Optical Core Networks Research in the e-Photon-ONe+ project

Franco Callegati, *Member, IEEE,*, Filippo Cugini, Paul Ghobril, Sebastian Gunreben, Víctor López, Barbara Martini, Pablo Pavón Mariño, János Szigeti, Marcell Perényi, Namik Sengezer, Dimitri Staessens, Massimo Tornatore

***Abstract*—This paper reports a summary of the joint activities about Optical Core Networks within the e-Photon-ONe+ project. The joint activities cover a number of issues related to Optical Core Networks spanning from performance evaluation to traffic engineering, resilience and control plane.**

Given the size of the e-Photon-ONe+ project, the works reported here provide a reasonable overview of the topics considered of interest by the European research community and support the idea of building joint research activities that can leverage on the expertise of different research groups.

***Index Terms*—Optical Networks, Traffic Engineering, Wavelength Routing, Protection, Restoration, Physical Impairment, Congestion Resolution, Optical Packet Switching, Service oriented networks, GMPLS.**

I. Introduction

EPHOTON/ONE was a *Network of Excellence (NoE)* on optical networks funded by the European Commission (EC) in the context of the 6th Framework Programme (FP6). At the time NoEs were so-called *new instruments*, having the primary goal of fostering the integration of the researchers and institutions in Europe [1], [2].

e-Photon-ONe was a large project, involving about 40 institutions and 500 researchers active in the broad field of optical networking. The project was originally funded for two years (2004-2005). It proved successfull and the funding was renewed for two more years (2006-2007) under the e-Photon-ONe+ name. Besides the scientific result the e-Photon-ONe consortium gained visibility and reputation worldwide. An example is the co-sponsorship (with COST and NSF) of the "US/EU Workshop on Key Issues and Grand Challenges in Optical Networking" held in June 2005, at the European Commission (EC) premises in Brussels. Objectives of the workshops were to determine future research directions in optical networking and explore methods to facilitate stronger research collaboration between US and EU researchers. [3]

The dimension and the wide technical scope of the NoE raised significant problems of management. A model based on departments was adopted. The concept of *Virtual Departments (VDs)* was defined as the container and prometer of activities aiming at achieving durable integration, such as promote *joint research activities (JAs)*, identify new *research topics*, create *common expertise and research methodology*, etc..

One of the VDs was devoted to optical core networks, titled *Virtual Department on Optical core networks and technologies (VD-C)*. This work reports a summary of the JAs developed by VD-C in the last two years (a reader interested to the previous period may find a summary of results in [4]), that are listed in Table I (the person compiling the contribution for this paper is highlighted).

The rest of the paper is organized as follows: Section II provides an overview of the reference network scenario. Then the following sections report the results achieved organized per topic: traffic engineering in Section III, network resilience in Section IV, optical packet switching in Section V and finally service oriented optical core networks in Section VI.

II. Reference Scenario

Optical Core Networks (OCNs) carry large amounts of aggregated traffic, incoming from legacy or optical local and metropolitan area networks. In the OCN architecture we envisage a data and a control plane. The Control Plane (CP) is responsible for all the networking functions, such as routing, resilience, management etc., while the data plane is responsible for user data flow forwarding and is mainly based on all optical switching technology. It is likely, at the current level of understanding, that both control and data planes of OCNs will be multi-faced. Therefore one of the most recent focus of research is on the problems arising when different technologies are interconnected.

For instance in the data plane alternatives for multiplexing and switching range over a wide set of alternatives:

- fiber switching: provides coarse bandwidth granularity, slow switching speed, limited flexibility in bandwidth allocation; suited for optical circuit switching (OCS);
- wavelength switching: provides a smaller switching granularity at the wavelength level (a fiber can carry up to

