SEVENTH FRAMEWORK PROGRAMME

Report on Y1 activities and new integration strategy

FP7-ICT-216863/DEIS-UNIBO/R/PU/D11.1

Project Number:	FP7-ICT-2007-1 216863
Project Title:	Building the Future Optical Network in Europe (BONE)
Contractual Date of Deliverable:	30/11/2008
Actual Date of Delivery:	10/12/2008
Workpackage contributing to the Deliverable:	WP11 : VCE on Network Technologies and Engineering
Nature of the Deliverable	Make your choice from: R
Dissemination level of Deliverable	Make your choice from: PU

Editors:	DEIS-UNIBO/F. Callegati, UAM/V.L. Alvarez, SSSUP/I. Cerutti, TUW/G. Franzl, UAM/J.A. Hernandez, UPC/Mirek Klinkowski, UPCT/P.P. Marino FUB/F.Matera, FT/H. Nakajima, UoP / Theofanis Orphanoudakis, FER/N. Skorin-Kapov, POLIMI/M. Tornatore,

Abstract:

This report focuses on the preliminary results of the joint activities started in the VCE on Network Technologies and Engineering. The report presents a summary of the organization of the work and of the results achieved in the executive summary. It then introduces the Joint Activities (JAs) as the most valuable technical contribution of the VCE, reporting a summary of the ongoing JAs.

Keyword list: optical networks, traffic engineering, OBS, OPS, protection, restoration, Carrier Grade Ethernet, technology survey.

Clarification:

Nature of the Deliverable

R	Report
P	Prototype
D	Demonstrator
O	Other

Dissemination level of Deliverable:

PU	Public
PP	Restricted to other programme participants (including the Commission Services)
RE	Restricted to a group specified by the consortium (including the Commission Services)
CO	Confidential, only for members of the consortium (including the Commission Services)

Disclaimer

*The information, documentation and figures available in this deliverable, is written by the BONE ("*Building the Future Optical Network in Europe*) – project consortium under EC co-financing contract FP7-ICT-216863 and does not necessarily reflect the views of the European Commission*

Table of Contents

1. Executive Summary

This document reports the activities of the VCE on Network Technologies and Engineering. A good level of integration in the research area of metro and core networks was achieved during the 6th FP with the e-Photon/ONe project. Unfortunately the split between the core and metro WPs raised some integration issue, since some activities, very similar in the concept, but referring to different application scenarios were running in parallel. This motivated the building of a single VCE related to "Network Technologies and Engineering" regardless of the application scenario.

The aim is to make a further step forward to achieve an integrated view of the problems of designing a high speed networks; focusing on traffic engineering, seamless QoS provisioning all over the network, multi-layer protection and fault management.

A very close link exists with the VCE on Services and Applications (WP12), that can be seen as a sort of "customer" of the solutions that are studied in WP11. For this reason the WP leaders made the effort to keep a close link in order to harmonize the activities of the two WPs.

Moreover in the BONE project several topics of interest for WP11 are addressed by:

TP Service aware optical network architectures

TP MPLS, GMPLS and routing

TP Edge-to-core adaptation for hybrid networks

TP Alternatives for multi-layer networking with cross-layer optimization

This is a problem and an opportunity at the same time:

- Problem: coordination to avoid replication of activities;
- Opportunity: provide valuable results obtained by means of a coordinated action that is well focused

The instrument chosen to correctly manage the issue was the organization of joint plenary meetings. During these meetings the WP leaders in particular and all participants in general had the opportunity to present and discuss the ongoing joint research activities. This was extremely useful to check for synergies, overlappings and positioning of JAs. As a consequence , after the first plenary meeting a big effort was done to move JAs from one WP to the other, merge JAs etc.

1.1 Participation

The WP counts the participation of 39 partners and 137 researchers subscribed to the WP on the BONE directory server. The involvement of partners is uneven, as could be expected when the number is so large. It ranges from pure auditory participation to the meetings to active involvement in one or several research activities.

The list of participating partners is:

UPC - BILKENT - TUW - SSSUP - UPCT - UAM - AIT - UCL - PoliTO - GET - USWAN - COM - UNIBO - CTTC - UNIMORE - Ericsson - KTH - UniRoma1 - UniRoma3 - BME - AGH - UoP - TID - UC3M - FT - UPVLC - UST-IKR - UDE - TUB - PoliMI - FUB - FER - UEssex - IT - IBBT - CORITEL - RACTI - HWDE – TELENOR

1.2 Publications

The various joint research activities achieved several interesting scientific results already, asproved by the many joint papers that have been accepted/submitted so far, as listed below.

- F. Matera (FUB), L. Rea (FUB), S. Pompei (FUB), A. Valenti (FUB), C. Zema (CORITEL), M. Settembre (CORITEL), "qualità of Service Control based on Virtual Private Network Services in a Wide Area Gigabit Ethernet Optical Test Bed", Fiber and Integrated Optics, Vol. 27, No. 4, pp. 301-306, Amsterdam, June 2008.
- K. Ramantas (RACTI), K. Vlakos (RACTI), O. G. de Dios (TID), C. Raffaelli (UNIBO), Window-based burst assembly scheme for TCP Traffic over OBS, OSA Journal of Optical Networking, Vol. 7, No. 5, pp. 487-495, May 2008.
- D. M. Forin (ISCOM), S. Di Bartolo (ISCOM), G. M. Tosi Beleffi (ISCOM), F. Curti (ISCOM), G. Cincotti (UniRoma3), A. Vecchi (UniRoma3), A. Raganà (UniRoma3), A. Teixeira (IT), Giga Ethernet free space passive optical networks, Fiber and Integrated Optics, Vol. 27, No. 4, pp. 229-236, to be published.

- J. Szigeti (BME), R. Romeral (UC3M), T. Cinkler (BME), D. Larrabeiti (UC3M), P-Cycle Protection in Multi-Domain Optical Network, Photonic Network Communications, to be published.
- J. Perello (UPC), S. Spadaro (UPC), I. Svinnset (TELENOR), E. Zouganeli (TELENOR), P. Cholda (AGH), J. Jajszczyk (AGH), K. Wajda (AGH), D. Verchere (Alcatel Lucent), R. Guenzinger (Alcatel Lucent), J. Fernández-Palacio (TID), O. González de Dios (TID), V. Chandrakumar (FT), The NOBEL approach to resilience in future transport networks, ECOC 2008 Proceedings, September 2008.
- N. Skorin-Kapov (FER), J. Chen (KTH), L. Wosinska (KTH), A tabu search algorithm for attack-aware lightpath routing, The Proc. of the10th International Conference on Transparent Optical Networks (ICTON 2008), pp. 42-45, Athens, Greece, June 2008.
- N. Skorin-Kapov (FER), O. Tonguz (CMU), N. Puech (GET), A novel optical supervisory plane model: The application of self-organizing structures, The Proc. of the10th International Conference on Transparent Optical Networks (ICTON 2008), pp. 10-13, Athens, Greece, June 2008.
- F. Matera (FUB), S. Pompei (FUB), A. Valenti (FUB), C. Zema (CORITEL), M. Settembre (CORITEL), Qualità of Service Control in Access Networks Based on Virtual Private LAN Services in a Wide Area Gigabit Ethernet Optical Test Bed, ICTON 2008, Athens June 22-26 2008, Vol. 4, No. 1, pp. 291-293, Athens, June 2008.

Another very important achievement in terms of publications is the book titled “Enabling Optical Internet with Advanced Network Technologies”. This is the joint effort of 8 scientists participating to WP11, who finalized during this year a work stemming from the e-Photon/One project. The book is organized into 6 chapters, each of them edited by a different person, covering the main issue optical networking. The focus is about an organized and homogeneous presentation of the most recent results achieved by the community. The book has been accepted for publication into the Computer Communications and Networks series published by Springer. The draft manuscript counts over 250 pages, acknowledges the contribution of the European Commission and is currently under production.

1.3 Mobility

Four mobility actions have been implemented during the first year for a total 9 man months.

The mobility actions involved mainly PhD students or young researchers that visited a partner to start or improve a research activity.

- “Network planning algorithms considering fault tolerance, security threats and periodic traffic patterns”, JiaJia Chen, PhD student at KTH, hosted by FER from 18/04/2008 to 05/05/2008
- “Benchmarking of network architectures for guaranteed service provisioning” Michele Savi, PhD student at UNIBO, hosted by UEssex from 28/05/2008 to 30/11/2008
- “Network planning algorithms considering fault tolerance, security threats and periodic traffic patterns”, Nina Skorin-Kapov, Assistant Professor at FER, hosted by KTH from 29/08/2008 to 14/09/2008
- “Network planning algorithms considering fault tolerance, security threats and periodic traffic patterns”, Nina Skorin-Kapov, Assistant Professor at FER, hosted by UPCT from 22/11/2008 to 07/12/

1.4 Meetings

The VCE organized 2 plenary meetings so far, jointly with WP12 WP21 WP22 WP24 WP26.

- Friday 27th June 2008, AIT, Athens Greece;
- Monday 20th October 2008, FUB, Rome Italy.

A third plenary meeting is planned for June next year. This will be a two days plenary meeting devoted to technical presentations of results achieved by the JAs, and will be very similar to a workshop in size and content.

1.5 Joint research activities

Joint Research Activities (JAs) are key to “integration” since they improve quality and quantity of research, strengthen the mutual knowledge of researchers that also learns to share and delegate problem to others. Therefore the main effort was devoted into starting new JAs that will provide valuable scientific results. At the same time care was taken into avoiding any overlapping with JAs in other WPs. For this reason some topic, that are indeed relevant for WP11, are not investigated in any JA. These topics are investigated in the framework of the TPs and WP11 will rely on the results produced there. Moreover it is worth reminding that, at least in the understanding of the WP leader according to his previous experience, a VCE does not have as major goal the steering of the direction of the research (more typical for a TP). The VCE should aim at favouring integration of the research by promoting the collaboration of the research groups about research topics that spontaneously stem from the community as a result of the research interests of the participants. In this way the research carried on in the VCE is a sort of "metric" of what the community believes is important in the field. Currently WP11 counts 13 JAs as reported in Table I, that can be classified into three main streams:

- Survey and economic perspective
- Network engineering
- Experiments on application

Activity #	Activity Title	Partners	Responsible
1	Common Architecture/Hierarchy for integration of OCS, OBS, and OPS in a common transport plane	TUW, DTU, AGH, GET/ENST, PoliTO, Uessex	Gerald Franzl
2	Comparative techno-economic network planning in OCS/OBS/OPS networks	UPCT, TID, UNIRM, FT, AGH	Pablo Pavon Marino
3	Experimental tests of Carrier Ethernet techniques	FUB, UNIROMA3, UNIROMA1, UNIBO, UNIMORE, KTH, CORITEL	Francesco Matera
4	Extension of the Flow-Aware Networking (FAN) architecture to the IP over WDM environment	GET/ENST, UaM, AGH	Victor Lopez Alvarez
5	Network planning algorithms considering fault tolerance, security threats and periodic traffic patterns	FER, KTH, UPCT, IT, AIT	Nina Skorin-Kapov
6	On exploring Admission Control Mechanisms for OBS networks	UaM, UC3M, UniMORE	José Alberto Hernández
7	PCE for Multi-domain Traffic Engineering	SSSUP, PoliMI	Filippo Cugini
8	Resilience analysis of double rings with dual attachments	UaM, TID	José Alberto Hernández
9	Traffic engineering and topology design in metro networks	SSSUP, GET/ENST	Isabella Cerruti
10	Survey on OBS routing methods for OBS networks	UPC, IT	Mirek Klinkowski
11	Optical Buffering Technologies Survey	FT, USWAN, TUW, UNIBO, UPC, KTH, Ericsson	Hisao Nakajima, Karin Ennser
12	Effects of outdated control information on routing in optical networks	POLIMI, CTTC	Achille Pattavina
13	Scalability of IP networks and routers	Telenor (FT, TID, UoP)	Evi Zouganeli

Table I. List of WP11 Joint activities

2. Technology Surveys

2.1 Optical Buffering Technologies Survey

2.1.1 Objectives

The lack of random access optical buffer (or memory) is the one important bottleneck for the implementation of meshed topology OBS networks within a practical operation scheme. The activity aims at surveying advanced optical buffering technologies, e.g. photonic band-gap devices, slow-light devices, nonlinear devices, which will potentially help the realisation of random access optical buffers in the future. The timeline may include not only the far future but also the near future (~2010). Practically the activity will consist in making a table of different technologies and establishing a forecast on each of them. The different steps leading to the goal are as follows:

- Establish a list of criteria (metrics, figure of merit) for the evaluation of performance, feasibility, integration, cost, etc.
- Testing the list with a few technologies for validation (M12).
- Collect information on different candidate technologies.

- Assess the candidate technologies according to the list, establish a forecast on each of them, and identify the most promising ones (M24).

The expected outcomes include:

- List of criteria for assessment,
- Forecast on optical buffering technologies,
- List of the most promising technologies.
- Papers, mobility actions.

2.1.2 Status of the JA

So far, the JA has produced the first draft of the list of criteria for assessment and the list of devices or technologies to be assessed. The first list is composed of 4 columns: Name of criterion, Category, Description and Applicability. Each criterion has to be explained why this criterion is needed, how it can be measured and to what it is applied with or without restriction. For instance, "Memory lifetime" belongs to "technological category" and describes "the duration of memory" that can be measured in "second", and is applied to "all memory types". The JA aims at filling in the list this way. The current list includes 14 criteria.

The second list has 3 columns: name, description and comment. A few items are currently on the list.

2.2 Survey on OBS routing methods for OBS networks

In 2008 the research groups of Universitat Politècnica de Catalunya (UPC, Spain) and Instituto de Telecomunicações (IT) have studied routing methods on optical burst switching (OBS) networks. The intention of the study was to put together a survey on the routing methods for OBS networks with the aim of providing a comparison of the performance of selected routing strategies.

There are a few motivations behind that work. Firstly, in the literature there is lack of a paper that would gather together and classify a variety of routing strategies that have been proposed for OBS networks. Secondly, to the best of authors' knowledge there is lack of papers comparing performance of routing methods in OBS. Finally, broad review of the literature would allow recognize still not-deeply addressed or open issues related to the routing problem.

In the first step of this joint activity, the focus was on the review of the literature. As a result, a detailed classification of routing methods was achieved. The next goal was to perform the evaluation of the performance of selected algorithms using UPC and IT simulation tools.

So far, a validation of the simulators in a unified network scenario has been completed and some initial results for selected multi-path algorithms have been achieved.

Current work concerns performing additional simulations and preparing a paper that would be submitted to a journal. The accomplishment of objectives is planned by the end of 2008.

A summary of the results obtained so far is presented in the following subsections. A list of references that have been used in this work is presented in Annex 1.

2.2.1 Routing methods in OBS networks

A great number of routing strategies have been proposed for OBS networks in the literature. In general, these strategies can be classified as alternative (deflection), multi-path (source-based), or single-path routing strategies.

A2.2.1.1 Alternative routing

A great part of research on routing problem in OBS networks concerns alternative (or deflection) routing. In alternative routing, when the burst contention occurs, a deflective mechanism reacts to it and re-routes a blocked burst from the primary to an alternative route. Deflection routing can be combined with other burst contention resolution mechanisms.

A great part of research on routing problem in OBS networks concerns alternative (or deflection) routing. In alternative routing, when the burst contention occurs, a deflective mechanism reacts to it and re-routes a blocked burst from the primary to an alternative route. Deflection routing can be combined with other burst contention resolution mechanisms.

Routing strategies considered for alternative routing in OBS networks can be either non-adaptive or adaptive.

In non-adaptive alternative routing both primary and alternative routing paths are fixed (static), and in most cases calculated with the Dijkstra algorithm. A number of alternative paths can be given from a node to the destination. Routing decision is taken in isolation, based only on local node congestion state information.

Adaptive alternative routing strategies apply a proactive calculation of alternative paths as well as their dynamic selection. The calculation of alternative paths is performed in an optimized way with the assistance of linear programming formulations. These methods need the information about network topology and traffic demands. In the case of dynamic alternative route selection some heuristics methods are used. In particular, either threshold-based or path-rank (priority) or probabilistic route selection techniques are applied. Dynamic route selection methods require the distribution of some link/node state information between respective nodes. Some of the alternative routing strategies support QoS provisioning by routing differentiation with respect to the quality class.

A2.2.1.2 Multi-path routing

Multi-path routing strategies in OBS networks aim the adaptive distribution of traffic over a number of routing paths in order to reduce network congestion. Although some proactive optimization techniques can be found, Dijkstra's shortest-path algorithm is still the most explored method for pre-calculation of routing paths. In most cases a small number of disjoint SPs with respect to the number of hops is calculated between each source-destination pair of nodes. In OBS multi-path routing, the selection of routing path is performed by the source node. The path selection can be either according to a given probability, like in the multi-path routing with traffic splitting, or according to the path congestion rank. Some authors propose centralized optimization methods for the calculation of the traffic splitting vector, whilst the others apply distributed heuristic methods. A ranking of the less-congested paths usually is obtained with some distributed heuristic algorithms. In both cases the distributed methods need updates about the network state information, which is usually disseminated from the intermediate/destination nodes to the source nodes. Such signalling messages can be either broadcasted or based on some events, like for instance, the burst dropping event.

A2.2.1.3 Single-path routing

Both non-adaptive (static) or adaptive (dynamic) strategies are considered for single-path routing in OBS networks. Static routing is usually based on Dijkstra's shortest path calculation with respect to the number of hops.

Adaptive single-path routing aims in burst congestion avoidance thanks to a proactive path calculation. The path calculation can be performed either in a centralized or in a distributed way. Centralized (or pre-planed) routing in OBS, in most cases, makes use of the optimization theory with (mixed) integer linear programming formulations. In each case it is supposed that a route computation unit has knowledge about network topology and (long-term) traffic demands. On the contrary, distributed routing uses some heuristics. Node state statistics are broadcasted, usually in a periodical manner, and used to calculate link weights (costs) in the respective nodes. Then a Dijkstra-like calculation is applied in order to find the lowest cost route. Some of the adaptive single-path routing strategies also support network resilience by the computation of backup paths.

2.2.2 Performance evaluation

The performance of routing algorithms is evaluated in a common network scenario in terms of overall burst blocking probability.

Figure 1 (left-hand side) presents the results obtained using IT and UPC simulation tools for shortest path routing (SPR). The results are consistent and correspond to analytical ones, which were achieved using a common OBS network burst blocking model.

Figure 1 (right-hand side) presents some comparative results achieved for SPR and two different multi-path routing algorithms, namely, the Load Balancing routing algorithm based on Lossless approximation (LBL) and the Multi-Path Routing algorithm based on a non-linear optimization model (MPR).

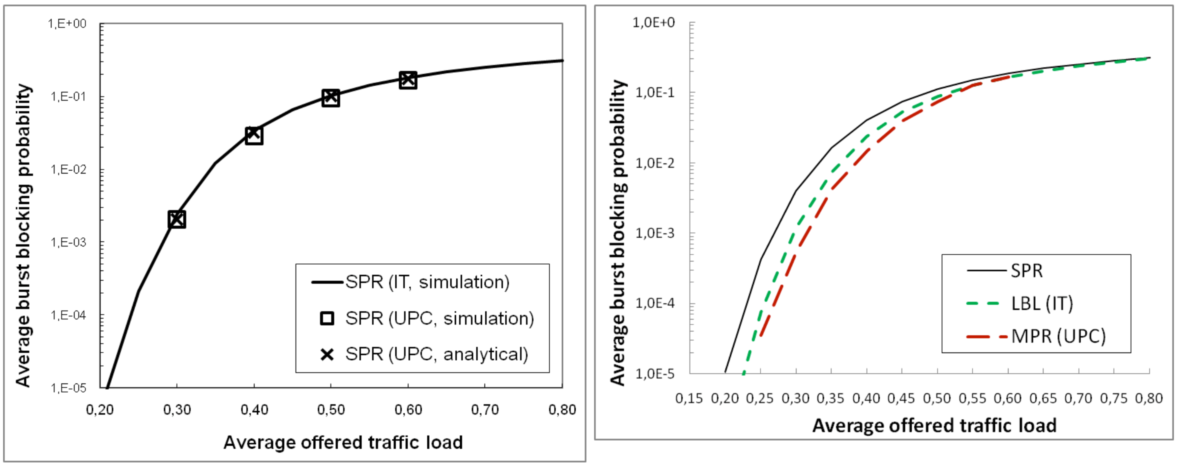


Figure 1. Evaluation of routing performance; (left-hand side) validation of IT and UPC simulation tools under shortest path routing (SPR); (right-hand figure) comparison of SPR and two multi-path routing algorithms (LBL, MPR).

