_v1.1am.pdf

[2] NOBEL, "Deliverable D27: Final report on Traffic Engineering and resilience strategies for NOBEL solutions", IST NOBEL, 2005.

[3] OMNET++ object-oriented discrete event simulation system. URL reference: http://www.omnetpp.org, 1996.

[4] P. L'Ecuyer, R. Simard, E.J. Chen, W.D. Kelton, An Objected-Oriented Random-Number Package with Many Long Streams and Substreams. Operations Research, 50(6):1073–1075, 2002. Source code can be downloaded from: http://www.iro.umontreal.ca/~lecuyer/papers.html

[5] Alan Gibbons, "Algorithmic graph theory", Cambridge University Press, 1985.

[6] P. Soproni, M. Perényi, T. Cinkler, "Grooming-Enhanced Multicast in Multilayer Networks", ONDM2007, 11th Int. Conference on Optical Network Design and Modelling, May 29-31, 2007, Athens, Greece.

[7] M. Perényi, P. Soproni, T. Cinkler, D. Larrabeiti, "Regular Reconfiguration of Light-Trees in Multilayer Optical Networks", ONDM 2008, 12th Int. Conference on Optical Networking Design and Modeling, Vilanova i la Geltrú, Spain, March 12-14, 2008.

[8] M. Ali, J.S. Deogun, "Cost-effective implementation of multicasting in wavelength-routed networks", Journal of Lightwave Tech., Vol. 18, Issue 12., 2000.

[9] Thomas H. Cormen et al., "Introduction to Algorithms", 2nd Edition, MIT Press and McGraw-Hill, 2001. Section 23.2: "The algorithms of Kruskal and Prim", pp.567–574.

[10] [P910] "Subjective video quality assessment methods for multimedia applications", ITU-T Recommendation P.910, September 1999.

[11] J. Welch, J. Clark, "A Proposed Media Delivery Index (MDI)", RFC4445, April 2006.

ICC 2013 awarded to Budapest

Rolland Vida

vida@tmit.bme.hu

A year ago or so the Scientific Association for Infocommunications (HTE – www.hte.hu), namely the Hungarian Sister Society of the IEEE Communication Society (ComSoc), decided to bid for the hosting of the International Conference on Communications – ICC 2013, one of the flagship conferences of IEEE ComSoc. Following the usual rigorous site selection procedure, this May the decision was finally taken and the organization of IEEE ICC 2013 was awarded to the Hungarian bid team.

ICC is one of the largest and most prestigious IEEE ComSoc conferences. Each year on the order of 3000 papers are submitted. Typically between 1300 and 1700 attendees participate in an intensive 5-day technical program, with 18 parallel tracks, tutorials, workshops, industrial forums and panel discussions, complemented by industrial exhibitions. In 2010 the conference will be held in Cape Town, in 2011 in Kyoto, and in 2012 in Ottawa. In Budapest the conference will be co-hosted by three 5-star hotels on the banks of the river Danube (the Marriott, the Intercontinental, and the Sofitel), complemented by the Pesti Vigadó, a beautiful theater that will host the plenary sessions.

The general chair of the conference will be *Christopher Mattheisen,* Chairman-CEO of Magyar Telekom, while the executive chair will be *Prof. Lajos Hanzó*, IEEE Fellow and Distinguished Lecturer from University of Southampton, UK. The local arrangements chair is *Rolland Vida* from HTE and the Budapest University of Technology and Economics, while the finance chair is *Péter Nagy*, Managing Director of HTE.

Our local community is thrilled by the opportunity of bringing this prestigious event to Budapest. Naturally, we are aware that this decision marks the beginning of an intensive preparatory phase. However, we are excited by the prospect of organizing a memorable, enlightening and enjoyable ICC 2013 in Hungary, which we expect to have a long-term benefit for the local research community.

See you all in Budapest for ICC 2013!

Péter Nagy (HTE), Nándor Mátrai (Asszisztencia Congress Bureau), June Leach-Barnaby (IEEE ComSoc), Rolland Vida (HTE-BME) and Gayle Weisman (IEEE ComSoc) during the site visit in Budapest, on the top-level terrace of the Marriott hotel