F. Callegati is with the D.E.I.S., University of Bologna, ITALY
F. Cugini and B. Martini are with CNIT, Pisa, ITALY
P. Ghobril is with France Telecom Research & Development, Lannion, FRANCE
S. Gunreben is with the Institute of Communication Networks and Computer Engineering at the University of Stuttgart, GERMANY
V. Lopez is with Universidad Autónoma de Madrid, SPAIN
N. Sengezer is with the Electrical and Engineering Department of Bilkent University, Ankara, TURKEY
Pablo Pavón Mariño is with the Department of Information Technologies and Communications of the Polytechnic University of Cartagena (UPCT), Cartagena, SPAIN
J. Szigeti and M. Perényi are with the HSNLab, Dept. Telecommunications and Media Informatics (TMIT), Budapest University of Technology & Economics (BME), HUNGARY
D. Staessens is with Dept. of Information Technology (INTEC) at Ghent University - IBBT, BELGIUM
M. Tornatore is with the D.E.I., Politecnico di Milano, ITALY

Activity Title	Contributing Author	Activity Collaborators
Traffic engineering		
Regular Reconfiguration of Light-Trees in Multi-layer Optical Networks	Marcell Perényi	P. Soproni, T. Cinkler
A Comparative Study of Single-layer and Multi-layer Traffic Engineering Strategies	Namik Sengezer	B. Puype, E. Karasan, M. Pickavet
Implementation and experimental verification of a multi-layer integrated routing scheme for traffic engineering	Sebastian Gunreben	F. Agraz, S. Spadaro, S. Gunreben
Dynamic optical circuit-switched transport networks	Paul Ghobril	E. Le Rouzic, H. Nakajima, R. Watza, J. Rzasa, W. Kabacinski, S. Hanczewski, S. Spadaro
Network Resilience		
p-Cycle Protection in Multi-Domain Environment	János Szigeti	R. Romeral, D. Larrabeiti, T. Cinkler
Survivability in Multidomain networks: applicability and signaling	Dimitri Staessens	R. Romeral
Effects of Outdated Information on Protected Routing in WDM Networks	Massimo Tornatore	A. Pattavina, R. Munoz, R. Casellas, R. Martinez
QoT-aware control plane	Filippo Cugini	C. Pinart, I. Martinez, E. Le Rouzic, M. J. Poirrier, P. Castoldi, L. Valcarenghi, N. Sambo, N. Andriolli, A. Giorgetti
Optical Packet Switching		
Key Parameters for Congestion Resolution in OPS/OBS Networks	Franco Callegati	W. Cerroni, L. H. Bonani, G. S. Pavani
Comparison of end-to-end packet ordering issues in synchronous and asynchronous OPS networks	Pablo Pavón Mariño	J. Veiga Gontán, J. García Haro, D. Careglio, M. Klinkowski, J. Solé-Pareta
Service Oriented OCNs		
Advanced connectivity service provisioning in GMPLS networks	Barbara Martini	V. Martini, F. Baroncelli, P. Castoldi, L. Rea, C. Zema, S. Pompei, F. Matera, L. Wosinska, H. Su, A. Valenti
Multilayer Switching Algorithm for an All-Optical Router	Víctor López	J. A. Hernández, J. Aracil, J. P. Fernández Palacios, O. González de Dios

TABLE I
AUTHORS OF SUMMARY AND PARTICIPANTS OF JOINT ACTIVITIES

hundreds of wavelengths), more flexibility as well as performance optimization exploiting wavelength conversion; again for OCS;

- sub-wavelength switching: typically makes use of multiplexing in the time domain; offers finer switching granularity, but requires faster switching speed, and is more complex; it opens a whole span of networking solutions, from optical time division multiplexing (OTDM), to optical burst switching (OBS), to optical packet switching (OPS).

The control plane exploits existing or new protocols. Because of the large traffic flows carried by the OCN the critical issues are:

- fast and flexible resource allocation scheme to answer the user needs;
- reliability and network survivability;
- traffic engineering and contention resolution.

These are mature research topics, here investigated with the pragmatic approach of looking for solutions to engineering problems. The focus is on multi-layer (ML) network architectures, to understand up to which extent the integration between layers may improve the overall network performance.