3. Engineering transport networks

3.1 Common Architecture/Hierarchy for integration of OCS, OBS, and OPS in a common transport plane

3.1.1 Reference scenario and objectives

The amenity of light transported over fibres is not solely the propagation speed, today, in the century of global presence and its enabler broadband communication, it is the tremendous bandwidth it provides and that men can even increase the available bandwidth simply by adding more fibres. One drawback of communication based on light propagation is coarse granularity: an optical bandwidth of 25/50/100 GHz per channel can hardly be efficiently exploited by a single service. Consequently in current systems electronics are used to split the capacity of optical channels into many digital communication channels. Independent of the scheme used, this demands to convert all signals at every node to the electrical domain. Considering that a single fibre can efficiently transport digital signals with a total (summed) bit-rate of several Terra-bit, we recognize that a reasonably sized switch needs a tremendous amount of state-of-the-art electrical processing units to handle all the passing signals individually. Options to evade surplus electrical processing of trespassing signals at intermediate network units, thereby evading this scalability limiting and cost driving factor, is to be evaluated, even though potentially stirring state-of-the-art practice.

Optical Circuit Switching (OCS) is the state-of-the-art technology of today, the technology used for modern Telecom as well as Datacom networks to interconnect the respective electrical switches and routers. OCS comprises features that can not be achieved with any other technology, i.e., transparency and real-time support to transport Radio over Fibre (RoF) as well as any proprietary signal format. Consequently it likely will exist for long to serve such special applications. Optical Packet Switching (OPS) is the suggestion of the ITU for next generation networks. Converting the theoretically efficient packet switching schemes from electronics directly into optics fails as for today we have no optical capacitor that can store and release photons on demand, i.e., a random access memory (RAM) is missing. Still, OPS is a promising candidate once real optical RAM and processing based on photons are technically feasible and commercially available. Meanwhile optical burst switching (OBS) is assumed a realistic alternative that combines wavelength and time multiplexing in the optical domain without the need for excessive digital processing albeit the tremendous bandwidth provided by fibre cables. However, OBS is a step beyond digital communication networks as in principle, at least for today lacking transparent all-optical 3R regeneration and forward error correction, no discrete regeneration of signals at intermediate nodes is foreseen.

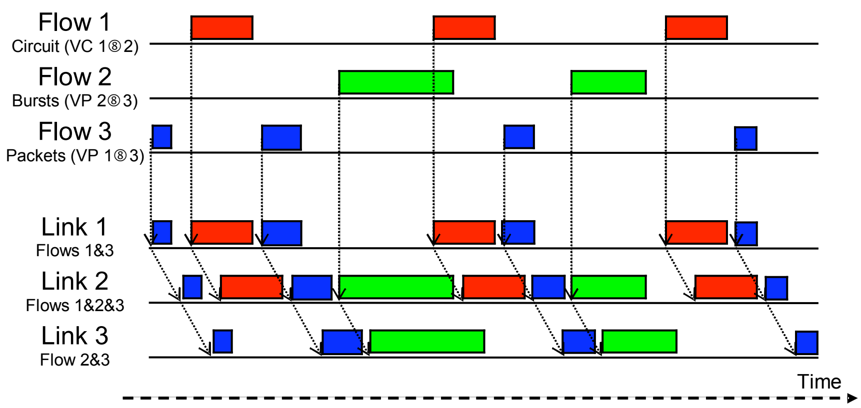


Figure 2. Multiplexing of different types of flows (time-slots, bursts, packets) sharing common transport capacities on connected links to perform transparent end-to-end transport

This Joint Activity (JA1) is studying and comparing options to integrate all three optical switching paradigms in a single transport plane as shown in Figure 2. In principle two approaches can be separated: Horizontal – providing all kinds of switching in the same network-layer without a hierarchy, and vertical – cascading the different switching options in a hierarchy of network-layers, as commonly done today, but integrating control and management of such multiple network-layer (mL3) architectures for better utilization of the available bandwidth. As typical there are a lot of possibilities in between, these shall not be excluded hereby, likely a compromise presents the optimum for ease of migration and implementation. The prime viewpoints considered shall be service's transport properties/characteristics and the so determined demand on Quality of Service (QoS) – application view, and the feasibly achievable utilization of resources – view of an operator that wants to provide appropriate QoS flavours.

3.1.2 Objectives

a) Several single partner studies on specific related issues preparing the ground for a sound proposal on unified transport plane providing potential for joint work, short term mobility and publications.

b) A survey on situation and options as currently realistically imaginable based on existing – at least in laboratory - hardware components.

c) A jointly written journal paper presenting one or more options for a transport plane serving all kinds of application demands - including performance analysis/forecast.

3.1.3 Status of development and summary of results

After a rather slow start some momentum was achieved in preparing first publications recently. Though only one was done jointly (authored by two partners) other were prepared based on feedback and discussions among at least some partners. Several study-topics were proposed respectively envisaged and likely some of these will be presented during 2009. Next find a summary of the topics that already are in a quite sound state.

A3.1.3.1 Foreseen service types respectively signalling demands:

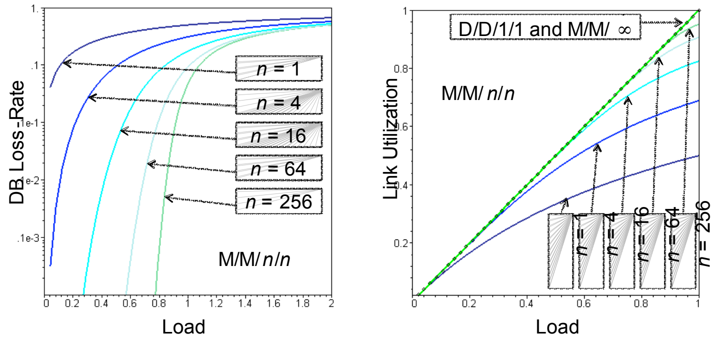
OBS being in between OCS and OPS is the first choice to achieve an integrated transport plane. While packets are in principle identical to short bursts, only the header needs to be separated and sent ahead of time (offset-time) to comply with OBS transport paradigm, integration of circuit-service demands extended OBS control options. However, a hierarchy of the three switching paradigms can effectively provide the same integration. In any case the following options are envisaged (in order of service quality/priority):

- Infinite duration and dedicated tear-down message to signal demand for an entire wavelength for an a priory unknown duration
- Request for repetitive switching of identical bursts with constant inter-arrival time (time-slot) for a CBR line service, again requiring the dedicated tear-down message
- Change request related to the request specified above in order to demand adjustments on timing and/or size for an adjustable virtual channel service

- Explicitly routed single burst/packet advertisement to provide virtual path service with/without acknowledgment
- Single burst advertisement with acknowledgment for assured delivery of single bursts – best used with scheduling window to improve routing success
- Just Enough Time (JET) one-way signalling for best effort and via offset-time differentiated single burst/packet switching

A3.1.3.2 OBS peculiarities and their impact on performance demands respectively bounds:

Considering OBS as core technology some principal behaviour peculiarities of OBS shall be considered. Firstly we note that aside potential fibre delay lines providing an array of fixed delays no real queuing of bursts is possible. This causes any OBS network to be a loss-system where in case two bursts demand the same resource at the same time one of the two is lost. Such networks are well studied, traditional telecom networks are of this particular type, and we can re-use the results, especially the Erlang equations, to study OBS performance issues.

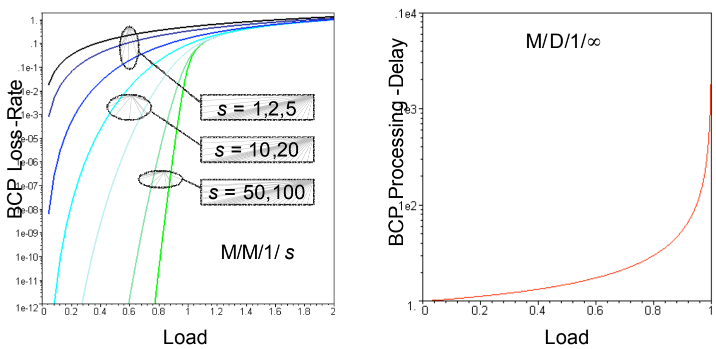


(a) (b)

Figure 3. Data-Burst loss-probability (a) and resource utilization (b) for different number of parallel channels over normailzed load (ρ=λ/nμ)

Figure 3 shows how parallel channels improve performance in terms of loss-rate and theoretically achievable utilization. Clearly visible is the strict demand for parallel channels.

Next we consider processing of the burst control packets (BCP). Noting that the offset-time once a burst is released into the network can not be changed on demand, we study how fast the processing needs to be to limit the loss of bursts caused by diminished offset-time to feasible amounts.



(a) (b)

Figure 4. Burst-Control-Packet loss-rate for different queue sizes (a) and processing-delay (b) per OBS core-node over normalized local load (ρ=λ/μ)

Figure 4 shows BCP loss-probability and the BCP delay caused by BCP processing in relation to load. To correctly calculate BCP-loss due to the actual delay bound determined by the actual offset-time we need to calculate the probability that a BCP needs to wait in the queue more than n times the sole processing time and assure that this is sufficiently small (e.g., <10-6). A problem is inter-arrival time distribution; if that is too bursty the M/D/1 and M/M/1 model underestimates the impact - for G/D/1 model the distribution needs to be correctly specified and equations derived. The proposed solution is to smooth the BCP traffic via shaping at the access (at edge/terminal).

A3.1.3.3 Integration options for OBS paradigm into GMPLS control scheme:

GMPLS proposes to be the final solution to all control issues. Developed for today's network hierarchies that might be true, integration of OBS technology is a different story. The GMPLS framework offers two generic functionalities: routing (finding an explicit path) and signaling (for explicit a priory resource reservation). GMPLS can cover the routing functionalities demanded to run an OBS network without any needed modifications. The remaining of the OBS control functionalities (e.g., BCP processing & on-time distributed resource provisioning) can not be covered by the GMPLS control plane. Thus there needs to be a separation between both, i.e. two control planes in parallel. Most literature that can be found on this topic focuses on the co-existence of two separate control planes as shown in the figure below.

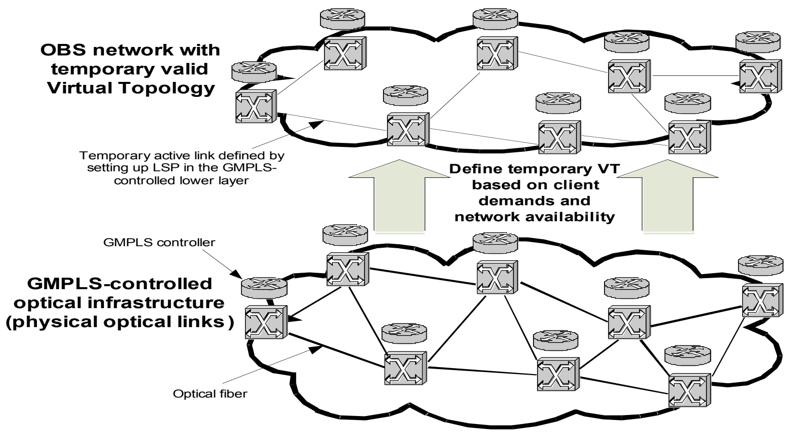


Figure 5. *GMPLS controlled optical infrastructure providing and managing the virtual topology to be used by OBS burst flows*

A general trend for handling the OBS network within the GMPLS framework is to label the BCP and consider both, BCPs and the strictly following bursts, as client-traffic in the data plane of the GMPLS controlled network on two absolutely parallel LSPs (one for BCPs, one for data) reserving an entire or even bundled wavelengths entirely to OBS use. According to this model, the BCPs are normal label-switched packets, i.e. it is assumed that a two-way signalling for end-to-end LSP establishment has been performed in advance.

A3.1.3.4 Flow transfer mode - a horizontal integration of switching paradigms:

The integration of switching paradigms via flow transfer mode (FTM) is based on OBS and extends what has been proposed as polymorphous agile transparent optical network (PATON) to integrate electrical last miles in an entirely burst switched network infrastructure. Flow Transfer Mode (FTM) enables different channel types to be multiplexed on the same transport resources without any restrictions causing decreased granularity respectively inaccessible idle resource shares. FTM exploits the entire potential for statistical multiplexing and therefore optimal resource utilization is not a priori rendered impossible.

Another intrinsic feature is the principal end-to-end transparency. Payload units are not restricted to certain formats and consequently FTM terminals can freely choose and even adopt modulation and digital format during operation without the need to inform the network as long as the end-to-end provided capacity, i.e., burst-size x burst-rate, is sufficient to transport the created data-stream. Not being designed entirely for the optical domain FTM enables the integration of electrical last miles as well as relay nodes performing O/E/O regeneration and burst re-scheduling to extinct constraints from limited optical reach. In the electrical domain FTM is a multi bit-rate TDM scheme and the nodes at the edge of the OBS infrastructure, as well as relay nodes, solely convert the electrical bursts 1:1 into optical bursts and vice versa without changing anything aside the carrier type, i.e., without performing burst assemblage. Bursts are created respectively disassembled at FTM terminals only.

Different FTM channels/services foreseen are:

Channel-/Service-Type	One-Way	Two-Way
Continuouse wavelength (line service)	possible	advised
Constant bit-rate (line service)	possible	advised
Adjustable bit-rate (virtual channel)	possible	advised
Assured variable bit-rate (virtual path)	not possible	mandatory
General variable bit-rate (virtual path)	best choice	possible
Assured single burst (possibly huge)	not possible	mandatory
Plain single burst/packet (rather small)	best choice	excessive

3.1.4 Open issues and topics planned for 2009:

1) Burst assembly strategies are intensely studied and many schemes recently published. A survey including considerations respectively comparison upon impact on performance shall be compiled to better distinguish different proposals respectively identify equivalent schemes.

2) A rather seldom addressed issue is burst-flow characteristics, i.e., inter-arrival time and burst-size distribution, and how smooth characteristics can be reached. From constant sized bursts to regularly timed bursts everything is possible and in case of CBR traffic deterministic burst-flows result. Shall timing and burst-size be upper/lower bound or shall it adapt to network state?

3) Network control and routing in OBS are still not completely defined. Most proposals assume a priory defined routes and shift routing problem to some not specified instance. Integrating two-way signalling recommends to include routing and, to optimise resource utilization, it is assumed that the actual scheduling time shall be determined from acknowledgement (scheduling window approach).

4) Techniques and new approaches to perform real 3R regeneration within an optically transparent environment allowing terminal defined carrier modulation to transport data-streams shall be envisaged. Clearly this is necessary to extend the optical reach to infinite (make OBS scalable in respect to distances bridged), however, for pioneering applications not mandatory.

5) As we assume that different channel types (services) shall coexist and share the same resources a study on optimum mixture of services respectively the impact of a dominating service-type on others shall be studied. In case of hierarchic architecture this 'reduces' to a rules-set to optimally configure the multiple network layers. Latter is covered by a different JA and therefore JA1 will concentrate on the first.

3.2 Comparative techno-economic network planning in OCS/OBS/OPS networks

3.2.1 Reference scenario and objectives

The aim of this JA is carrying out network planning studies in realistic scenarios, for optical technologies to be integrated in the short/medium or long term. This JA has evolved in two lines:

- A cost comparison of OPS/OBS/OCS networks in a metro ring network scenario
- Techno-economic evaluation of alternatives for the deployment of a photonic mesh

A3.2.1.1 A cost comparison of OPS/OBS/OCS networks in a metro ring network scenario

The aim of this JA is carrying out network planning studies in realistic scenarios (realistic network topologies, traffic demands) which compare from a cost perspective the merits of the OCS, OBS and OPS networks. At a first stage, this line focuses on a CAPEX/OPEX comparison in a metro ring network scenario, supported mainly by UPCT and TUW groups. The objetive is comparing the three alternatives: OPS/OBS/OCS, for carrying the same traffic in the same topology. The steps taken in this action have been:

1. Select a network topology in terms of number of nodes, distances among them, and number of wavelengths. Select a traffic model, from realistic measures.
2. Define a node model for OPS, for OBS and OCS. The CAPEX cost of each node is estimated attending to the weighted sum of the components of the model. OPEX cost of the model takes into account the power consumption.
3. Dimension the OPS, OBS and OCS network for similar measures of quality. Traffic loss objective with class differentiation is the disinguished performance measure of interest. Simulation and planning analysis are being used. Simulation is preferred for OPS and OBS analysys. Planning through MatPlanWDM tool (developed by UPCT) is preferred for OCS networks.
4. Calculate the CAPEX and OPEX of the network in the three alternatives, in order to compare, extract conclusions, and feedback for new comparisons.

A3.2.1.2 Techno-economic evaluation of alternatives for the deployment of a photonic mesh

TID group addresses a techno-economic comparison between DQPSK and DP-QPSK technical approaches for deploying a photonic meshed backbone at 100 Gbps. Given the different theoretical characteristics of both modulation technologies, the study includes a cost model to select the most suitable technology depending on the backbone size.

Several solutions have been proposed to achieve 100G transmission. The first and simplest way is the inverse multiplexing that allows 100G traffic to be split into multiple lower data rate carriers such as 10G. However, it has been proven that this solution makes network management more difficult and shows important problems at the receiver due to parallel transport synchronization complexity. On the other hand, 100G serial transmission must cope with challenging penalties arisen from the very restrictive optical impairments in such a high bitrate.

The best solution to deal with these restrictions is evolving towards more sophisticated modulation schemes with more than one bit per symbol so that the symbol rate becomes lower, relaxing optical impairments. DQPSK and DP-QPSK seem to be the most supported strategies by vendors.

Thanks to DP-QPSK lower symbol rate, optical reach is greater than DQPSK one; thus, all-optical links may be pretty longer in an optical mesh topology. This study analyses the cost feasibility for an all-optical mesh by comparing transparent optical links developed over either DP-QPSK or DQPSK. Furthermore, for achieving the same link length, DQPSK solution with electronic regeneration may be cost-effective because of the lower cost of DQPSK transponders in relation with DP-QPSK ones. As a result, a practical comparison between the two strategies is given regarding the length and the number of cascaded ROADMs that are traversed in a traffic path.

3.2.2 Status of development and summary of results

A3.2.2.1 A cost comparison of OPS/OBS/OCS networks in a metro ring network scenario

UPCT has developed the network-wide version of the oPASS simulator, suitable for evaluating OBS and OPS networks of heterogeneous topologies. The tests for the ring topologies are being carried out. UPCT has defined the architectures under test for both OPS and OBS nodes. They are:

- The OPS knock-out switch proposed in [1]. This is an example of OPS/OBS architecture able to emulate output buffering.
- The large-scale IBWR switch, scheduled by the PDBM scheduler [2]

TUW partner is working on the cost model for the nodes. Its contribution will be in estimating of power consumption for the configurations dimensioned, which has a direct influence on OPEX. TUW is estimating the power consumption of the networking equipment but also to take into account the power consumed by the room cooling equipment. Additionally, an UPS (uninterruptible power supply) unit with VRLA batteries for backup can be assumed, which ensures, as an example, one hour of continued operation in case of an interruption of supply from power grid. The housing of batteries increases the requirements on room size, which consequently contributes to an increased OPEX. These results can be combined with other parameters related to OPEX and CAPEX in order to determine the total cost of a network deployment within the planning phase.