At the same time new topics are emerging as envisaging future mass market bandwidth-greedy applications, such as Grid Computing and Service Delivery Platform running on world-wide scale. They will require "on demand" network services with requirements in terms of bandwidth, availability, end-to-end delay, etc. Service architectures have been defined by the principal standardization bodies, such as the IP Multimedia Subsystem (IMS) by 3GPP [6] and the Next Generation Network (NGN) by ITU-T [7], and application-to-network mediation layer have been proposed in order to support a wide range of applications with QoS-enabled connectivity requirements.

Unfortunately none of them foresee any exploitation of the GMPLS CP capabilities. This opens a whole set of new problems that are particularly important for optical networking, as it is envisioned to be the enabling transport technology for these new networking scenarios.

All these topics have been addressed within VD-C, with the sole exception of Optical Burst Switching. A workpackage was devoted to this topic in the e-Photon/ONe+ project that worked in close contact and collaboration with VD-C but formally reported results separately. A summary of the activities on OBS can be found for instance in [5].

III. TRAFFIC ENGINEERING

Traffic engineering describes the process of analysing and (re)structuring traffic flows in a network in order to optimize the overall network performance with respect to pre-defined metrics. Four joint activites addressed traffic engineering in e-Photon-ONe+ , mainly focusing on ML issues.

The first activity deals with multicast in an optical ML network and shows the benefit of reconfiguring the mutlicast trees periodically to cope with traffic variations.

The second activity compares three different ML routing strategies in an optical ML network. It shows the benefit of traffic flow reconfiguration in comparison to a statically configured network.

The third activity focuses on the realization issue of a TE algorithm for ML optical networks which is implemented in an optical test-bed. The focus here is to test and match the assumptions of the algorithm with the test-bed limitations.

The last activity has a slightly different focus. The goal is to propose a comparison framework to assess the advantages/drawbacks of a more dynamic switching technology (for instance OBS) with respect to a more coarse wavelength switching, in a network scenarios where different switching granularities are available.

A. Regular Reconfiguration of Light-Trees in Multi-layer Optical Networks

In recent years the traffic in transport networks keeps growing due to multicast (MC) applications, including very important broadband services such as digital media broadcasting, VoD streaming, distance learning or virtual private LAN services [9].

In spite of its benefits in terms of bandwidth (BW) savings, today most commercial ISPs do not implement multicasting in the lower layer. As a consequence, a huge amount of bandwidth is wasted due to MC delivery based on application-layer multicasting (ALM) i.e. unicast-based distribution.

Nonetheless the MC service is an essential feature, which is currently not present in the core, but it should be, since it is the key to the scalable implementation of the triple-play concept: TV channels are usually multicasted from a content distributor to local caches/relays near the end users.

In general, MC should be implemented in the lowest layers of the network to minimize waste of BW [10], however, the physical layer may pose limits on the number of supported connections. This is the case in WL-routed optical networks, where limits are set by:

- the number of available WLs;
- the amount of multipoint units in optical nodes;
- the nodesŠ fan-out and the optical power budget.

Therefore optimizing light tree construction is quite a relevant challenge in next generation MC-capable optical networks [16]. In our approach the problem is addressed with reference to dynamic multicast trees, where the member tree leaves are continually changing. New destination nodes may log in to the tree to receive the content, while other nodes may leave the tree and return at a later time. This corresponds to a scenario where IP membership drives optical tree set up.

The continuous changing of tree leaves causes the degradation of the MC tree in terms of switching and transmission resources as the tree diverges from the optimum. Regular reconfiguration can solve this degradation issue and take advantage of spared resources. Unfortunately reconfiguration has also drawbacks: computation effort, short disruption in the data transmission flow, additional signaling overhead. The activity analyses the cost/benefit trade-off of reconfiguration of the MC tree of a two-layer network, where the upper, electronic layer is packet switching capable, while the lower, optical layer is WL (space) switching capable.[1]

The traffic consists of dynamic, multicast delivery demands. In a real scenario, the tree would be due to the aggregation of multiple multicast sessions or it could be given by a selected set of individual ultra-broadband multicast sessions. An example can be a digital media distribution service, where the audience is varying in time. New customers appear, who subscribe to the content, and other customers with expired subscription leave the network. In this case a customer may be a residential user as well as a local provider (e.g., a local cable-TV provider).

Since the exact solution of the design of the MC tree is NP-complete, heuristic algorithms are defined and applied to route and maintain MC trees in the network, comparing their costs and performances.

- *Accumulative shortest path* (ASP) algorithm simply connects newly arriving endpoints to the MC tree (applying Dijsktra's algorithm) and clears branches leading to departed endpoints.
- *Minimal Path Heuristic* (MPH, [11]) transforms the original WL graph into a virtual graph and applies Prim's algorithm [12] to form a minimum cost spanning tree.
- *Tree routing* algorithm is similar to MPH, except that it operates in the original WL graph, not in a derived *upper layer* virtual graph.

The performance of the heuristics are compared to the optimal solution obtained by an Integer Linear Programming (ILP) formulation [10]. Although ILP is time consuming, it provides a close-to-optimal solution, but it has the drawback that, in presence of a traffic variation, the new solution may be very different compared to the original one, requiring reconfiguration of existing paths. The heuristics, on the other hand, have in common that they can route new demands without interrupting current active demands.

The number of WLs was 8 on each physical link. We assumed that every incoming lightpath (LP) in a network node can be routed up to the electronic layer to perform (tree-) branching, WL conversion or grooming (multiplexing of multiple trees with sub-wavelength bandwidths). However, O/E (optical to electronic) and E/O conversion was assumed to be two times more expensive than switching a LP purely in the optical layer, i.e. the usage of a WL, optical switching of a LP, O/E, E/O conversion had all a cost of 1 unit (resulting an O/E/O conversion to cost 2 units).

The heuristics were simulated on the COST 266 European reference network [13] using the CPLEX optimizer [14] to solve ILP instances.

For space reasons we present results for the ASP algorithm compared with the ILP solution. Other results and details are available in [15]. Reconfiguration of dynamic light-trees is clearly beneficial for transport network operators (see Figure 1 as an illustration). Lots of network resources, e.g. O/E converter units and WL capacity can be spared by restoring the optimal topology of the tree.

[1]The electronic layer can perform traffic grooming. The two layers are assumed to be interconnected according to the peer model, i.e. the control plane has information about both layers and both layers are involved in routing.

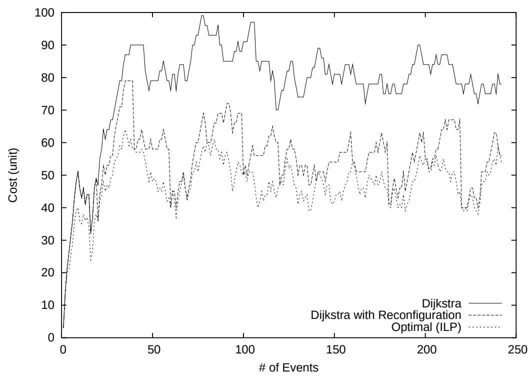


Fig. 1. The cost of routing as a function of elapsed events for Dijsktra algorithm with and without reconfiguration compared with optimal ILP solution.

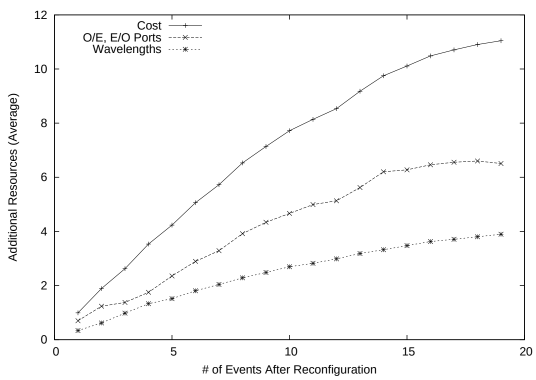


Fig. 2. Average additional cost of routing after reconfiguration.