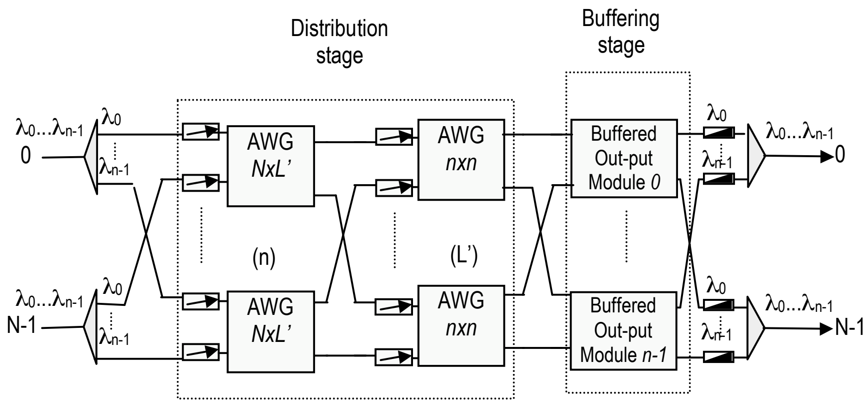


Figure 6. The Knock-out switch

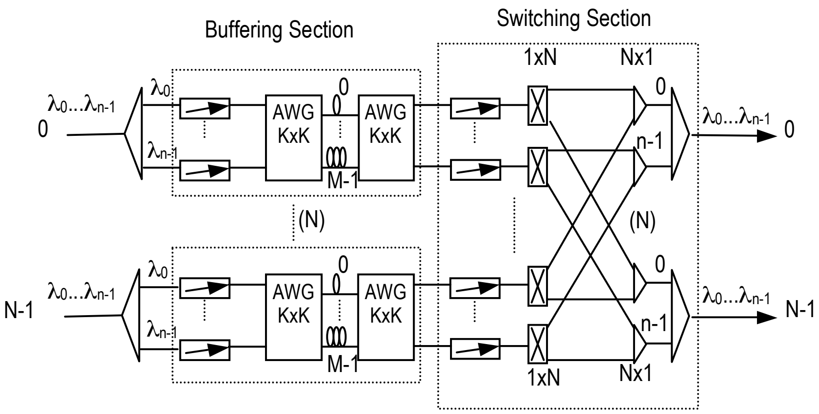


Figure 7. The IBWR switch

A3.2.2.2 Techno-economic evaluation of alternatives for the deployment of a photonic mesh

The first conclusions of the comparison have been published in [3]. In that paper, the study defines a ratio between the cost of the transponders. While DP-QPSK transponders duplicate DQPSK modulation components and need a polarization multiplexing device, DP-QPSK hardware works at half the frequency. Thus, it is assumed that cost of DP-QPSK transponder is twice the cost of DQPSK. This information is combined with the different lightpath length which can be obtained with both approaches. For a feasible interval up to 4900 Km, the optimal cost solution depends strongly on the lightpath distance. Further than 4900 km, DP-QPSK will always be the best solution in terms of cost. Furthermore, It is worthy of notice the fact that DQPSK consumes twice the OF capacity, which may be a significant advantage to fulfil forthcoming traffic demands. Besides, given that DQPSK modulation format involves a 100 GHz optical grid, new ROADMs with 32 x 100 GHz channels should be employed instead current 64 x 50 GHz.

Next figure summarizes cost ratio of DP-QPSK / DQPSK for a mean link distance ranging from 10 to 9000 Km. Above one, cost of DP-QPSK will be higher and DQPSK more cost-effective. Below one, DQPSK will be higher and DP-QPSK more cost-effective.

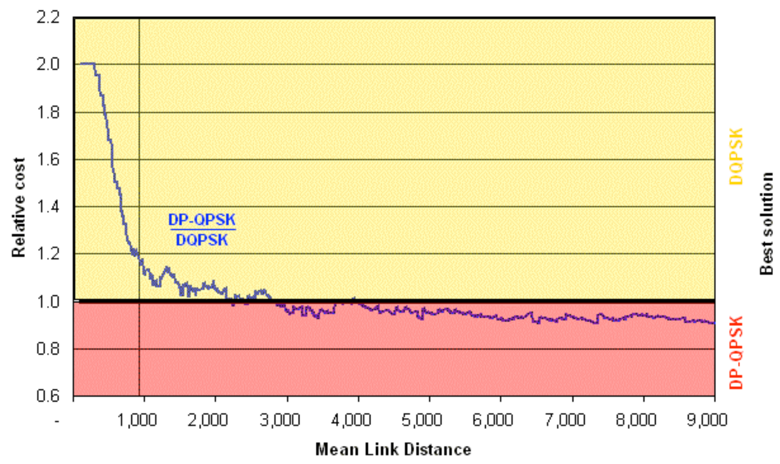


Figure 8. DP-QPSK / DQPSK cost ratio as a function of mean lightpath distance

In small and medium sized countries with low covered surface (link lengths up to 3000km), DQPSK solution is preferable due to its lower deployment cost. However, in cases of big countries with longer lightpaths (USA) or intercontinental backbones, DP-QPSK will be the best choice due to the longer reach and the saving in regenerators.

The study in [3] concludes that, in green-field deployments, the required investments for each option would strongly depend on the network size, so that while DQPSK would require lower investments in medium size scenarios, DP-QPSK would be a more cost effective option in the biggest networks. On the other hand, DP-QPSK channels can be transported over existing 50 GHz optical grids, therefore it would be a better option than DQPSK when 100Gbps channels are introduced over existing 10/40 Gbps photonic networks.

The next table summarizes the above conclusions.

	Mean Link Reach < 3000 Km	Mean Link Reach > 3000 Km
Greenfield photonic mesh at 100 Gbps	DQPSK	DP-QPSK
Migration from an existing 10/40 Gbps photonic mesh	DP-QPSK	DP-QPSK

3.2.3 References

[1] Pavon-Mariño P., Garcia-Haro J., Malgosa-Sanahuja J., Cerdan F., "A Performance Study of a Knock-Out Large-Scale Optical Packet Switching Architecture Under Scattered Wavelength Path Operational Mode", Proc. of 2003 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM 2003), Victoria (Canada), August 2003, pp. 478-481.

[2] P. Pavon-Marino, J. Garcia-Haro, A. Jajszczyk, "Parallel Desynchronized Block Matching: A Feasible Scheduling Algorithm for the Input-Buffered Wavelength-Routed Switch", Computer Networks, vol. 51, no. 15, pp. 4270-4283, October 2007.

[3] Santiago Andrés Azcoitia, Francisco Laura Mateo Quero, Óscar González de Dios, Juan Pedro Fernández-Palacios, "Techno-economic evaluation of alternatives for the deployment of a photonic mesh". NOC 2008 Conference.

3.3 Extension of the Flow-Aware Networking (FAN) architecture to the IP over WDM environment

3.3.1 Motivation and objectives

The research in Quality of Services (QoS) technologies started many years ago, but now the networking engineers start to understand the keys to QoS provisioning. The traditional traffic engineering is based on demand predictions, while the modern Traffic Engineering request for measurement-based approaches or state based approaches (automatic approaches).

Flow-Aware Networking (FAN) appears as a new approach to offer QoS. Flow-Aware Networking (FAN) observes and makes decisions at flow level instead of packets level. Flow-based traffic engineering uses more tractable mathematical models than packet-based traffic engineering (self-similar models). Flow-based traffic engineering takes more into account the probabilistic relation between demand-performance-capacity (represents almost human behaviour and is near to telephone systems).

Operator networks are migrating to an IP over WDM scenario where IP layer is directly connected to the optical layer. Moreover, the optical layer is becoming more and more intelligent and it is able to establish and tear down lightpaths automatically. This new topology creates not only new problems but also solutions, such as the utilization of the optical layer to solve congestion problems at IP layer. We call this solution Multilayer Flow-Aware Networking (MFAN).

This joint activity is focus on the enhancement of the congestion avoidance in Flow-Aware Networking technology. The main results of this activity are:

- Define the node architecture for multi-layer networks as well as the policies to deal with the incoming flows to the system.
- Define new congestion control mechanisms.

The solution proposed is shown in Figure 9. A MFAN node is composed by a FAN queue at the IP level and a module that is able to ask for extra resources to the optical layer when congestion is detected at the IP level. When FAN queue can not deal with the incoming flows, extra resources are applied to the optical layer.

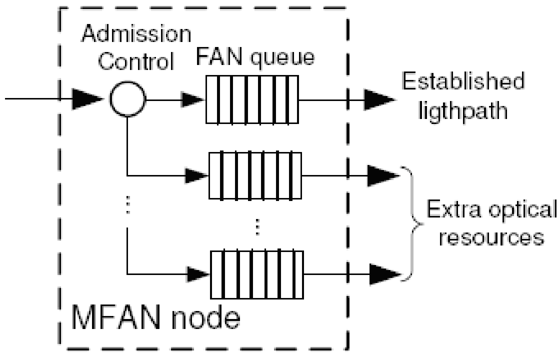
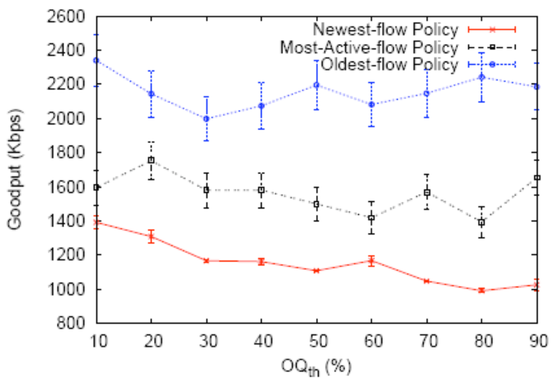


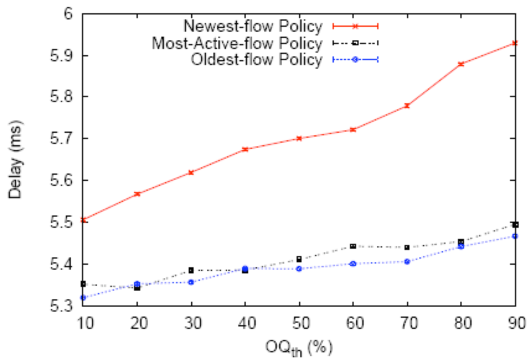
Figure 9. Multi-layer Flow-Aware Networking (MFAN) node architecture

3.3.2 Summary of results

The work is this topic started in previous NoE ePhoton/One+, where three policies were defined to deal with the incoming flows: Newest-flow, Most-Active and Oldest policies. Such policies were evaluated in a TCP dominated scenario. We discovered that Oldest policy was able to detect the biggest flows and this helps to improve the goodput in the optical domain and to decrease the delay as it is shown in Figure 10.



TCP flows goodput in the optical queue



Mean delay of the UDP packets in the optical queue

Figure 10. Performance in a TCP dominated scenario in a dumb-bell topology

In this joint activity, we have evaluated the performance of Multilayer Flow-Aware Networking (MFAN) in a UDP dominated scenario, where multimedia applications such as VoIP were the main traffic source. In previous work we assumed that the congestion was due to TCP traffic, but UDP traffic can congest the system. It is necessary to know the performance of the policies when this situation occurs (Figure 11).

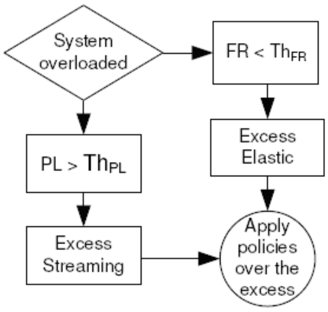
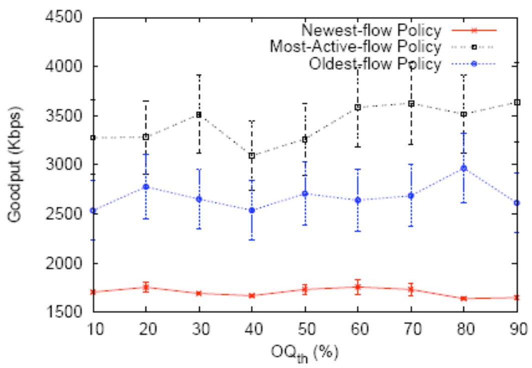
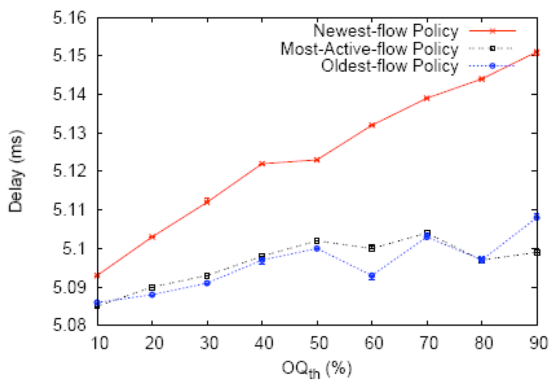


Figure 11. Optimal decision of the flows.

The performance of MFAN in a UDP dominated scenario is different than in the TCP scenario. Due to the nature of the UDP flows, Most-Active policy is able to extract the more suitable flows from the IP layer allowing a better performance in the optical queue. This can be seen in Figure 12. Delay is the similar to Oldest policy but the goodput achieved is higher.



TCP flows goodput in the optical queue



Mean delay of the UDP packets in the optical queue

Figure 12. Performance in a UDP dominated scenario in a dumb-bell topology

To check this behavior, the amount of TCP traffic was varied from an 80% to a 20% (Figure 13). This experiment concluded that the performance of Oldest policy is better when the amount of TCP traffic is more than a 50%, while Most-Active policy is more suitable when UDP traffic is congesting the IP level.

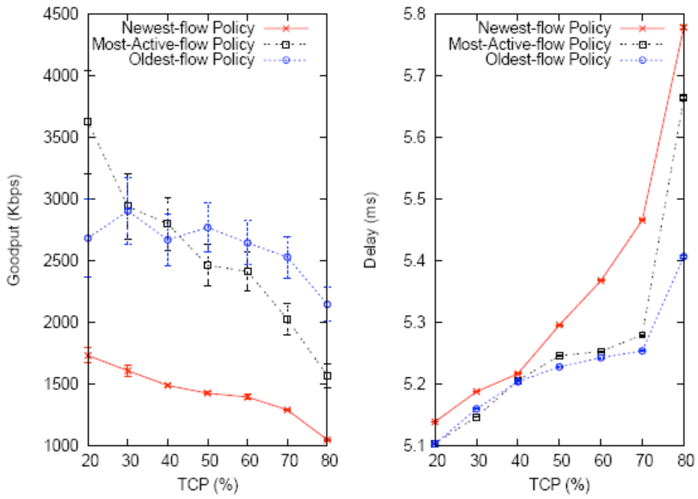


Figure 13. Policies performance evaluation when the percentage of TCP varies

This experiment opens the idea of a "Hybrid policy", which combines the benefits of Oldest and Most-Active policies, which should be the solution for all kind of traffic profiles.

3.3.3 Next steps

Once MFAN solution is studied in typical TCP and UDP scenarios, MFAN is going to be evaluated in Grid environments. GridFTP is extensively used to send big amount of traffic, so MFAN is going to be evaluated with this traffic profile.

Measurement-Based Admission Control (MBAC) block and scheduling process are proposed to be used in FAN environments, so we are working on the evaluation of this admission control mechanisms in Grid environments.

3.4 Network planning algorithms considering fault tolerance, security threats and periodic traffic patterns

3.4.1 Reference scenario and objectives

In Transparent Optical Networks (TONs), all-optical data connections, i.e. lightpaths, are established between pairs of nodes forming a virtual topology over the physical interconnections of optical fibers. Network planning in such networks are mainly focused on solving the Virtual Topology Design (VTD) problem which includes: (1) determining the set of lightpaths to be established, (2) finding physical routes and (3) assigning wavelengths to the selected lightpaths, and finally, (4) routing packet switched traffic over the established virtual topology. The 2nd and 3rd sub-problems of VTD are commonly referred to as the Routing and Wavelength Assignment (RWA) problem. Each lightpath must be routed over the physical topology and assigned a certain wavelength subject to the following constraints. The wavelength clash constraint states that lightpaths traversing a common physical fiber cannot be assigned the same wavelengths, while the wavelength continuity constraint states that the entire lightpath must be assigned the same wavelength. The RWA problem has been shown to be NP-complete and, thus, several heuristic algorithms have been proposed to solve it sub-optimally. Common objectives for the RWA problem are to minimize the number of wavelengths [1] used or minimize lightpath congestion [2]. We propose incorporating new objectives into the RWA problem, which is based on improving security vulnerabilities in TONs.

Namely, security and reliability issues are crucial in transparent optical networks due to the extremely large fiber throughput involved and the vulnerabilities associated with transparency. Some physical layer attacks can propagate due to transparency and cause system wide damage [3]. Furthermore, performance monitoring must be done in the optical domain using optical monitoring equipment, which is expensive and cannot be assumed at all nodes. Fast and successful reaction and restoration mechanisms in the presence of faults, as well as deliberate attacks, can prevent from loosing large amounts of critical data, which can cause severe service disruption. Dealing with these issues to assure service availability is especially difficult considering the limitations involved in optical performance monitoring [4].

We propose a novel approach to optical networks security, which is aimed at minimizing the potential damage caused by physical-layer attacks through careful RWA without the need for expensive specialized equipment. The objective is to arrange the set of lightpaths in such a way as to minimize the possible disruption, i.e. minimize the maximum number of lightpaths, which can be disrupted by a number of attack (or fault) scenarios. Consequently, if fewer lightpaths can be disrupted by a single attack (and we know ahead of time which connections could be attacked by which other connections), not only is network service disruption minimized, but detection and localization algorithms can be faster since they search for the source among fewer potential lightpaths. Furthermore, isolation of the attack, which is usually done in the reaction phase, is to some extent already partially done in the planning process.

In addition to Attack-Aware RWA, in the framework of this joint activity, we have considered the problem of Scheduled VTD. Namely, lightpaths in TONs can be static (set up semi-permanently), dynamic (connection requests arrive unexpectedly with random holding times), or scheduled (set up and torn down according to a predefined schedule). Due to the periodic nature of traffic, it may be possible to pre-define a schedule for establishing and tearing down lightpaths in order to more efficiently utilize network resources. Approaches to solving Scheduled RWA with the objective to minimize congestion have been proposed in [5] and [6]. However, to the best of our knowledge, determining the schedule of lightpaths on the basis of estimated periodic traffic, in conjunction with RWA and traffic routing, has not yet been studied. This means searching for the temporal evolution of virtual topologies and RWA schemes, which efficiently adapt to known traffic variations. Since this problem is NP-complete, heuristic algorithms are needed to help solve it sub-optimally. Besides developing efficient algorithms for scheduled VTD, the goal of this investigation is to assess the potential advantage of reconfigurable equipment, i.e. its benefit over the static case.

3.4.2 Status of development and summary of results

A3.4.2.1 Attack-Aware RWA

In the frame of this joint activity, we have investigated various attack and fault scenarios in transparent optical networks and proposed a new objective criterion for the Routing and Wavelength Assignment problem [7][8]. This objective, called the Lightpath Attack Radius (LAR) is aimed to arrange a given set of lightpaths in such a way as to minimize the maximum number of lightpaths, which can be disrupted by various attack scenarios. In addition to minimizing service disruption, detection and localization algorithms may be more scalable using such an approach since they search for an attacker among fewer potential lightpaths and locations.

We define the LAR as the maximum number of lightpaths any one lightpath is adjacent to, i.e., with which it shares at least one common physical link, including itself. Namely, this is the maximum number of lightpaths that will be disrupted in case of the potential physical-layer attacking scenarios, which include the following. If a high-powered jamming signal is injected on a legitimate lightpath, it can disrupt the lightpath it is injected on, as well as the other lightpaths with which it traverses common links as a result of gain competition in amplifiers and inter-channel crosstalk on fibers. Namely, an amplifier has a finite amount of gain available, which is divided among its incoming signals. Thus, by injecting a high-power jamming signal within the amplifier passband, an attacker can deprive other signals of power while increasing its own, allowing it to propagate through the network causing service degradation, or even service denial. In fibers, long distances and high-power signals can introduce nonlinearities causing crosstalk effects between channels on different wavelengths, called inter-channel crosstalk. We assume that only the original attacking signal can cause crosstalk effects, i.e. that attacked channels do not acquire attacking capabilities themselves. Furthermore, a low power tapping attack can be realized as follows. Namely, an attacker can request a legitimate data channel and then not send any data on it. As a result, the corresponding wavelength carries only leakage it picks up from neighboring channels via crosstalk. This weak leakage signal is then amplified at an amplifier and delivered directly into the hands of the attacker via his legitimate data channel. We currently only consider the set of attacks achieved by exploiting the vulnerabilities of fibers and amplifiers, but future work will also include the vulnerabilities associated with optical switches.

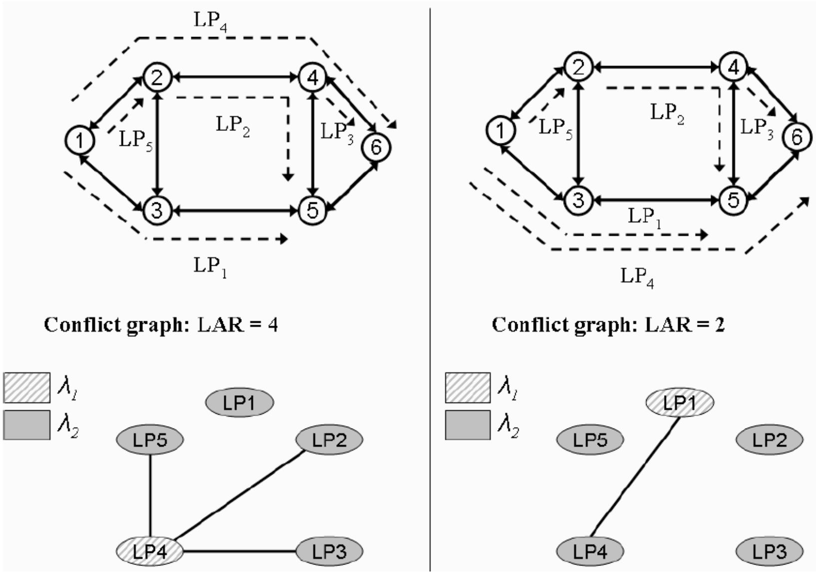
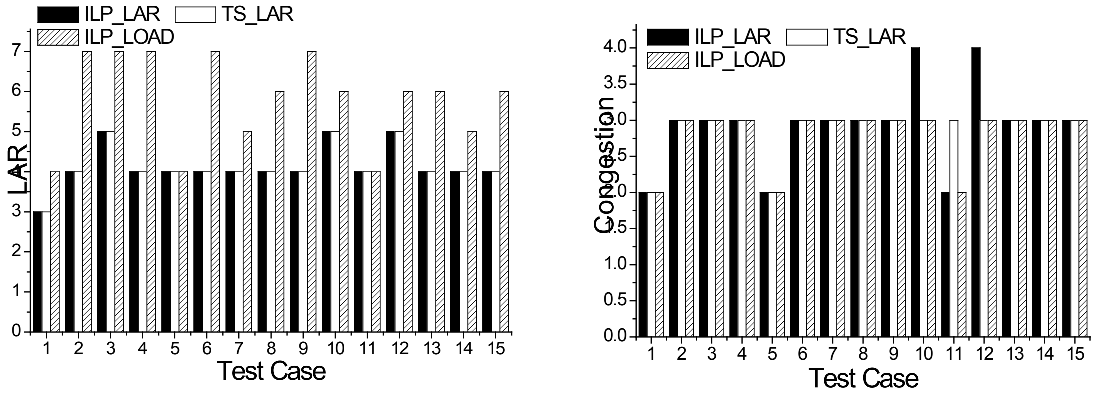


Figure 14. An example of two RWA schemes with the same number of wavelengths and the same congestion but with different values for the LAR.

Each routing solution obtained by solving the routing sub-problem of RWA corresponds to one so-called conflict graph, where nodes represent lightpaths and two nodes are connected if they are routed over a common link in the physical topology. The LAR is the maximum degree in such a graph, incremented by one, as shown in Figure 14 for a small example. Furthermore, solving the wavelength assignment sub-problem is equivalent to solving the graph coloring problem on the conflict graph. Since it is known that the maximum degree of a graph incremented by one is an upper bound on the number of colors needed to successfully color a graph, it follows that the LAR is also an upper bound on the number of wavelengths required. Consequently, if our objective is to minimize the LAR, we simultaneously minimize the worst case on the number of wavelengths needed for wavelength assignment. Furthermore, the LAR is also an upper bound on lightpath congestion (i.e., the maximum number of lightpaths routed over any one physical link). Note that one aspect of congestion is that it represents the maximum number of disrupted lightpaths in case of a fiber cut or other component malfunction along any link. Thus, minimizing congestion could potentially minimize the need for rerouting in case of such failures.

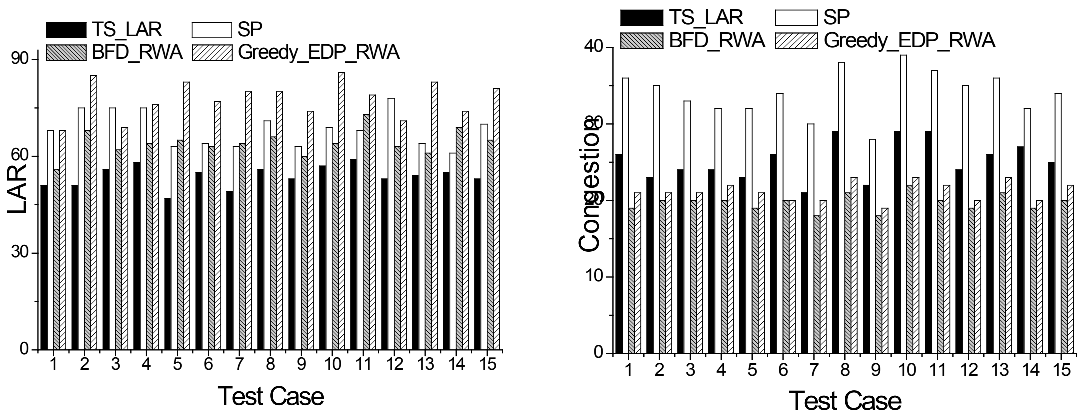
We formulated the routing sub-problem of the RWA problem with the objective to minimize the LAR as an exact Integer Linear Program (ILP), denoted as ILP_LAR. For comparison purposes, we consider a formulation which minimizes congestion from [2], denoted as ILP_LOAD. Since the problem is NP-complete, we propose a tabu search heuristic TS_LAR to help find suboptimal solutions for larger problems. Tabu search is a meta-heuristic which guides a simpler search procedure through the solution space and prevents it from getting stuck in local optima by introducing a memory structure called a tabu list. We ran the formulations and the proposed heuristic for a small 6-node network and compared the obtained routing solutions with respect to their LAR and congestion, shown in Figure 15(a) and (b), respectively. We can see that the ILP_LAR formulation achieves significantly better results for the LAR while obtaining optimal or near-optimal congestion in all cases. The TS_LAR algorithm obtained the optimal LAR in all cases, indicating its effectiveness. All three routing approaches were run in conjunction with a wavelength assignment algorithm from [9] in order to obtain complete RWA schemes. In the cases tested, the number of wavelengths was equal to the congestion and is, thus, omitted for the sake of brevity.



(a) (b)

Figure 15. The (a) LAR and (b) congestion obtained by the proposed ILP_LAR formulation, the proposed heuristic TS_LAR, and the ILP_LOAD formulation from [2] for the 6-node network.

Since the ILP formulations are complex and, thus, highly intractable, heuristic algorithms need to be run to find sub-optimal solutions for larger problems. We tested TS_LAR on the 30-node reference basic topology from the COST Action 266 project [10] and compared with Shortest Path routing (in conjunction with graph coloring for WA), and two greedy RWA algorithms, Greedy_EDP_RWA [10] and BFD_RWA [11], both aimed to minimize wavelengths. The results for the LAR and congestion (which closely corresponds to the number of wavelengths needed) are shown in Figure 16.(a) and (b), respectively. We can see that the proposed heuristic obtains superior solutions with respect to the LAR, but as a trade-off with an increase in congestion. This trade-off seems justified by the extremity of the vulnerabilities associated with optical networks security.



(a) (b)

Figure 16. The (a) LAR and (b) congestion obtained by TS_LAR, Shortest Path (SP) routing, Greedy_EDP_RWA [10], and BFD_RWA [1] for the 30-node network from [10].

Preliminary results were presented in [7] and [8], while new results are currently in preparation for a journal paper. Future work will include incorporating intra-channel crosstalk occurring in optical switches, and its propagation capabilities, into our model. Furthermore, the proposed formulation and heuristic approach will be extended to include attack-aware wavelength assignment. Careful power equalization placement in order to thwart jamming attacks and further reduce the LAR at a minimum cost will also be considered.

A3.4.2.2 Scheduled Virtual Topology Design (UPCT, FER)

We formulated the Scheduled Virtual Topology Design problem as a Mixed Integer Linear Program (MILP) with the objective to minimize the number of transmitter and receivers necessary to establish a scheduled set of lightpaths which can satisfy given estimated traffic demands [11]. The traffic demands are given as a set of 24 traffic matrices, each representing the estimated traffic for one hour of the day. The MILP formulation was implemented and tested in the MatPlanWDM, a MATLAB-based software developed at UPCT and publicly available at the MATLAB central web site. It is composed of an application kernel, a set of libraries of related algorithms and a graphical user interface whose general goal is to aid the implementation and evaluation of optimization algorithms for lightpath-based optical networks. Due to the high complexity of the problem, heuristic approaches are necessary to find good solution for large problems. We have developed a tabu search algorithm and greedy approaches for solving Scheduled VTD. Testing to assess their efficiency and comparison with the MILP formulation using MatPlanWDM is currently underway.

3.4.3 References

[1] N. Skorin-Kapov, “Routing and Wavelength Assignment in Optical Networks Using Bin Packing Based algorithms,” European Journal of Operational Research 177, 1167-1179 (2007).

[2] D. Banerjee and B. Mukherjee, “A Practical Approach for Routing and Wavelength Assignmnet in Large Wavelength-Routed Optical Newtorks”, IEEE Journal on Selected Areas in Communications 14, 903—908 (1996).

[3] C. Mas, I. Tomkos and O. K. Tonguz, Failure Location Algorithm for Transparent Optical Networks. IEEE Journal on Selected Areas in Communications 23, 1508-1519 (2005).

[4] Kilper, D.C et al., “Optical Performance Monitoring.”, Journal of Lightwave Technology Vol. 22(1), 2004, pp. 294—304.

[5] J. Kuri, N. Puech, M. Gagnaire, E. Dotaro, and R. Douville, “Routingand wavelength assignment of scheduled lightpath demands,” IEEE J. Select. Areas Commun., vol. 21, pp. 1231–1240, Oct. 2003.

[6] N. Skorin-Kapov, “Heuristic Algorithms for the Routing and Wavelength Assignment of Scheduled Lightpath Demands in Optical Networks,” IEEE IEEE J. Select. Areas Commun. Vol. 24, No. 8, pp.2-15, August 2006.

[7] N. Skorin-Kapov, J. Chen, L.Wosinska, “A tabu search algorithm for attack-aware lightpath routing” in the Proc. of ICTON’08, (Athens Institute of Technology, Athens 2008) pp.42-45.

[8] N. Skorin-Kapov, “MILP formulation for routing lightpaths for Attack-Protection in TONs”, Proceedings of NAEC'08, Riva del Garda, Italy, Sept. 2008, pp. 55-62.

[9] D. Kirovski and M. Potkonjak, “Efficient coloring of a large spectrum of graphs,” in Proc. 35th Conf. Design Automation, San Francisco, CA, June 1998, pp. 427–432.

[10] R. Inkret, Kuchar and B. Mikac, Advanced Infrastructure for Photonic Networks: Extended Final Report of COST Action 266", Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, pp. 19-21, 2003.

[11] P. Manohar, D. Manjunath and R. K. Shevgaonkar, "Routing and Wavelength Assignment in Optical Networks From Edge Disjoint Paths Algorithms", IEEE Comm. Letters, Vol. 6, No. 5, pp. 211-213, May 2002.

[12] N. Skorin-Kapov and P. Pavon-Marino, “Scheduling, Routing and Assigning Wavelengths to Lightpaths in Optical Networks”, Dubrovnik, Croatia, May 29-31, 2008.

3.5 On exploring Admission Control Mechanisms for OBS networks

3.5.1 Motivation and objectives

Typical OBS networks employ Just Enough Time (JET) signaling, a one-way reservation protocol whereby Burst Control Packets (BCPs) request resources at intermediate nodes but these are note confirmed back to the source. Hence, data bursts are transmitted without any guarantees, and it sometimes occurs that these are dropped at a certain hop in the source-destination path, hence wasting all those resources already consumed at previous hops. This effect is specially harmful if some connections are abusing of the global shared resources, violating their respective Service Level Agreements (SLAs), thus causing: (1) global performance degradation; and, (2) unfair service received by other connections.

In this joint activity, we have proposed and studied an admission control mechanism called “Random Packet Assembly Admission Control” (RPAAC) which moderates the two problems mentioned above. The mechanism monitors the network load status, detects which links are heavily loaded and decides which flows among the total traversing them require throughput decrease, on attempts to alleviate congestion and benefit other flows which are not abusing from the network resources. Such throughput decrease consists of preventive packet dropping during burst assembly at the ingress nodes of the OBS network, thus making no use of the network core resources. The numerical results show a substantial increase in the throughput experienced by well-behaved flows, and fundamental fairness achievement in the use of the shared optical resources.

A3.5.1.1 Description of RPAAC

Essentially, RPAAC proposes the architecture below, which requires the cooperation of two entities: The border nodes which regulate the traffic load offered to the network, via the burst-assembly process; and the core nodes which collect measurements about their status (network load and blocking probability). The second check whether or not the blocking probability observed at each network link is below a blocking threshold, which is designed not to be exceeded if all end-to-end traffic flows are meeting their respect SLAs. Thus, when a given link experiences excessive blocking, then it becomes clear that one or many flows traversing such link are violating their contracted SLAs, hence overloading the network, and possibly causing performance degradation to other flows which are indeed meeting their respective SLAs. At this point, RPAAC takes place to detect such misbehaving flows, and reduce their throughput (offered traffic) accordingly. Such throughput decrease is performed via the random packet selection probability per end-to-end flow, which is proportionally adjusted to the amount of excessive traffic observed in the congested link. Additionally, the random nature of the burst-assembly timer ensures that the traffic offered to the network follows a Poisson process, which helps for mathematical tractability of the problem.

In this light, for each misbehaving end-to-end flow fj, its admission control probability pAC is computed based on the measured link blocking probabilities and the flow's contracted SLAs. If the flow is find to behave correctly (its offered traffic does not exceed its SLA), then pAC =1 (all packets accepted to the burst assembly process); otherwise its pAC <1. Such pAC is computed as follows:

$$p_{AC} = 1 - \frac{I_{li} - I_{li}}{I_{fj}} \frac{I_{fj} - SLA_{fj}}{\sum_k (I_{fk} - SLA_{fk})}$$

The algorithm identifies the number of misbehaving flows fk over link li, and reduces their throughput proportionally with respect to what it is specified by their SLA. Such proportion is calculated as:

$$\frac{I_{fj} - SLA_{fj}}{\sum_k (I_{fk} - SLA_{fk})}$$

Which gives a value in the range [0,1]. Finally, such quantity is multiplied by:

$$\frac{I_{li} - I_{li}}{I_{fj}}$$

In order to give the percentage of traffic necessary to be dropped for flow fj.

A3.5.1.2 Experiments and results

The following presents a summary of the results obtained for RPAAC applied to the NSFnet topology (Figure 17). To evaluate the benefits arisen after applying RPAAC, we define the following metrics:

Network throughput:

$$I_{tot} = \sum_k I_{fk} (1 - B_{fk})$$

Throughput fairness, per flow and total:

$$f = \frac{\max_j \left[\frac{I_{fj}}{\min\{I_{fj}, SLA_{fj}\}} \right]}{\min_j \left[\frac{I_{fj}}{\min\{I_{fj}, SLA_{fj}\}} \right]}$$

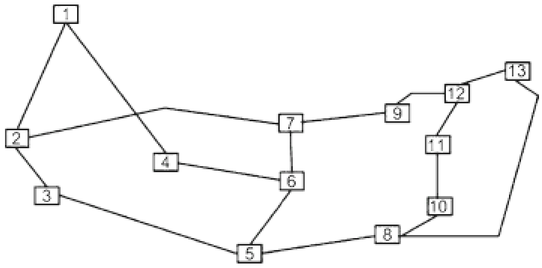


Figure 17. The NSFnet topology

According to this, the global throughput represents the total amount of traffic carried by the network, which shall be evaluated in the two cases of interest: With and without RPAAC. On the other hand, the fairness metric gives an idea of how such achievable throughput is distributed among the different end-to-end flows in the network. Ideally, flows whose carried traffic is close to their SLA experience fairness values close to one which is desired, whereas misbehaving flows exceeding their SLA will obviously experience fairness values greater than one. The closer the fairness index is to unity, the more fair sharing of network resources among flows.

In the NSFnet topology, it is assumed that each node in the network injects a random value of traffic to all other nodes in the network of average 0.15 Erlangs. All links are assumed to have eight wavelengths, and blocking probability threshold set to $B_{li}^{thres} = 0.026$. The simulation results are shown in the next figures. Figure 18 shows the blocking probability values obtained for all links before and after applying RPAAC, sorted from largest to smallest. As shown, after applying the Admission Control policy, no link exceeds the threshold blocking value defined above. It is also worth noticing that the blocking difference between the highest and the lowest blocking values has been significantly reduced after applying RPAAC, thus improving fairness. Indeed, after applying RPAAC, all links appear similarly loaded and no link appears as a bottleneck (much higher blocking probability than average) to the network.

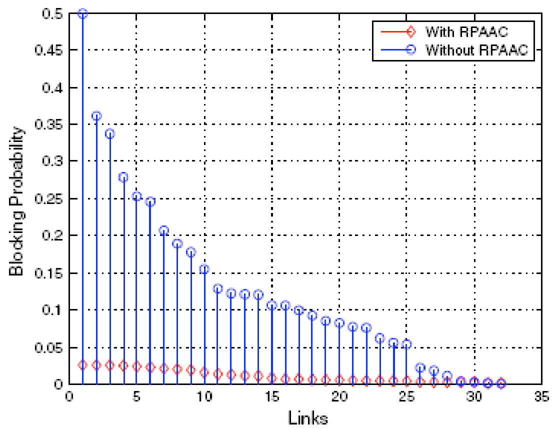


Figure 18. Blocking values of individual links before and after applying RPAAC

Basically, with RPAAC, well behaved flows experience a substantial throughput increase after applying RPAAC, since network resources, which were previously monopolized by misbehaving flows, are freed after applying the admission control policy. Additionally, RPAAC goes a bit further in the treatment of misbehaving flows since it penalizes much more those flows which are particularly abusing from the network. Hence, throughput penalization is proportional to the amount of resource abuse caused by each individual flow. This is shown in the next figure, where well-behaved flows are depicted on the left and misbehaved flows are on the right hand-side of the figure.

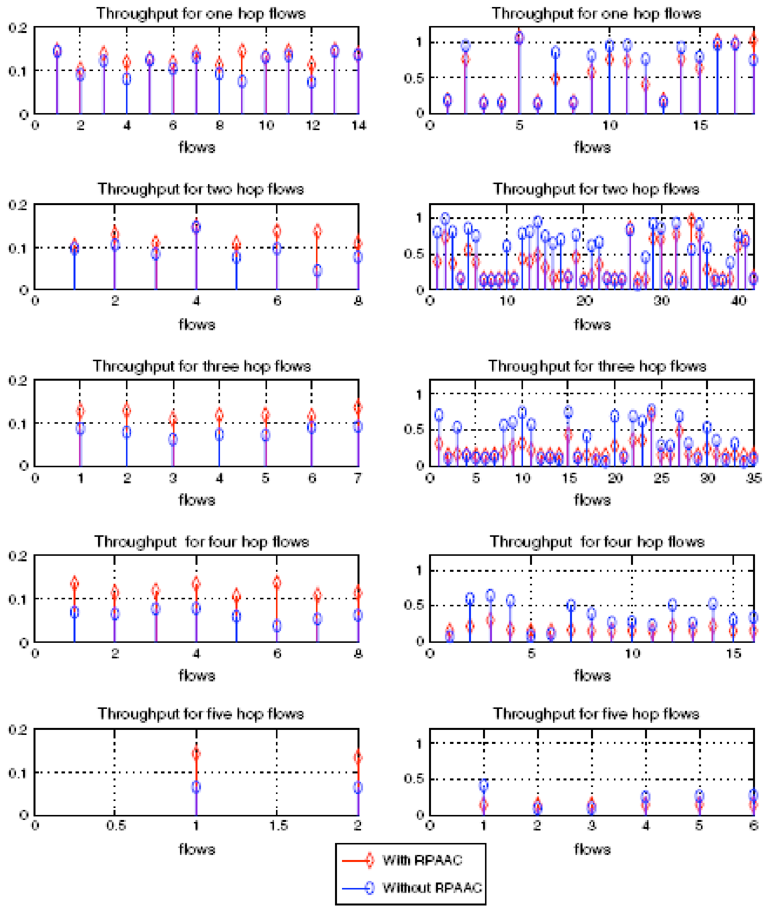


Figure 19. End-to-end throughput of well-behaved (left) and misbehaved (right) flows both before and after applying RPAAC

Finally, concerning global network fairness, this experiment shows a value of f=13.52 before RPAAC, which is improved to f=4.52 after RPAAC is applied to the network.