Figure 2 shows details of the average additional cost split in routing cost, number of O/E, E/O conversion ports and number of WLs as a function of the number of elapsed events after reconfiguration. The arrival of a new demand or the departure of an existing one is called an *event*. The more the events after reconfiguration the more the MC tree diverges from the optimal and the additional cost increases.

We showed that the reconfiguration period has an optimal length (Figure 3), if we take into account the negative aspects of reconfiguration as a penalty cost in addition to the network cost. Reconfiguration seems to be especially useful if grooming is not possible. Still, a number of technical challenges must be addressed to make reconfiguration practical, like the seamless switch-over of traffic from the old to the new tree.

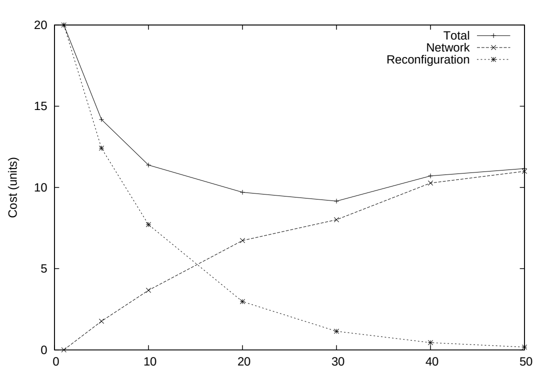


Fig. 3. Total, network and reconfiguration costs as a function of the reconfiguration period. The minimum of the total cost curve suggests the optimal length of the reconfiguration period.

B. A Comparative Study of Single-layer and Multi-layer Traffic Engineering Strategies

Many Traffic Engineering (TE) routing strategies have been proposed. In this activity the expertise of various research groups was put together with the aim to compare the strategies in a Multi Protocol Label Switching (MPLS)/Wavelength Division Multiplexing (WDM) network. The assumption is to have a label switched path (LSP) for each source/destination node pair, which carries the total traffic between these nodes. The routing and TE algorithm determines the path of the LSPs and matches it with lightpaths in the WDM layer. To this end three algorithms were compared.

a) Fixed-Topology TE (FTE): Provides TE in the electronic layer only. It assumes the knowledge about the traffic expectations on an hourly basis and designs the virtual topology accordingly [2]. This strategy has two subsequent phases.

- *Offline* designs the topology applying a heuristic search algorithm based on the tabu search metaheuristic. The algorithm starts from a randomly generated topology with the given number of lightpaths and gradually improves the topology by changing the routing of the lightpaths.
- *Online* At the beginning of the online phase, the algorithm calculates a given number of shortest paths for each source-destination pair. When the bandwidth requirement of an LSP changes, the best path is chosen. A dynamic path cost function determines the best path [17] and the LSP is (re)routed. The cost function favors shortest paths when the loads on the links are small and paths with higher available capacity when the links are heavily loaded.

b) TE with Periodic Topology Reconfigurations (PTR): Performs topology generation and routing together by using a load based link cost function. The topology changes dynamically and, after each topology update, LSPs can be routed also on the newly established lightpaths.

The applied cost function considers unopened lightpaths and aims to keep the load between certain thresholds. Below the lower threshold, the function has a high cost region; this allows grooming the traffic on fewer lightpaths. Above the upper threshold, the cost increases rapidly to prevent overloading of the links. The cost in the high load region dominates because preventing blockings is prioritized over the grooming objectives. After the routing is done, the lightpaths that are not used are torn down [18].

At the beginning of each hour, this strategy generates a new topology with the new routes of the LSPs, and performs the required reconfigurations. An additional cost is defined for opening new lightpaths to reduce the number of reconfigurations between consecutive topology updates.

[2]The operator may obtain this information from the network management plane

c) TE with Periodic Topology Reconfigurations and Route Updates (PTR-RU): Combines the mechanisms of FTE and PTR. The WDM layer is reconfigured periodically as in PTR, hence the virtual topology is adapted to the hourly traffic changes. In between the hourly reconfigurations, the LSPs routes are updated dynamically to cope with the shorter term traffic variations using the same dynamic cost function and the rerouting mechanism as in FTE.