A3.5.1.3 Conclusions

In OBS networks, it sometimes occurs that data bursts consume network resources (in terms of wavelength timeslots) but never reach their destination due to congestion at an intermediate link. This effect is especially harmful as the network load increases, and it is sometimes observed that a few misbehaving end-to-end connections or flows are causing such behavior, which is suffered by the rest of flows. To alleviate this problem, this joint activity has proposed the Random Packet Assembly Admission Control mechanism (or RPAAC), a policy which is designed to drop packets at the border OBS nodes when the network load exceeds a certain threshold. Essentially, incoming packets at the optical network are either accepted or rejected at the burst assembler with a certain probability, which is adjusted by the admission control algorithm. The algorithm includes a monitoring stage whereby all network links are examined whether or not they exceed a load level, in order to calculate such amount of excess traffic per link, and it further detects which particular flows are causing such excess. This information is sent back to the border nodes which adjust the admission control probability for each individual flow. Additionally, the RPAAC follows a proportional penalization policy, that is, those flows which are abusing of the network resources suffer throughput decrease proportionally to the amount of traffic excess observed. This gives a more fair resource sharing between the connections existing in the network.

A3.5.1.4 Status of development and summary of results

This joint activity is already finished and its results are published online (not on text yet) with reference:

P. Reviriego, J. A. Hernández, J. Aracil: “Assembly admission control based on random packet selection at border nodes in Optical Burst-Switched networks”, Photonic Network Communications.

3.6 Traffic engineering and topology design in metro networks

3.6.1 Reference scenario and objectives

In this JA, an optical metropolitan area network (MAN) is considered. The optical MAN consists of WDM links and optical nodes with optical packet switching (OPS) capabilities. The network is based on connection-oriented paradigm, i.e., optical connections with sub-wavelength granularity can be established. Connection packets are all-optically switched in intermediate nodes.

The main aim of this JA is to optimize the design and the traffic engineering of the considered optical MAN. The innovative nature of the problem is the design of an OPS network while keeping into account the physical and technical limitations deriving from the physical components.

During this year, the JA has focused on an optical MAN with unidirectional ring topology. The unidirectional rings are time-slotted WDM rings with optical add-drop multiplexers nodes able to optically process the passing-through packets and adding (and dropping) packets from (to) the local ports. Different traffic connections can thus be aggregated on the same wavelength by a proper allocation of the slots to the traffic to be carried. Nodes that are transmitting (receiving) data on a given wavelength need to be provided with transmitters (receivers), that convert the electronic signal into optical signal (or vice-versa).

During the study, the following assumptions were made:

- no wavelength conversion is available, i.e., a traffic demand transmitted on a wavelength must be routed and received on the same wavelength;
- requested connections are static and known in advance;
- to account for the time-division multiple access (TDMA) nature of the ring, the traffic demands are expressed in terms of bit/s (in alternative, the bit rate of the traffic demands can be normalized to the wavelength capacity);
- given the TDMA nature of the ring, it is possible to split a traffic demand (from a given source to a given destination node) on two or more different wavelengths;
- different traffic demands can be aggregated on the same wavelength channel up to the available wavelength capacity;
- traffic demands can be added and dropped from the wavelengths, provided that (when adding) the above capacity constraint is ensured;
- the overall wavelength cost is proportional to the number of wavelength used;
- each node is provided with a number of tunable transmitters sufficient to support the requested traffic;
- nodes receiving traffic are provided with multi-wavelength receiver modules, i.e., a set of receivers each one operating on a distinct wavelength of a pre-defined wavelength-band.

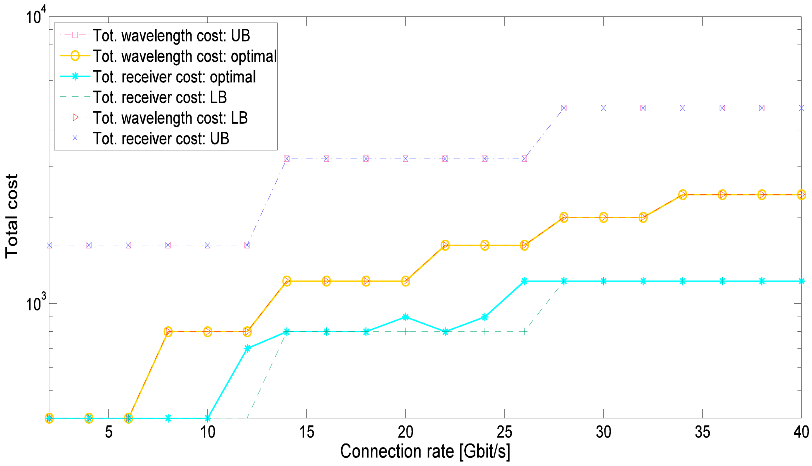
Under the above assumptions, the objective of this year activity was the definition and the formalization of the optimization problem, i.e., design of the unidirectional ring with the number of multi-wavelength receivers and wavelengths required to support the requested connections.

3.6.2 Status of development and summary of results

The work is the result of the collaboration between GET/TELECOM Bretagne and SSSUP. GET/TELECOM Bretagne defined the scenario and the optimization problem. SSSUP helped in the modeling of the optimization problem, with a mixed integer linear programming (MILP) formulation. Also, GET/TELECOM Bretagne is actively collecting the optimal results, obtained by running MILP solvers.

The definition of the problem required the identification of the network costs to be accounted and the identification of the sub-problem to be solved. The ring design problem consists in two different sub-problems that needs to be solved jointly: the aggregation of different connections to fit wavelength capacity and the wavelength assignment along with the aggregation of wavelengths into wavelength-bands.

The problem needs to be solved by taking into account the two contrasting costs to be minimized, i.e., wavelength cost and multi-wavelength receiver costs. The minimization of the wavelength cost leads to a minimization of the wavelength-bands (and thus a minimization of the used wavelengths). The minimization of the multi-wavelength receiver costs leads to a maximization of the wavelength-bands such that each node could operate on distinct wavelength-bands and, thus, be provided with the minimum number of multi-wavelength receivers. The optimal solution is thus a tradeoff between the two minimum cost designs.



A plot with preliminary results is presented here, but not yet published nor disclosed. The results are obtained for a 4-node unidirectional ring, in which a connection is required between each node pair, and the cost for a wavelength is equal to the cost for a receiver module. Receiver modules operate on 4 wavelengths. The figure shows the total cost for wavelengths and the total cost for receiver modules versus the connection rate. Optimal results are compared against upper bounds (UB) and lower bounds (LB) that can be easily computed in polynomial time. The curves of the optimal costs show that the overall cost for receiver modules is lower than the wavelength costs. This cost-saving is due to the multi-wavelength feature of the receivers. Also, the increase of the optimal cost for receiver modules is lower than the increase of the optimal wavelength cost. This indicates that, thanks to the traffic multiplexing performed by the optical nodes with OPS capabilities, it is possible to achieve cost-saving as fewer receivers are required in the network. Finally, the optimal costs match or are very close to the computed lower bounds.

Future work includes a thorough analysis of the collected results and derivation of useful considerations for the ring design. By exploiting such study, polynomial time algorithms that trade optimality for computational complexity could be proposed. Then the work could be extended to consider optical MAN with mesh topology, realized by connecting the studied unidirectional ring architecture.

3.7 Effects of outdated control information on routing in optical networks

In 2008 the research group of Politecnico di Milano (POLIMI) has investigated the effects of outdated control information in control-plane-enabled optical networks .

New signalling suites for a distributed control plane in WDM networks, such as GMPLS and ASON, allow to cope with the increasing variability of traffic patterns to be supported by operators by providing a means to dynamically setup and release connections. The dissemination of link-state information (usually provided by routing protocols such as OSPF-TE) is essential in this kind of control-plane enabled networks: in particular, this information has to be continuously updated to allow routing algorithms to efficiently carry on the path computation. This information is then collected by each node in a repository termed as the Traffic Engineering Database (TED) [1].

The information that should be distributed to populate TED and build an accurate net- work image depends on various factors, mainly on which protection technique is applied and if wavelength conversion is enabled. E.g., while the standard distributed information (e.g., link state advertisement in OSPF-TE) is the link-state and the free capacity available on each link, in a not-wavelength-convertible case, the single channel status has also to be distributed if the wavelength continuity constraint needs to be satisfied.

The problem of inaccurate routing due to control-plane delays in next generation optical networks has been receiving increasing research attention. Past research has focused mainly on inaccurate routing related to the periodic update of TEDs [2][3][4]. Our investigation introduces innovative elements with respect to the previous body of research: first, to the best of our knowledge, this is the first work covering in its analysis the peculiar effect (a) on protection (especially, backup sharing in shared protection) and (b) comparing the cases with and without wavelength conversion. Second, most of the previous works were focused on possible countermeasures to be included in the routing algorithm to soothe the effect of outdated information, focusing on a small range of values. Here we explore a wide range of values of control delays, analyzing the deriving performance loss.

In this joint activity, our generic delay model is based on the following assumptions:

- Constant Control Delay
- Negligible Set-Up Time
- Identical network vision among all nodes

Our simulative study is aimed at covering a large set of scenarios, exploring three different dimensions: (i) the value of control delay, (ii) the routing strategy adopted and (iii) the conversion capabilities.

In Table 1 we summarize the routing algorithms that we have applied to solve these six problems: unprotected (), dedicated path protected and shared path protected routing with and without wavelength conversion. In the following the wavelength convertible case will be referred as Virtual Wavelength Path (VWP), as opposite to the not wavelength convertible case referred to as Wavelength Path (WP).

As mentioned before, TEDs in source nodes are not always updated and routing may be carried on according to an outdated image of the network. The routing process can be divided in two separated phases: first, the OCP evaluates a possible routing path according to the network image stored by its TED; second, protocols such as RSVP-TE carry on the actual reservation of resources by signalling the resources along the path that a connection is being established.

So, for sake of terseness, in the following we will refer to the (possibly) outdated representation of the network in the TEDs as Network Image (NI) and to the actual state of the network as Present Network (PN), while τ is the value of control delay.

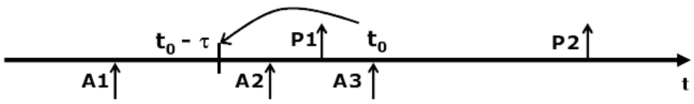


Figure 20. Network Image and Present Network.

If NI and PN coincide, e.g. when τ is so small that no changes in the network state occurs before TEDs are updated, blocking of a connection is due only to capacity lack. When introducing relevant control delays, new contributions to blocking can be identified:

- • RC - Resource Conflict: a connection is routed according to the NI along a path that is no more available on PN, because in the meanwhile those resources have been taken. The connection is blocked even if there are chances that alternative free paths were available.
- • FS - False Saturation Conflict: a connection can not be routed according to the NI due to capacity lack (no routes are available between source and destination nodes). The connection is blocked even if, in the meanwhile, on the PN some resources have been released and at least a path has been made available.
- • SR- Sub-optimal Routing: a connection is routed along the best route according to the NI state. The connection is not blocked but in the meanwhile, on the PN, due to some resources release/occupation, conditions are changed and the new best route

Figure 20 shows a simple example that clarifies the previous classification: at t_0, the arrival instant of connection A3, connection A_1 has been already released, while connection A_2 is still in the network. This is the PN state,which represent the actual network occupation. The path computation for connection A_3, due to the outdated network image at the source node, will be executed based on the NI status, i.e. the status of the network at instant $t-\tau$: note that, according to Fig. 1, during the computation of a path for A_3 connection A1 will be considered still in the network (causing possible false saturation conflicts), while the connection A_2 will be ignored (causing resource conflicts).

Applying this very general control-delay representation, we are able to provide wide-range simulative study to quantify the effect of signalling on routing performance, mainly by using the blocking probability (P_b) metric: since blocking stems from different causes, the influence of the various P_b components are discussed.

In Figure 21.(a), P_b curves are plotted for 200 and 240 arrivals per second (correspondent to a network load of 0.446 and 0.55, respectively) in the UN case. These curves can be described by three well distinct phases:

- • I phase - Not influential delay: in this first phase P_b is stationary and it is not influenced by the delay. Essentially performance are equivalent to the ideal (zero-delay) case.

- • II phase - Linear increase of P_b: delays in network-state reception start affecting the quality of source-routing, causing a significant and linear increase on P_b
- • III phase - P_b saturation: P_b is no more affected by a further increase of the network-state-information delay, since the network image in the source-routing node is totally uncorrelated to the actual network state. The situation is equivalent to route connections with no information about network state.

In Figure 21.(b) results are shown in the WP case. P_b values are higher than in VWP case: while in the first phase this is a well-known effect of wavelength continuity constraint, in the saturation phase we should also consider that an higher P_b is achieved because the routing algorithms for VWP cases are more tolerant to outdated information in the VWP case than the WP case.

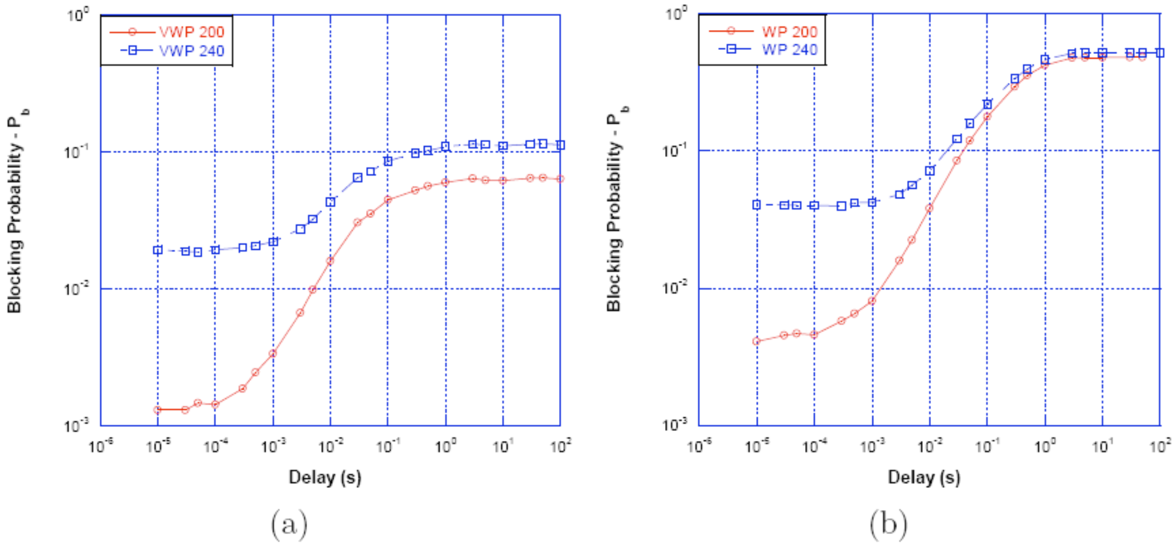


Figure 21. Total Pb for unprotected routing in (a) VWP and (b) WP case.

In Figure 22.(a) we have drawn P_b for increasing values of control delay in case of routing with shared and dedicated path protection under WP and VWP assumption. The effect of outdated information on P_b also in those cases follows the three phases described in the UN case. The values of P_b (both starting value in the first phase and the final value in third phase) are much larger in SPP case and even more in DED case due to the higher capacity requirements of the protected routing.

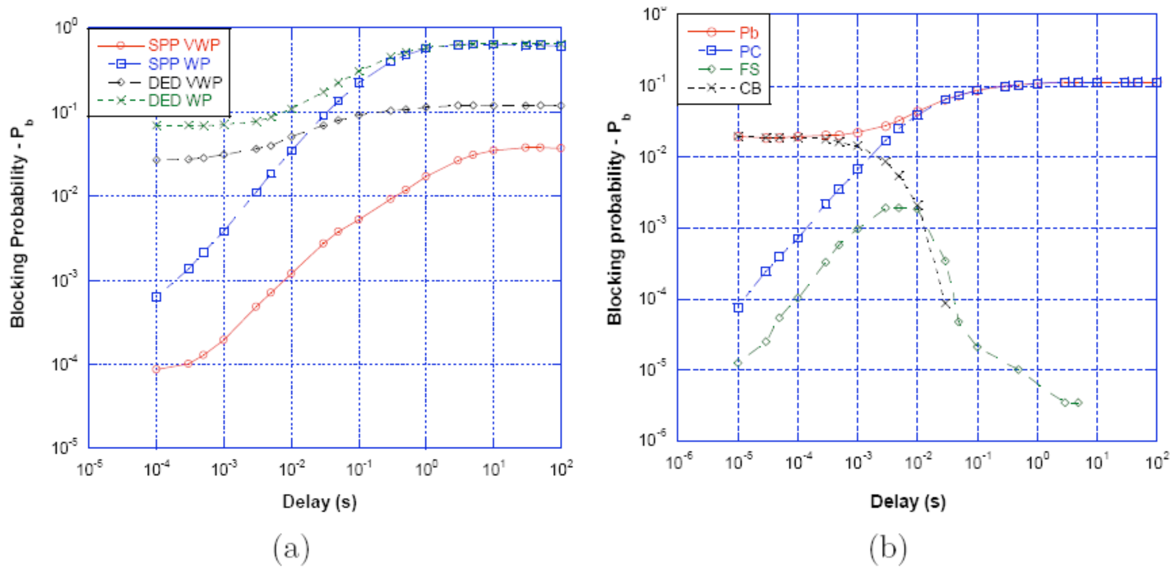


Figure 22. Total Pb for unprotected routing in (a) VWP and (b) WP case.

In Figure 22(b) the contribution of the various components is drawn for an offered load of 240Erl; for the sake of clarity also the overall blocking probability has been reported, which is the sum of all the other components in the graph.

The trends of the three components are strictly interdependent; in particular:

- for very small delays, FS and PS curves grow linearly, while the CB component is almost constant (actually it slightly decreases of a value equal to the sum of PS and FS that are very small in this phase) and largely constitutes the overall probability.
- when the PC contribution equals the CB contribution (at about D = 10^{-3}), the CB curve start rapidly decreasing, while the PC curve keeps growing linearly until it stabilizes for value of D comparable with the average holding time (i.e., about 1).
- also the FS contribution keeps growing until it reaches comparable values with CB. Then it decreases consistently to stability on very small probabilities below 10^{-5}.

The overall P_b is dominated initially by CB, then for higher delay values PC is the dominant term; FS plays a significant role only on a limited range of delays values after PS has intersected CB.

The main contribution of this activity has been achieved during this first year, leading to the submission of our results to a journal [5]. As future steps we aim at investigating some of these further aspects of the problem, namely:

- *Effects of Outdated Information for different traffic dynamism*. For a fixed value of traffic load in a network, the outdated information may induce different effects according to the characteristics of the traffic. It is intuitive that the same control delay has a weaker impact on a traffic composed by long and rare connections than on a traffic composed by short and numerous connections.
- *Comparison of simulative results with emulation results*. Shared path protection has been implemented over the test-bed ADRENALINE in CTTC. We plan to verify "on-field" the results emerging from our sulmulative analysis
- *Comprehensive analysis of resource overbuild for dedicated and shared path protection with and without wavelength conversion.*

REFERENCES

[1] IETF Network Working Group, "OSPF extensions in support of GMPLS," (2002). Internet Draft.

[2] A. Giorgetti, N. Sambo, I. Cerutti, and P. Castoldi, "Impact of link-state advertisemen in GMPLS-based wavelength-routed networks," in "Proc. of Optical Fiber Communication (OFC) Conference," (2008).

[3] X.Masip-Bruin, R.Munoz, S.Sanchez-Lopez, J.Sole-Pareta, J.Domingo-Pascual, and G. Junyen, "An adaptive routing mechanism for reducing the routing inaccuracy effects in an ASON," in "Proc. of ONDM," (2003).

[4] S. Darisala, A. Fumagalli, P. Kothandaraman, M. Tacca, and L. Valcarenghi, "On the convergence of the link-state advertisement protocol in survivable WDM mesh networks," in "Proc. of ONDM," (2003), pp. 433–447.

[5] M. Tornatore, F. De Grandi, R. Mu\~noz, R. Martinez, R. Casellas, A. Pattavina, "Effects of Outdated Control Information in Control-Plane-Enabled Optical Networks", submitted to OSA Journal on Optical Networking

3.8 Benchmarking of network architectures for guaranteed service provisioning (Netbench)

The selected methodology defined a set of requirements and assumptions to be used in the evaluation phase for the definition of performance evaluation scenarios. According to the selected parameters for benchmarking, several end-to-end networking scenarios can be modeled and simulated. The objective is to collect a set of results and expose several advantages or weakness of specific dynamic network reconfiguration, resource allocation and traffic forwarding mechanisms under specific operation conditions.

3.8.1 Benchmarking parameters

The first and one of the most significant parameter affecting the performance evaluation methodology is the set of requirements imposed by the network topology. Details about the reference network topology should be described in terms of: topology (connectivity), path distances and capacity. Depending on the case to be demonstrated either existing backbone network topologies (NSF network, European network) for which details are available in the literature or "ideal" reference topologies serving specific cases to be demonstrated can be used. Novel network architectures as CANON [1], an approach where the nodes of a wide area network (typically 20-30 nodes) are organized in clusters mainly based on vicinity, traffic, legacy infrastructure and administrative criteria, are also assumed.