An event driven simulation tool was developed to evaluate the performances of the TE strategies under common traffic demands.

A 24 hour traffic pattern, the details of which are described in [17], is used to generate the hourly expected traffic values. Since one LSP is set between any source-destination pair, the traffic changes are in terms of bandwidth requirements of the LSPs.

The actual traffic between a node pair deviates from its expected value according to a zero mean Gaussian noise component that is added to the expected bandwidth. The standard deviation of the added noise is linearly dependent on the instantaneous value of the expected traffic. The ratio between deviation and expected value is called *traffic unpredictability*.

Two scenarios are investigated.

- *Scenario 1*. The maximum bandwidth expectation per hour is known for each LSP. Algorithms FTE, PTR and PTR-RU design the virtual topology using this information.
- *Scenario 2* No knowledge of the traffic demands is available in advance. The algorithms design the virtual topology based on the instantaneous bandwidth information. This excludes the FTE strategy as it requires the knowledge of the traffic pattern in advance.

Figure 4 shows the blocking probability (as percentage of the total required bandwidth that is lost) of the TE strategies in the first scenario for different values of the traffic unpredictability. PTR and PTR-TU set up an average of 20.5 lightpaths for each topology within 24 hours. For a fair comparison, we designed a topology with 21 lightpaths and 23 lightpaths.

According to the results, the TE routing strategies performing topology reconfigurations (PTR and PTR-RU) outperform FTE and, for lower values of the traffic unpredictability, PTR is also better than PTR-RU. This is due to the fact that, in that region, the topologies and the routes are calculated using more accurate traffic information, and LSP route updates are not needed. As traffic unpredictability increases above the value of 0.17, PTR-RU starts to give the best performance, which implies that when the traffic unpredictability is high, dynamic route updates improves the network's performance.

The blocking performances of PTR and PTR-RU in the second scenario are presented in Figure 5. When no statistical information is available and the TE actions are performed using the instantaneous traffic information, the benefits of dynamic route updates come forth and PTR-RU displays a better performance than PTR. This is in consistency with the results of the first scenario.

Comparing the two scenarios, it is possible to see that making use of the statistical information about the traffic

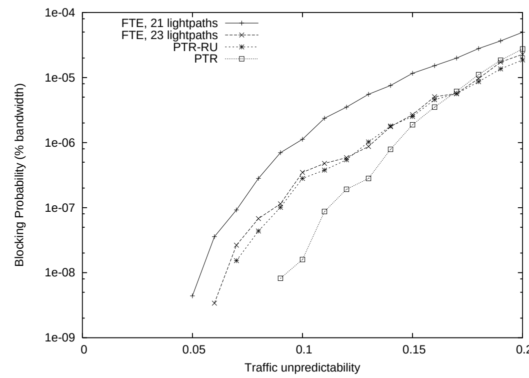


Fig. 4. Comparison of TE routing strategies: Scenario 1

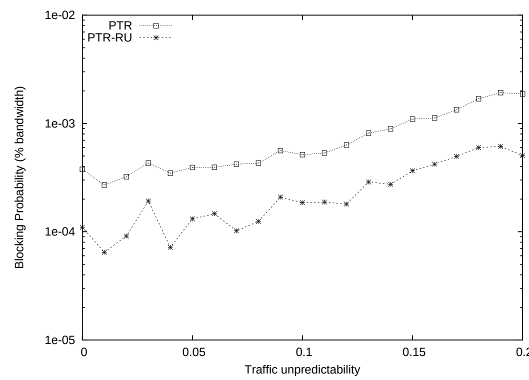


Fig. 5. Comparison of TE routing strategies: Scenario 2

(scenario 1) brings to far better performance than using only instantaneous traffic information (scenario 2) as intuition would suggest.