The functionality of the nodes should also be specified. Most optical switching techniques require an adaptation unit employed at the network edge and potentially transparently switch traffic at the inner core network. As described in [1] in the CANON approach the nodes are separated as core and edge node depending on being an intra-cluster node (a node inside the cluster) or an inter-cluster node (Master Node, a node to connect with other clusters). From the reference above it can be easily pointed out that electrical buffers exist only in the periphery thus once a packet enters the optical domain it will travel until the destination without any further queuing while two granularity levels are used: the transceivers of the cluster node are operating on a slot-by-slot basis while data are exchanged between clusters in much longer fixed-size frames. Thus the switching node architecture and the wavelength conversion capability has also to be studied. Solutions like those presented in [1] will be evaluated.

In all cases, the traffic load used as an input has to be dynamic fluctuations of traffic approximating realistic conditions over a wide-area backbone network. As an example exponential distributions for interarrival times can be used or a characterization of traffic based on the RedIris network. Also different QoS classes will also be examined.

Since the benchmarking methodology extends to an overall network solution, the control plane architecture cannot be ignored. The usage of OBS ([4]) alternatives (TAW, JIT, JET) or GMPLS based protocols, create a major effect on the performance evaluation of the system.

Finally the performance metrics have to be pointed out. They can be divided in two major sections

- CAPEX: the switching node and the overall network cost. The cost can be extended in number of components used, cost for implementing the network system, or even in power consumption and switching or node dimensions 3.8.3.
- OPEX: traffic performance. Performance evaluation of the system that can be showed by metrics of delay, packet loss probability, jitter, utilization, etc .

3.8.2 Selected models and preliminary results

A study of the alternative CANON hierarchical network architecture in the NSF network has been evaluated where the 14 nodes of the NSF network were considered as MNs collecting regional traffic from clusters of 5 nodes (Figure 23).

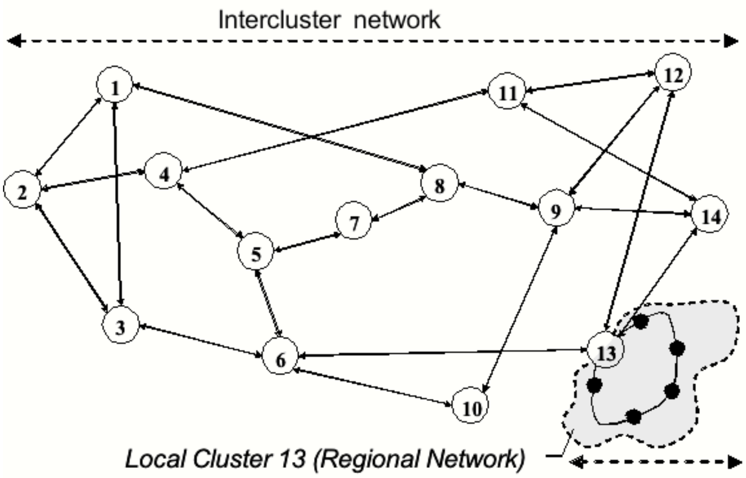


Figure 23. The inter-cluster network topology connecting 70 nodes organized in 14 clusters

On each link, W wavelengths of 10Gbps each carry data in frames of 5ms (two scenarios for W=4 and W=16 are shown here). All MNs are assumed to have full wavelength conversion capability (where W>1). The distances among the MNs were rounded up to the nearest multiple of frame durations (see [2] for further details). At the cluster nodes, Poisson sources simulate generation of packets with uniformly distributed destinations. An intra-cluster RTT of 5ms, corresponding to rings of 1000km periphery has been assumed.

The performance of the proposed scheme is compared to that of just-enough-time (JET) OBS ([4]) and OBS-INI schemes ([5]). Specifically OBS-JET is only examined for the 2nd step (inter-cluster dynamic reservations) and not as an end-to end solution for the overall network of 70 nodes, which would be totally impractical. To compare our results directly with those presented in [5] the comparison with OBS-INI has been made on the basis of using a single wavelength (W=1) and presenting aggregate results for intermediate initiating nodes residing 3-hops away (which is the best case presented in [5]).

In Figure 24 the average end-to-end delay for traffic generated in clusters 3 (highest average distances) and 8 (representing an average case) and destined to all other clusters is measured against the offered load as a percentage of link capacity. Infinite FIFO buffers per destination at the MNs are assumed in this scenario so as to evaluate the worst case for the overall delay. As observed CANON and OBS-INI introducing 2-way reservations exhibit higher average delay than plain OBS-JET, remaining though within acceptable limits for any type of service, especially if we take into account that we are considering an optical network of 14 nodes for OBS_INI and OBS-JET and 70 nodes in total for CANON counting the nodes of all clusters (including MNs).

In Figure 25 the loss performance of the proposed approach across all 70 nodes is compared with that of OBS-JET and OBS-INI only across the 14 core nodes respectively for different combinations of W and B. To bring losses at the MNs to measurable values, we constrain the edge MN's buffer capacity B per destination MN to small values. The proposed scheme, even with limited buffers, achieves significantly lower losses than when using OBS-JET, even only between MNs, whereas the slightly higher delay compared to OBS-INI is compensated by the lower loss probability. It is worth noting that due to the exclusive use of two-way reservations in CANON, losses are observed only due to buffer overflows. Increasing the buffer size can provide further improvements leading to very low loss probability.

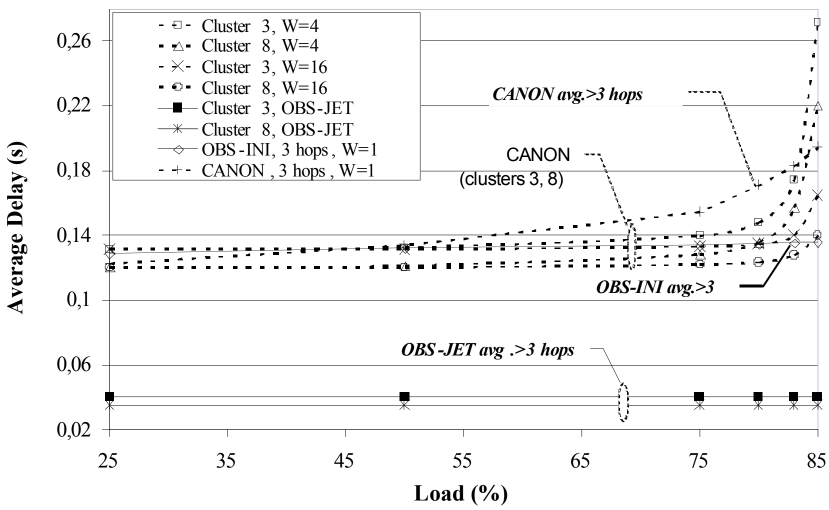


Figure 24. Average end-to-end delay vs. load

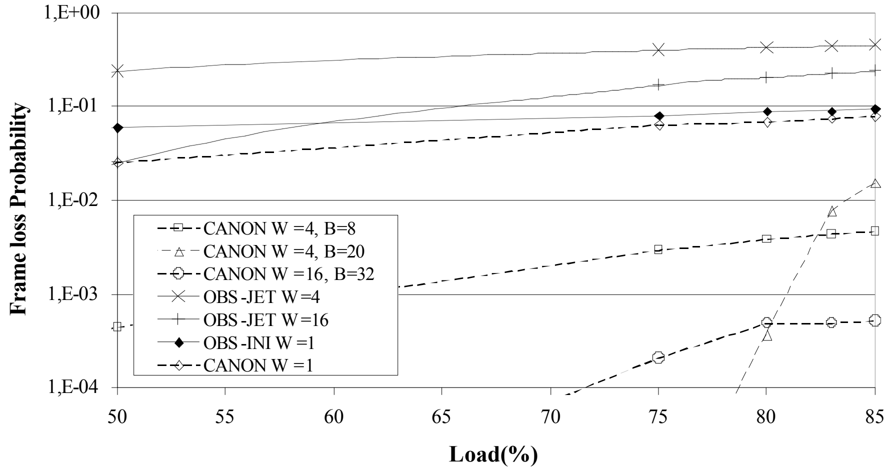


Figure 25. Loss probabilities for CANON vs. OBS-JET and OBS-INI

3.8.3 References

[1] A. Stavdas, T. G. Orphanoudakis, A. Lord, H. C. Leligou, K. Kanonakis, C. Matrakidis, A. Drakos, J. D. Angelopoulos, “Dynamic CANON: A Scalable Multi-domain Core Network”, to appear in IEEE Communications Magazine, Special Issue on Multi-Domain Optical Networks: Issues and Challenges, June 2008.

[2] J. D. Angelopoulos, K. Kanonakis, H.C. Leligou, Ch. Linardakis, I. Pountourakis, A. Stavdas, “Slotted Optical Switching with pipelined two-way reservations”, IEEE Journal of Lightwave Technology, vol.24, no.10, Oct. 2006, pp. 3616- 3624.

[3] C. Raffaelli, M. Savi, “Cost Comparison of All-Optical Packet Switches with Shared Wavelength Converters”, IEEE ICTON, Invited paper, Rome, Italy, July 2007.

[4] C. Qiao, M. Yoo, “Optical burst switching (OBS) — a new paradigm for an optical internet”, Journal of High Speed Networks, Vol. 8, No. 1,1, pp. 69–84, 1999

[5]V. Vokkarane, “Intermediate-node-initiation (INI): A generalized signaling framework for optical burst-switched networks”, Optical Switching and Networking, Vol. 4,pp. 20–32, Feb. 2007

4. Implementation and evolution scenarios

4.1 Experimental tests of Carrier Ethernet techniques

Aim of such Joint Activity is the experimental demonstration of techniques for core and metro networks regarding all the routing processes that are migrating from “layer 3” to “layer 1” with particular interests for protection, restoration, traffic engineering, congestion resolution and resources allocation. In this migration towards layer 1 “Carrier Ethernet” approach will have a fundamental role and therefore such a topic will dominant in such JA.

The activities performed in the first year in this JA have been meanly focused to the management of Quality of Service (QoS) in networks based on Ethernet transmission with introduction of novel optical devices as All Optical Wavelength Conversion (AOWC). These activities have been based on experimental investigations carried out by means of the multi-access multiservice IP test bed, implemented in the framework of the E-Photon/One project. In particular, in this year our investigations have regarded some developments of VPLS, assumed as a technique close to layer 2 processing, both for traffic control in accesses based on ADSL2+ and to enable the introduction of the AOWC in optical networks. From the theoretical point of view an investigation was performed on GMPLS-controlled Ethernet segment protection & BFD over Ethernet and on the evolution of networks based of Ethernet transmission towards the Optical Burst Switching (OBS) paradigm.

4.1.1 Experimental Test-bed

The optical network Test-Bed is shown in Figure 26 [1]. The core part is composed of four Juniper M10 routers, with ZX GbE interfaces (1550 nm), fully meshed using the fibers contained in the cable Roma-Pomezia-Roma (50 Km with round trip in Pomezia). The edge (metro) part is composed of three Cisco 3845 edge routers, connected with three Juniper routers by means of GbE fiber transmission. The access part consists of xDLS and Fiber To The x (FTTx) systems. The FTTx consists of an Ethernet Passive Optical Network (EPON), AN5116-03 ePON FiberHome, with an Optical Line Termination (OLT) and up to eight Optical Network Units (ONU). The OLT and the ONUs are connected by means of a single mode fiber with the wavelength for downstream at 1490 nm and the wavelength for upstream at 1300 nm.

The seven routers were set up according to OSPF (Open Shortest Path First), MPLS and BGP (Border Gateway Protocol) protocols. A network generator/analyzer (SmartBits 6000) could be inserted in the network to overload links with a total traffic around 1 Gbits/s with modality showed in ref. [1], permitting to analyze the QoS in the presence of network congestion. The reliability of such a network can be evaluated by means of QoS measurements both with network tests (objective measurements) and with subjective, or perceptive, tests (subjective measurements) as showed in ref. [1]. In particular network tests were carried out by using a network analyzer software: NetIQ Chariot that allows us to evaluate some network parameters like throughput, jitter and data loss.

In the context of MAN/WAN, the Virtual Private LAN Service (VPLS) [1] is currently one of the technique that realizes Ethernet connections employing MPLS Label Switched Paths (LSPs) and that permits to achieve excellent network performances in terms of traffic management and Quality of Service (QoS), with multicasting capabilities that result particularly suitable for triple play services. VPLS is a kind of Layer 2 Virtual Private Network (VPN) where the customers are connected by Ethernet Line/LANs; with respect to the conventional Layer 2 VPNs where the customers are connected in a point to point way, in VPLS customers can be connected by a multipoint Ethernet LAN. The main feature is that users sharing VPLS seem to be in the same LAN, independently by their own geographic position.

The network elements involved in VPLS are Customer Edge (CE) and Provider Edge (PE) (Figure 26): the CE is completely unaware of VPLS, it only provides to set “attachment circuit” (VLAN). PE instead is the real maker of VPLS, it elaborates the original frame Ethernet adding a VPLS label.

According to these considerations the test bed is configured to operate with CoSs based on VPLS, that are extended to the CE sites by means of VLAN Tagging technique [2][3]. Therefore, a tunnel based on VPLS&VLAN Tagging between a Server (PC1) connected to CE1 and a user PC connected to ONU2 can be established, as shown in Figure 26 by the bold line, according to specific service requirements. In this Test Bed eight different CoSs can be defined, according to the values of three bits, in agreement with IEEE 802.1p and DiffServ over MPLS standards. In particular the VLAN TAG CoS field (3 bits) is employed to guarantee QoS among CEs and PEs; the PE adds an outer MPLS label and the QoS among PEs is guaranteed using the MPLS LABEL EXP field (3 bits). Such an architecture results well suitable to forward the traffic on dedicated paths allocated in selected wavelengths. This aspect blends with future WDM networks, where each wavelength will be characterized in terms of transmission impairments and, as a consequence, traffic with higher priority could be allocated on wavelengths with higher performance. Therefore, VPLS helps to introduce a new “dimension” for Service Level Agreement based on the wavelength channel characteristics. The WC devices allow us to widen the use of the QoS based on WDM since wavelengths can be changed along the network propagation.

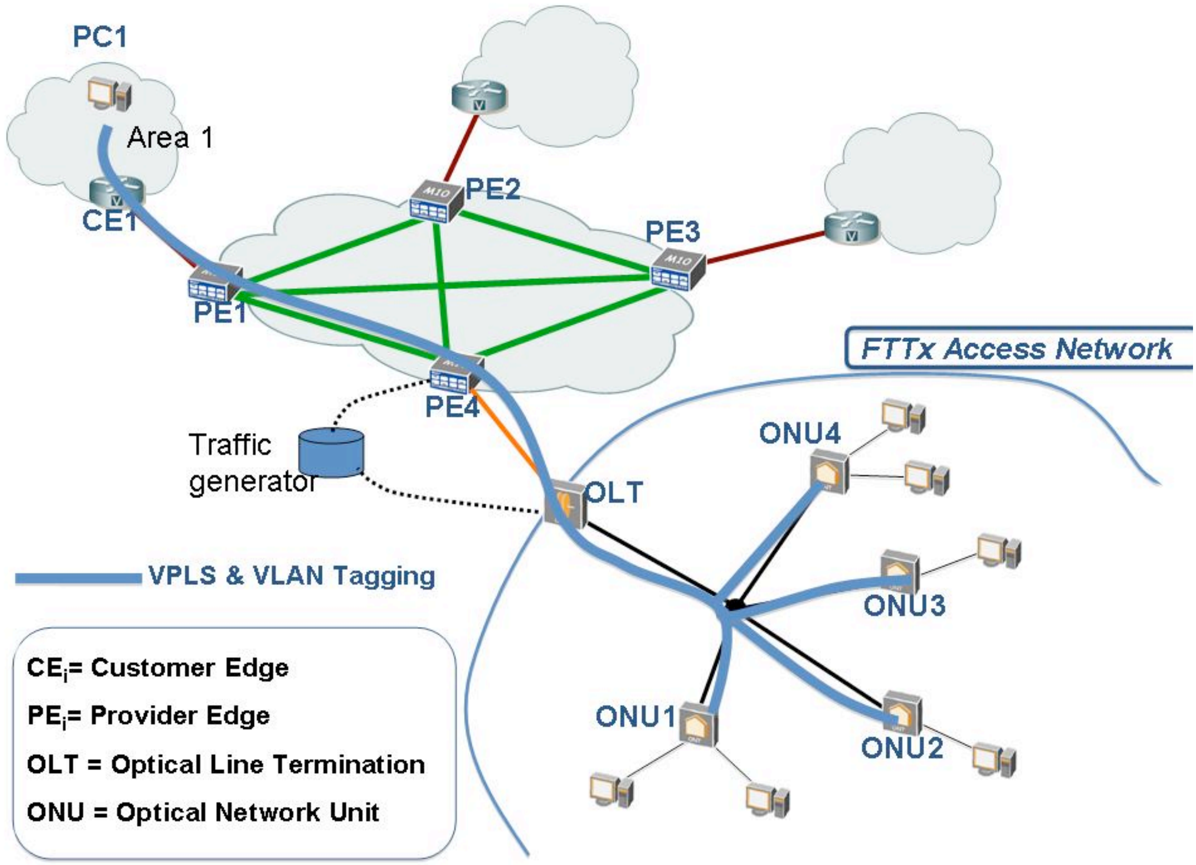


Figure 26. Test-Bed design. The routers, OLT and ONUs are optical fiber interconnected with GbE interfaces.

QoS of each CoS depends on network congestion in different ways; the worst performance is shown by the *Best Effort* (BE), conversely the *Gold* behaves as the best one and permits to guarantee the necessary performance (throughput, jitter, data loss) independently of the network congestion [3].

4.1.2 VPLS for traffic control in the access network

Here we report an investigation on the role of VPLS to control the traffic in the edge routers to guarantee the flux in the access part of the network having bandwidth limitations, as in the case of xDSL terminations. The technique that we used to guarantee QoS consists in the association between the VLAN TAG, that determinates a particular "attachment circuit" of the VPLS, and a suitable CoS. The aim of this proposed technique is to employ a particular VPLS instance for any kind of service; we suppose three different services, a VIDEO traffic, an High Speed Internet traffic and Voice traffic, and we associated three different VLANs; for each VLAN we associated three different priority classes, respectively: Gold, Silver and Bronze. The VPLS employs MPLS protocol; to guarantee QoS within MPLS protocol, each tunnel employs the E-LSP technique. Actually this method consists of marking the 3 bit EXP field of the MPLS shim header according with some priority rules. Using the sum of VLAN TAG based QoS technique and E-LSP technique, we have a strictly association between the kind of services and QoS parameters; besides we obtain an end to end Quality of Service.

Such a method was used to select the traffic towards the ADSL2+ modem. Two data fluxes were addressed to the modem with a total bit rate higher than 20 Mb/s: a first one with an input throughput of 20 Mbps that includes a video and the latter corresponding to a video with a throughput of 5 Mbps. In the absence of traffic treatment with CoS severe degradation could be experienced by the user for both videos caused by the broadband bottleneck. By means of the VPLS approach the user can select the required service in order to better exploit the available bandwidth.

The results of such approach are reported in Figure 28, where we show the throughput at the user location versus time for the two fluxes. The first (20 Mb/s) is treated with a BRONZE CoS and the latter with a GOLD CoS. Figure 28 confirms the importance of such a method and in fact after 5 sec, when the gold video is sent, a congestion occurs, but, while the gold flux is maintained the bronze is limited. As a consequence the QoS of the Gold is excellent while the Bronze one is strongly degraded. To better illustrate the effects of such traffic treatment two video screenshots are also reported to show the evident perceivable difference between the low priority and the high priority traffic.

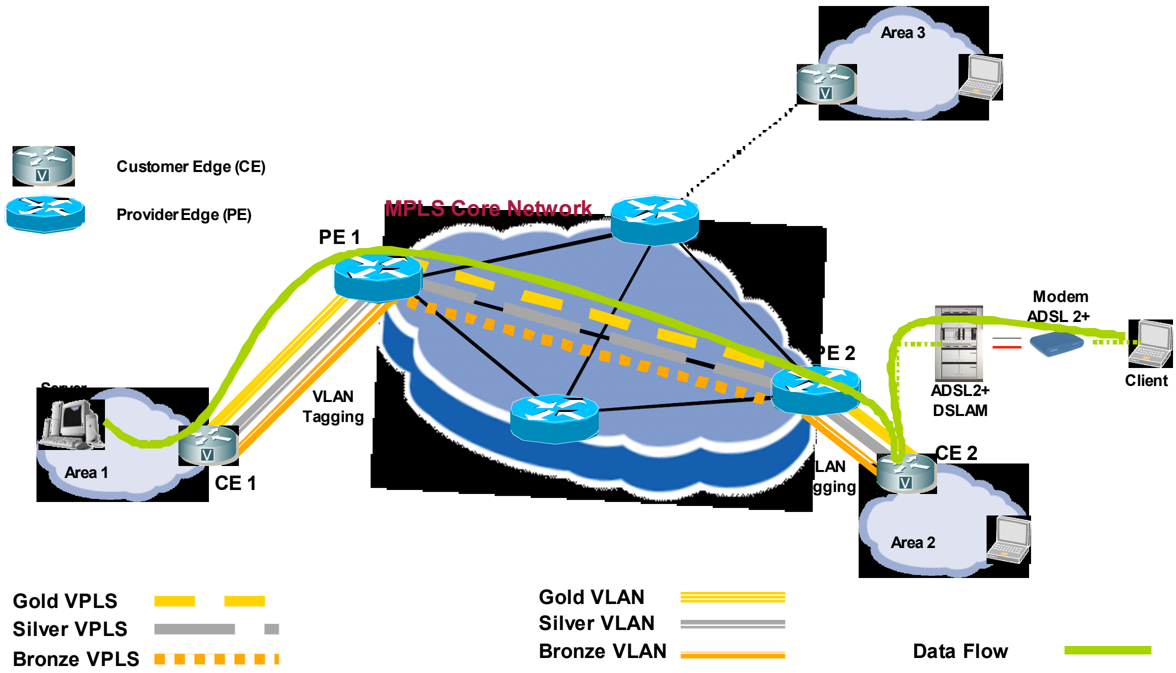


Figure 27. VLAN based QoS technique

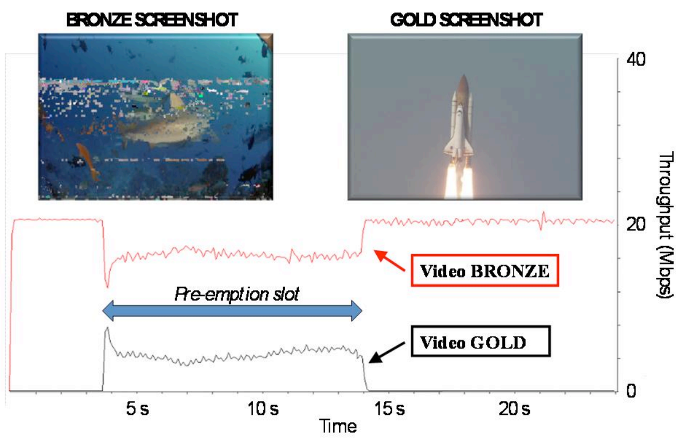


Figure 28. Low priority and High priority traffic treatment with VPLS.

4.1.3 All optical wavelength conversion in GbE networks based on VPLS

As reported in ref. [3] VPLS is an architecture that results well suitable to forward the traffic on dedicated paths allocated in selected wavelengths. This aspect blends with future WDM networks, where each wavelength will be characterized in terms of transmission impairments and, as a consequence, traffic with higher priority could be allocated on wavelengths with higher performance. Therefore, VPLS helps to introduce a new “dimension” for Service Level Agreement based on the wavelength channel characteristics. The WC devices allow us to widen the use of the QoS based on WDM since wavelengths can be changed along the network propagation.

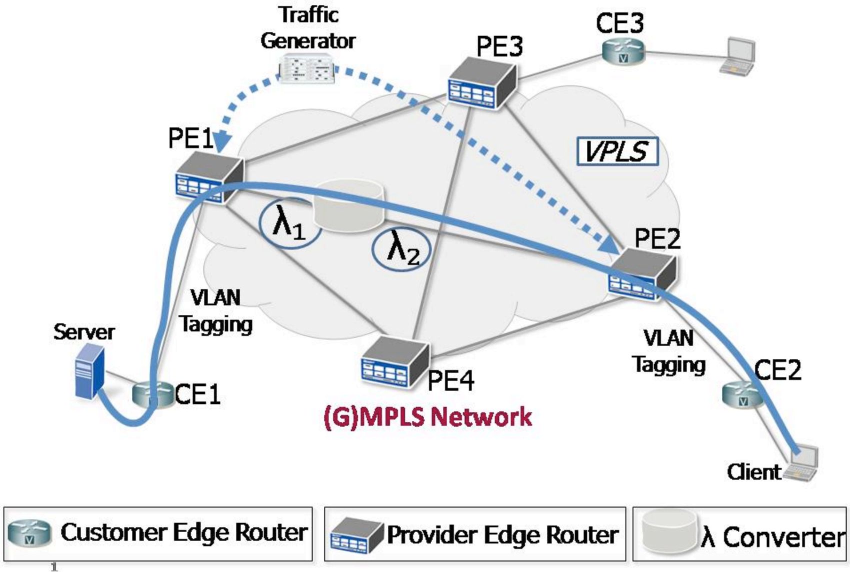


Figure 29. Network test bed based on VPLS (in the cloud) and VLAN tagging with WC (λ converter).

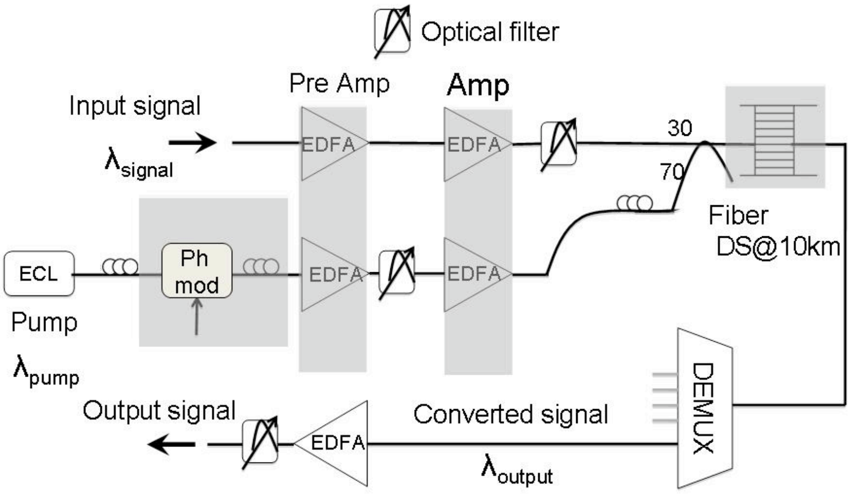


Figure 30. Wavelength converter set-up: Ph Mod phase modulator to control Brillouin effect.

Network tests for QoS measurements were done using a network analyzer, including the traffic generator, that allows us to evaluate some network parameters like throughput, jitter and data loss. We analyzed streaming from the Server (linked to CE1) to the user PC representing the client (linked to CE2) as shown in Figure 29. The WC is located at the output of PE1 and its scheme is reported in Figure 30.

The input signal (λ_1=1553,30 nm) from PE1 is amplified by a dual optical amplificatory stage and it is mixed inside the DS fiber (10 km long) with the amplified pump coming from an external cavity laser (ECL) (λ_p=1553,76 nm). The signal power at the DS fiber input is 6 dBm and 7.6 dBm for signal and pump respectively. The output wavelength is selected by means of the optical demultiplexer (DEMUX) that permits to switch between λ_1 and the converted λ_2. DEMUX is controlled by a software running on a LINUX server .

In Figure 31 we report the optical spectrum at the output of the DS fiber that shows the signal replicas generated with high efficiency by means of the FWM effect; in our experiment we use the wavelength at 1554,12 nm for λ_2. As reported in [2], the efficiency of such a process can be much high and it depends on the pump power; furthermore, it is a wide bandwidth effect since it can be achieved in all the C bandwidth. The signal at the

DEMUX output is amplified and filtered (Fig. 4), it is transmitted in the link between PE1 and PE2, and received by PE2.

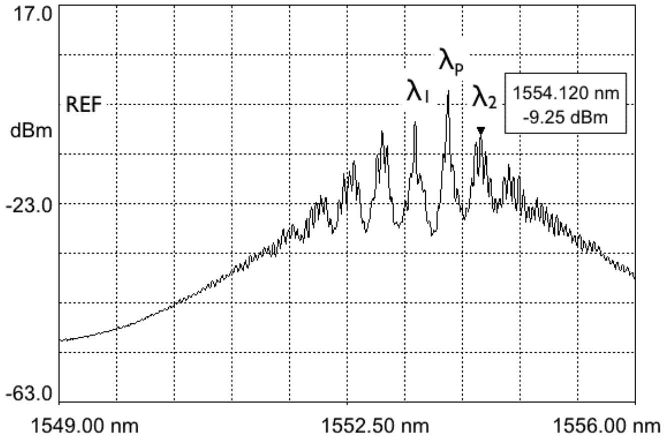


Figure 31. Fig. 7. Optical spectrum at the DS fiber output. Resolution:10.0 dB/D, 0.05 nm

During the wavelength switching operation an out of service of 20 ms is manifested, as shown in Figure 32, due to the physical switching (3 ms) and to the restoration procedures among the routers (17 ms). This time (20 ms) is a negligible out of service for most of the services, but it could be an impairment for High Definition (HD) IPTV that requires a maximum duration single error time lower than 16 ms (Broadband-Forum TR 126). However it has to be pointed out that such an impairment could degrade HDTV based on MPEG2, conversely MPEG4 is able to recover packet loss. Apart this out of service, we did not reveal further degradations due to WC. In Figure 33 network tests in the presence of WC are reported: it is shown the throughput at client site versus time for 10 Mb/s data streaming, that is the bit rate required for HDTV based on MPEG4 (Broadband-Forum TR126). BE and *Gold* classes are monitored before, during and after the wavelength conversion process, both in the presence and in the absence of the traffic congestion. The time of the wavelength switching is indicated in the figure as *λ conversion point*. Degradation is shown only for BE class and only during the congestion interval, independently of the WC process. Measurements on jitter revealed that it was always lower than 2 ms for Gold and lower than 4 ms for BE, that is lower than the maximum acceptable (50 ms). Data loss measurements showed that, apart the loss during the out of service, for Gold no data loss was observed; conversely, for BE we observed no data loss in absence of congestion and an average data loss equal to 5% (maximum 8%) during the congestion, independently of the presence of WC. These data confirmed that WC maintains the QoS properties of VPLS. Measurements on the other six CoSs showed that a degradation was always present only during the traffic congestion, even though in different weaker ways with respect to BE.

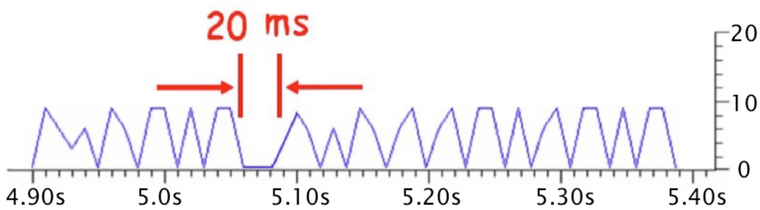


Figure 32. Out-of-service time during wavelength switching for 10 Mb/s data streaming.

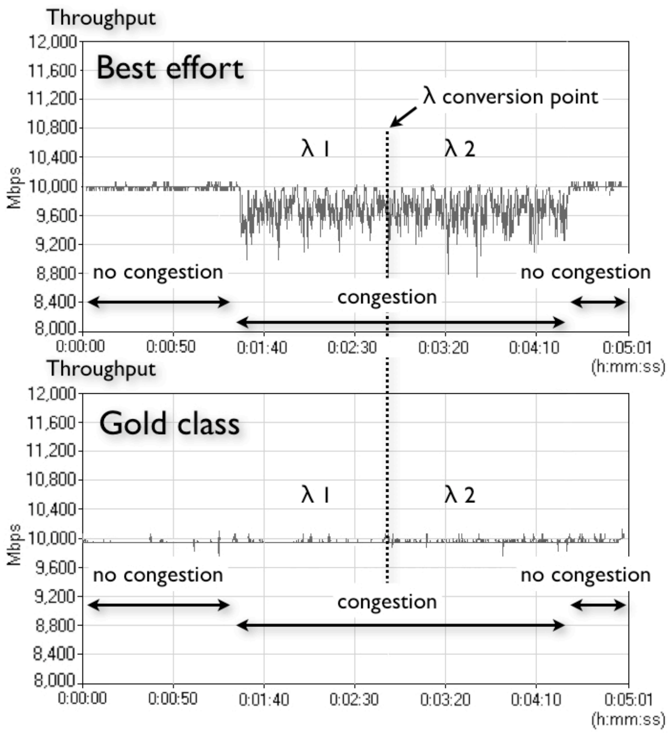


Figure 33. Throughput vs elapsed time for Best Effort and Gold classes in VPLS environment.

These network measurements were also confirmed by perceptive (or subjective) measurements reported in terms of the average evaluation given by a panel of 16 viewers by means of the technique reported in [1]. To stress the effects of network impairments we tested High Definition TV 1080i streaming with MPEG2 codec from Server to PC.

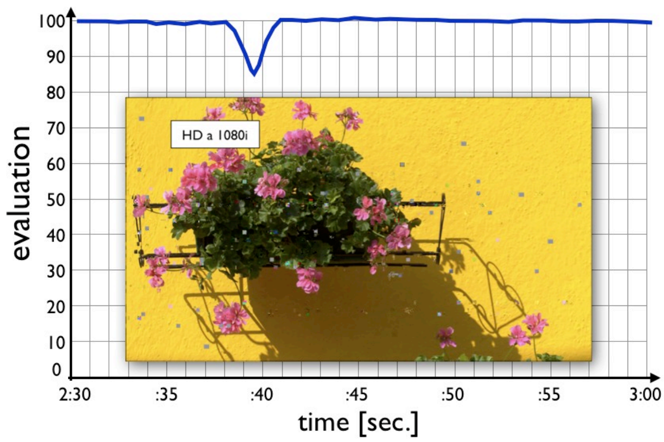


Figure 34. Perceptive evaluation vs elapsed time for Gold case. A screenshot taken from HD video streaming during the WC switching process is also included

In BE case, degradation was observed in the presence of congestion independently of the presence/absence of WC. In Gold case only during the wavelength conversion switching the viewers manifested a weak evaluation decreasing, due to some jeopardized frames, as shown in Figure 34.

In conclusion we have shown the introduction of the wavelength conversion in a wide area optical Gigabit Ethernet network, verifying that such a process does not introduce any signal degradation and it can be achieved in a very fast time, compatible with wide bandwidth real time services, including high definition videos. Furthermore, we have shown that forwarding processes based on VPLS can introduce a QoS traffic management based in the wavelength domain.

These results confirm that the Ethernet architectures can receive several advantages from the WDM techniques, in particular in terms of OAM and QoS. The wavelength conversion will permit a wider use of OAM functionalities and a better QoS management, especially if simpler and cheaper wavelength converters will be adopted, as for instance by adopting the semiconductor optical amplifiers.

4.1.4 Loss measurements in Optical Cross Connect

In wide area networks, GbE transmission will have to cope also with novel devices as OXC. Before the introduction of OXC in a GbE test bed Ericsson made characterization of OXCs in terms of losses and in this report documents the testing of 16x16 OXC for wavelength dependence. Two OXCs were tested for wavelength dependence of the insertion loss within the wavelength bands of 1530 – 1570 nm, 1570 – 1610 nm and 1525 - 1610 nm. 36 interconnection paths of each OXC were characterized.

WDL (dB)	1530-1570 nm	1570-1610 nm	1525-1610 nm
Average	0.11	0.13	0.22
Standard Deviation	0.06	0.06	0.09
Maximum	0.33	0.33	0.50
Minimum	0.02	0.02	0.04

A4.1.4.1 Summary of Results

Wavelength dependent loss (WDL) of a channel is the difference between the maximum and the minimum insertion loss of that channel when the wavelength of the light is varied within the specified wavelength band.

The data below shows the WDL statistics as measured in the three wavelength bands for the two OXCs.

A4.1.4.1.1 Test details

This test was conducted on two 16x16 OXC randomly selected for this test. The OXCs had insertion loss measured at different wavelengths, the maximum and minimum value in each band was recorded, and the peak-to-peak insertion loss difference was calculated.

A4.1.4.1.2 Data Summary:

The graphs below show the insertion losses for the individual mirrors in the two wavelength ranges. The responses of the 36 different paths are shown in 4 separate graphs.

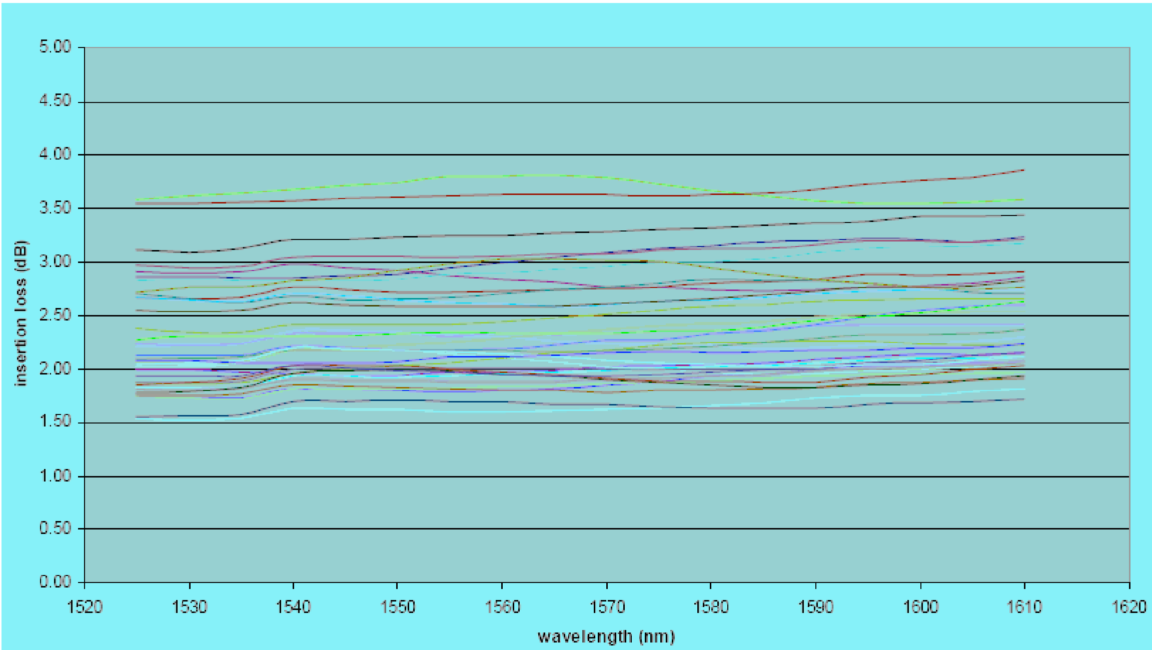


Figure 35.   wavelength range 1525 - 1610 nm (36 paths)

The histograms below show the histograms of the wavelength related peak-to-peak insertion loss variations for both OXCs combined.

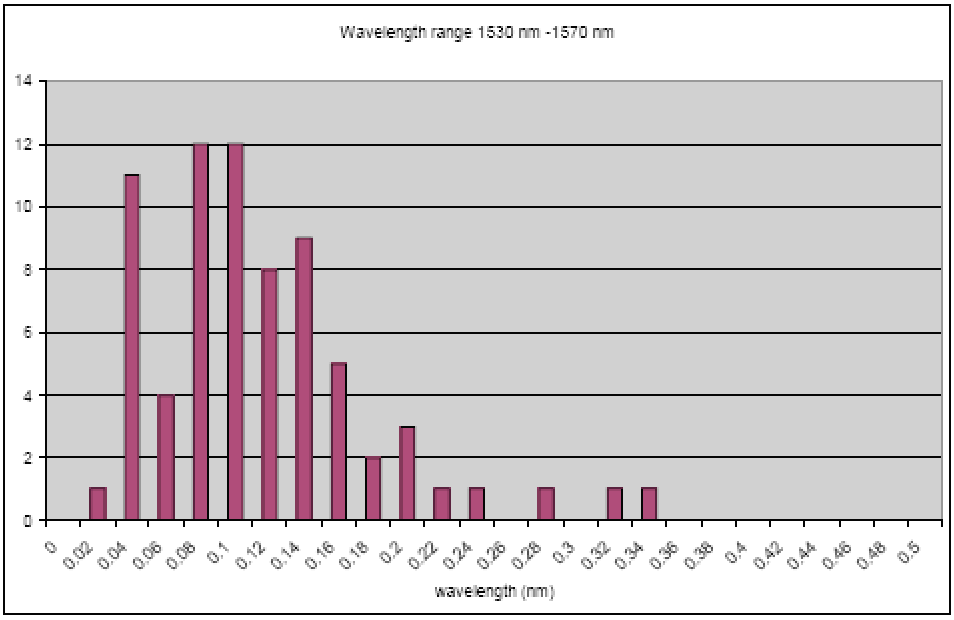


Figure 36.   Range 1530 – 1570 nm

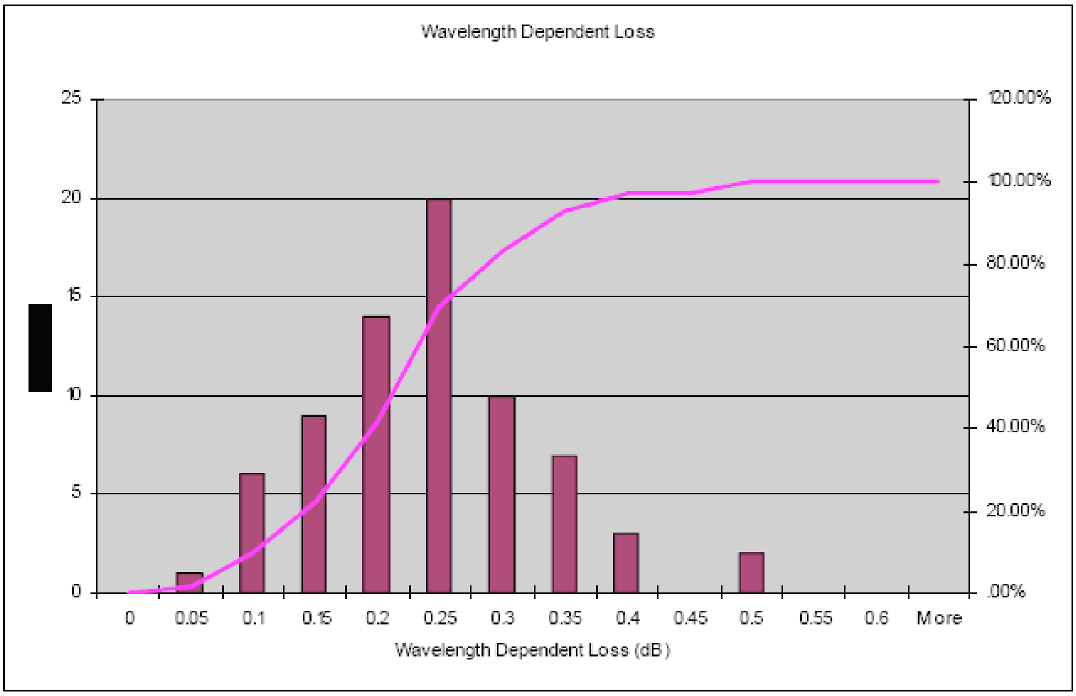


Figure 37.   Range 1525-1610 nm.

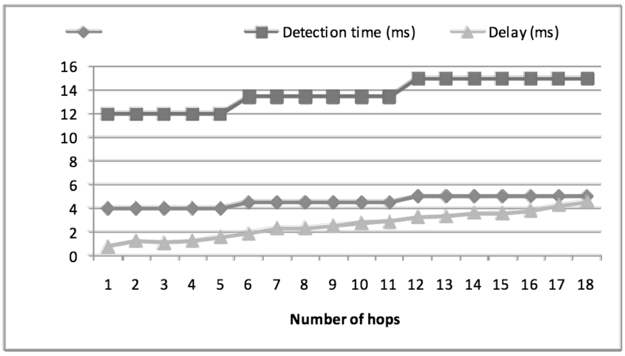
4.1.5 GMPLS-controlled Ethernet segment protection & BFD over Ethernet

Ethernet Label Switching (ELS) is a new evolving connection-oriented Carrier Ethernet technique. As is the case with PBB-TE, it disables RSTP, MAC learning and flooding and allows the usage of GMPLS as control plane such as to enable the setup of connections allowing traffic engineering and advanced OAM functionality. ELS forwarding uses the translation capability of IEEE 802.1ad-based switches to allow for connections based on link local S-VID (Service VLAN ID) switched frames. This way a frame going through an ELS connection, undergoes a series of push and pop operations on the SVID field similar as in MPLS. The resulting connection is called an Ethernet Label Switched Path (ELSP).

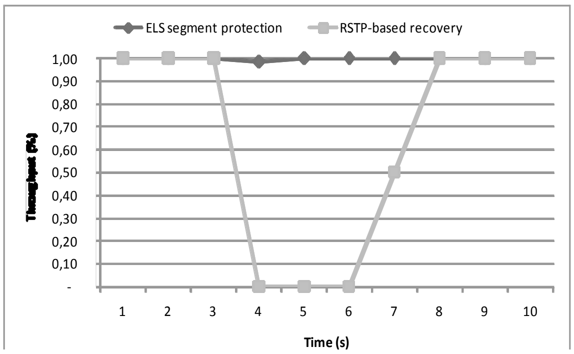
Using this forwarding architecture, GMPLS-controlled segment protection of ELSPs now becomes possible as in defined in IETF RFC 4873. This way, any segment between two points in a network can be protected by a precomputed backup segment, provided that a cycle in between both points is available. Before failure switchover can be settled, it needs to be detected. A lightweight protocol to detect failures is Bidirectional Forwarding Detection (BFD), as specified in [1]. This Hello-protocol is able to detect failures in the order of tens of ms. In extension to existing BFD specifications for IPv4 and IPv6, we defined an extension such as to enable BFD sessions running directly over Ethernet, omitting IP/UDP encapsulation, and directly using MAC-addresses as BFD endpoints.

An emulation platform based on open-source software running on Linux was set up such as to benchmark the performance of GMPLS-controlled ELS and its capability to provide segment protection. We emulated the ELS data plane and BFD over Ethernet using the Click Modular Router Platform ([2]). The control plane software used for tests is based on VLSR software of the Dragon project ([3]). We extended this GMPLS package with protection signaling and notification as defined in RFC 4873 and the signaling of ELSPs.

The emulation results in Figure 38 were obtained using our BFD over Ethernet implementation and comparing standard RSTP Ethernet recovery with GMPLS-controlled ELS recovery.



Minimal BFD failure detection time



Throughput on failure: RSTP vs. ELS segment recovery

Figure 38. Emulation results

4.1.6 Numerical Investigation on Hybrid GbE-optical Burst Switching Networks

In 2008 the research group of University of Modena and Reggio Emilia (UNIMORE) has started investigating the performance of hybrid optical networks.

In particular, TCP performance have been evaluated when high speed Ethernet over Passive Optical Networks (EPON) are interconnected by means of a core optical network based on the Optical Burst Switching paradigm.

Passive Optical Network (PON) is a promising technology to solve the last mile problem. EPON has been regulated by IEEE and it has been one of the subjects of first year activities.

The inter-working unit, or edge node, between these two networks has to be properly studied and designed and its main function is the burst assembly: it must collect incoming EPON frames, extract IP datagrams and assembly them into optical bursts according to proper assembly algorithms.

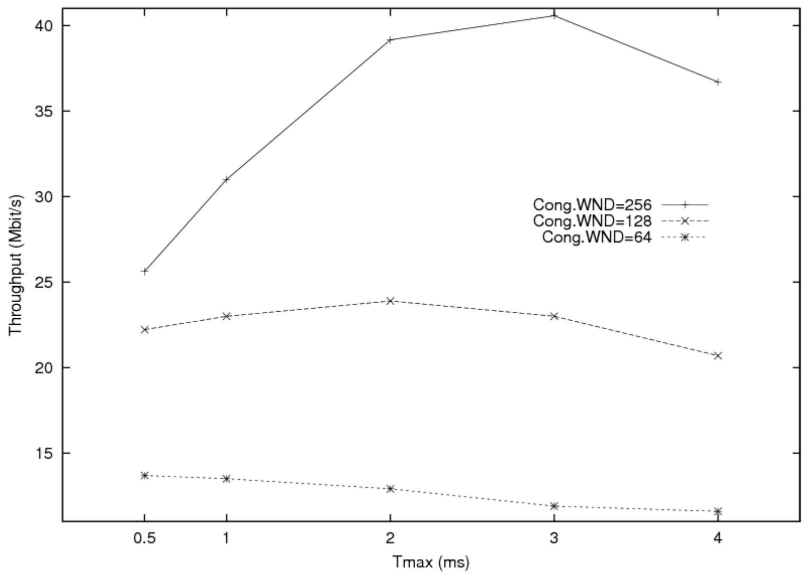


Figure 39. Average throughput for 10 TCP clients (ONUs) vs. T_{max} with Burst Loss Prob.=10^{-3} and different values of the TCP congestion window.

In particular, a time-based assembly algorithm has been employed and analyzed for operating with a EPON which is characterized, on the other hand, by some other parameters such as the cycle time, for managing the upstream transmission.

This hybrid network has mainly been studied through simulations by means of the ns2 simulation tool. Figure 11 and more numerical investigations have revealed that: (i) the cycle time in EPON has negligible impact on TCP performance, (ii) the TCP congestion window is very critical and it leads to remarkable performance differences, (iii) for different values of the congestion window a given assembly time value exists which maximizes performance, (iv) by increasing the assembly time a better intra-fairness index can be obtained, for a better bandwidth sharing.

4.1.7 Open issues and topics planned for 2009:

In 2009 we will continue the investigation on techniques for core and metro networks regarding all the routing processes that are migrating from "layer 3" to "layer 2", with the goal to reach a complete transport based on Ethernet (Carrier Ethernet). Therefore specific studies will regard all the techniques that, in different ways, are approaching to Carrier Ethernet: VPLS, T-MPLS, IEEE 802.1q (Q in Q), IEEE 802.1ad (MAC in MAC) and PBB TE, taking into account topics, of protection, restoration, traffic engineering, congestion resolution and resource allocation, especially for architectures based on GMPLS control plane.

Furthermore assuming the use of Ethernet transmission on wider and wider geographical areas, we will take into consideration the impact of optical processing techniques, typical of backbone networks, on Carrier Ethernet, and therefore we will investigate on GbE WDM transmission, on the control based on GMPLS, on the introduction of all optical devices as OXC, wavelength converters and 3 R regenerators, and we also investigate on interplay with future paradigm as Optical Burst Switching (OBS).

Comparison in terms of network performance will be carried out between Carrier Ethernet approach and conventional circuit switching (SDH/SONET) one.

4.1.8 References:

[1] S. Pompei, L. Rea, Luca; F. Matera, A. Valenti, "Experimental Investigation on Optical Gigabit Ethernet network reliability for high-definition IPTV services," J. of Optical Networking, Vol.7, n.5, pp. 426-435, 2008

[2] F. Matera, L. Rea, M. Venezia, L. Capanna, G. Del Prete, "Fast Restoration based on alternative wavelength paths in a wade area optical IP network,"Fiber and Integrated optics, Vol.26, No.2, pp 123-130, 2007

[3] A. Valenti, S. Pompei, L.Rea, F.Matera, G. Tosi Beleffi, M. Settembre, "Network Performance Investigation in a Wide Area Gigabit Ethernet Test Bed adopting the All Optical Wavelength Conversion" in press on Photonics Technology Letters

[4] D. Katz, IETF draft “Bidirectional Forwarding Detection”, draft-ietf-bfd-base-08.txt
[5] E. Kohler, “The Click modular router”, ACM Transactions on Computer Systems, 1999, Volume 18 number 3, p236 –297
[6] X. Yang, “GMPLS-Based Dynamic Provisioning and Traffic Engineering of High-Capacity Ethernet Circuits in Hybrid Optical/Packet Networks”, INFOCOM 2006, April 2006, p1-5, Barcelona, Spain.
[7] M. Casoni (UNIMORE), A Simulation Study of the IPACT protocol for Ethernet Passive Optical Networks, ICTON 2008 International Conference on Transparent optical Networks, Vol. 4, pp. 309-311, Athens (Greece), June 2008.

Annex 1: Bibliography of the survey on OBS routing methods for OBS networks

[ACP04] A. Agusti and C. Cervello-Pastor. A new contentionless dynamic routing protocol for obs using wavelength occupation. In Proceedings of the 12th IEEE Mediterranean Electrotechnical Conference (MELECON 2004), Dubrovnik, Croatia, May 2004.

[AdDA07] C. G. Argos, O. González de Dios, and J. Aracil. Adaptive multi-path routing for obs networks. In Proceedings of the 9th IEEE International Conference on Transparent Optical Networks (ICTON2007), Rome, Italy, July 2007.

[Bou03] N. Boudriga. Optical burst switching protocols for supporting qos and adaptive routing. Computer Communication Journal, 26(15):1804.1812, 2003.

[CEJ05] T. Coutelen, H. Elbiaze, and B. Jaumard. An efficient adaptive offset mechanism to reduce burst losses in obs networks. In Proceedings of Proceedings of IEEE Global Communications Conference (GLOBECOM 2005), St. Louis, MO (USA), December 2005.

[CMC06] Q. Chen, G. Mohan, and K.C. Chua. Route optimization for efficient failure recovery in optical burst switched networks. In Proceedings of IEEE High Performance Switching and Routing workshop (HPSR 2006), Poznan, Poland, June 2006.

[CWXQ03] Y. Chen, H. Wu, D. Xu, and C. Qiao. Performance analysis of optical burst switched node with deflection routing. In Proceedings of IEEE Inter-national Conference on Communication (ICC 2003), Seattle, USA, May 2003.

[CZZ04] C. Cameron, A. Zalesky, and M. Zukerman. Shortest path prioritized random deflection routing (sp-prdr) in optical burst switched networks. In Proceedings of ICST International Workshop on Optical Burst Switching (WOBS), San Jose, USA, 2004.

[DPZQ06] Y. Du, T. Pu, H. Zhang, and Y. Quo. Adaptive load balancing routing algorithm for optical burst-switching networks. In Proceedings of Optical Fiber Communication Conference (OFC 2006), Anaheim, CL (USA), March 2006.

[GBIQ04] S. Ganguly, S. Bhatnagar, R. Izmailov, and C. Qiao. Multi-path adaptive optical burst forwarding. In Proceedings of IEEE High Performance Switching and Routing workshop (HPSR 2004), Phoenix, AR (USA), April 2004.

[GZ06] D. Gao and H. Zhang. Information sharing based optimal routing for optical burst switching (obs) network. In Proceedings of Optical Fiber Communication Conference (OFC 2006), Anaheim, CL (USA), March 2006.

[HAM+05] Pan H., T. Abe, Y. Mori, Ch. Young-Bok, and H. Okada. Feedback based load balancing routing for optical burst switching networks. In Proceedings of the 11th Asia-Pacific Conference on Communications (APCC 2005), Perth, Western Australia, October 2005.

[HHM05] Y. Huang, J.P. Heritage, and B. Mukherjee. Dynamic routing with preplanned congestion avoidance for survivable optical burst-switched (obs) networks. In Proceedings of Optical Fiber Communication Conference (OFC 2005), Anaheim, CL (USA), March 2005.

[HLH02] C. Hsu, T. Liu, and N. Huang. Performance analysis of deflection routing in optical burst-switched networks. In Proceedings of the 21st Joint Conference of IEEE Computer and Communications Societies (INFOCOM 2002), New York, NY (USA), June 2002.

[HN04] E. Hyytia and L. Nieminen. Linear program formulation for routing problem in obs networks. In Proceedings of the 9th IEEE Symposium on Computers and Communications (ISCC 2004), Alexandria, Egypt, June – July 2004.

[HTM06] Y. Hirota, H. Tode, and K. Murakami. Cooperation method considering wavelength assignment and routing problem in optical burst switched networks. In Proceedings of Optical Fiber Communication Conference (OFC 2006), Anaheim, CL (USA), March 2006.

[IYS05] D. Ishii, N. Yamanaka, and I. Sasase. Self-learning route selection scheme using multipath searching packets in an obs network. Journal of Optical Networking, 4(7):432.445, 2005.

[KG03] M. Köhn and C. Gauger. Contention-based limited deflection routing in obs networks. In Proceedings of Global Telecommunications Conference (GLOBECOM 2003), San Francisco, CL (USA), December 2003.

[KKK02] S. Kim, N. Kim, and M. Kang. Contention resolution for optical burst switching networks using alternative routing. In Proceedings of IEEE International Conference on Communications (ICC 2002), New York, NY (USA), April-May 2002.

[LKSG03] S. Lee, H. Kim, J. Song, and D. Griffith. A study on deflection routing in optical burst-switched networks. Photonic Network Communications, 6(1):51.59, 2003.

[LLGC06] J. Lu, Y. Liu, M. Gurusamy, and K.C. Chua. Gradient projection based multi-path traffic routing in optical burst switching networks. In Proceedings of IEEE High Performance Switching and Routing workshop (HPSR 2006), Poznan, Poland, June 2006.

[LMC05] J. Li, G. Mohan, and Kee Chaing Chua. Dynamic load balancing in ip-over-wdm optical burst switching networks. Computer Networks, 47(3):393.408, 2005.

[LY06] J. Li and K. L. Yeung. Burst cloning with load balancing. In Proceedings of Optical Fiber Communication Conference (OFC 2006), Anaheim, CL (USA), March 2006.

[LYH+06] K. Long, X. Yang, S. Huang, Q. Chen, and R. Wang. Adaptive parameter deflection routing to resolve contentions in obs networks. In Proceedings of the 5th International Conference on Networking (Networking 2006), Coimbra, Portugal, May 2006.

[OA05] N. Ogino and N. Arahata. A decentralized optical bursts routing based on adaptive load splitting into pre-calculated multiple paths. IEICE Transactions on Communications, E88-B(12):4507.4516, 2005.

[OTYC05] L. Ou, X. Tan, H. Yao, and W. Cheng. A study on dynamic load balanced routing techniques in time-slotted optical burst switched networks. In Proceedings of the 3rd International Conference on Networking and Mobile Computing, (ICCNMC 2005), Zhangjiajie, China, August 2005.

[PMP07] J. M. Pedro, P. Monteiro, and J. J. O. Pires. Efficient multi-path routing for optical burst-switched networks. In Proceedings of Conference on Telecommunications (ConfTele), Peniche , Portugal, May 2007.

[TR05] J. Teng and G. N. Rouskas. Traffic engineering approach to path selection in optical burst switching networks. Journal of Optical Networking, 4(11):759.777, 2005.

[TVJ03] G.P. Thodime, V.M. Vokkarane, and J.P. Jue. Dynamic congestion-based load balanced routing in optical burst-switched networks. In Proceedings of IEEE Global Communications Conference (GLOBECOM 2003), San Francisco, CA (USA), November 2003.

[VJ02] V. Vokkarane and J. Jue. Prioritized routing and burst segmentation for qos in optical burst-switched networks. In Proceedings of Optical Fiber Communication Conference (OFC 2002), Anaheim, CA (USA), March 2002.

[WMA02] X. Wang, H. Morikawa, and T. Aoyama. Burst optical deflection routing protocol for wavelength routing wdm networks. Optical Networks Magazine, 3(6), November/December 2002.

[YR06a] L. Yang and G. N. Rouskas. Adaptive path selection in optical burst switched networks. IEEE/OSA Journal of Lightwave Technology, 24(8):3002.3011, August 2006.

[YR06b] Li Yang and George N. Rouskas. A framework for absolute qos guarantees in optical burst switched networks. In Proceedings of IEEE Broadnets 2006, San Jose, CA (USA), October 2006.

[ZLW+04] J. Zhang, H. J. Lee, S. Wang, X. Qiu, K. Zhu, Y. Huang, D. Datta, Y. C. Kim, and B.Mukherjee. Explicit routing for traffic engineering in labelled optical burst-switched wdm networks. In Proceedings of International Conference on Computational Science (ICCS 2004), Krakow, Poland, June 2004.

[ZVR+04] A. Zalesky, H.L. Vu, Z. Rosberg, E.W.M.Wong, and M. Zukerman. Modelling and performance evaluation of optical burst switched networks with deflection routing and wavelength reservation. In Proceedings of the 23rd Joint Conference of IEEE Computer and Communications Societies (INFOCOM 2004), Hong Kong, China, March 2004.

[ZWZ+04] J. Zhang, S. Wang, K. Zhu, D. Datta, Y.-C. Kim, and B. Mukherjee. Preplanned global rerouting for fault management in labeled optical burstswitched wdm networks. In Proceedings of Global Telecommunications Conference (GLOBECOM 2004), Dallas, TX (USA), December 2004.