C. Implementation and experimental verification of a multi-layer integrated routing scheme for traffic engineering

This research activity faces a multi-layer (ML) optical network. It focuses a ML routing algorithm and its real life evaluation in the optical test-bed CARISMA [21]. The scenario includes an optical transport layer (WDM), providing coarse transport capacity and an electronic layer (SDH) providing a finer transport capacity. It assumes a common control plane based on Generalized Multi Protocol Label Switching (GMPLS).

The applied Weighted Integrated Routing (WIR) scheme has already been proposed in [19], [20]. It assumes a co-location of the electrical and optical nodes and operates in a two-step approach.

1) The shortest path algorithm searches for a path on the optical layer only. The result consists of source and destination nodes as well as the intermediate nodes along this path.
2) The algorithm choses a set of nodes including source and destination nodes and a subset of the intermediate nodes. With this set of nodes, the algorithm checks the feasibility of a path taking into account both layers. Step 2) is repeated for every subset of the intermediate nodes.

All path alternatives form a set of possible solutions to the initial routing problem. The criteria of the optimal path depend on the cost of the path. Cost metrics may include the number of wavelength conversions, electrical or optical hops or the link occupancy.

The test-bed shows a dual ring structure with two fibers and three optical nodes with reconfigurable optical add and drop multiplexing capacity. The control plane comprises up to 20 Linux-based routers, acting as optical connection controllers [21]. These routers implement the GMPLS protocol set, including RSVP-TE for connection establishment, OSPF-TE for resource state dissemination, and the Link Managemet Protocol (LMP) for resource management and discovery.

The CARISMA test-bed is essentially a wavelength-routed optical network. It allocates resources for end-to-end connections in whole wavelength granularity. For fine grain connections the Forwarding Adjacency (FA) concept permits the aggregation of higher-order LSPs into these lower-order ones. The routing protocol advertises these lower-order LSPs as FA-LSP. Any other node may use this FA-LSP for path computation, nesting lower-order LSPs into these already created FA-LSPs [22]. Thus, the FA concept enables ML routing support in the control plane of a pure optical network.

The implementation of the routing algorithm in the test-bed faced some restrictions imposed by the test-bed as well as by the assumptions of the routing algorithm.

- *Test-bed limitations:* The algortihm assumes a ML SDH/WDM network. As the test-bed implements the WDM layer only, we distinguish both layers only in the control plane using OSPF.
- *Wavelength continuity:* Wavelength continuity requires the knowledge of the wavelength in the control plane. This knowledge violates the concept of GMPLS. The test-bed implementation includes this information using a non-standard OSPF-TE extension.
- *Number of optical hops:* The FA does not provide any information of the underlying optical links. Consequently, the FA hides the number of optical nodes and the path metric of the number of optical hops is not applicable.
- *Cost metrics:* Technical limitations and the restrictions of the control plane limit the possible cost metrics of the proposed algorithm. We implemented the proposed algorithm taking into account the following criteria: number of electrical hops, number of optical hops (if available).

The FA support required changes at three main components of the control plane: the Link Resource Manager, the Routing Controller and the Connection Controller .

- *Link Resource Manager:* The link resource manager module implements the LMP and keeps the information about the physical resources of the node. Moreover, we developed operations for supporting FA establishment, tear-down and bandwidth reservation/release.
- *Routing Controller:* The Routing Controller implements the route calculation and the OSPF-TE mechanism to store and announce link state information. In this context, the FA support required a new set of structures and processes needed for route calculation.

Fig. 6. Network topology

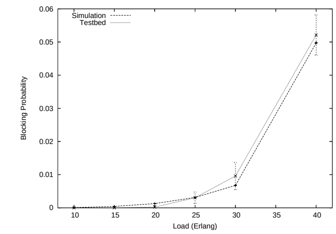


Fig. 7. Comparison of the connection blocking probability

- *Connection Controller:* The Connection Controller coordinates the LSP and FA establishment and management. For these tasks we developed and implemented several changes and add-ons, mainly regarding RSVP-TE in the CARISMA control plane.

The implemented algorithm was tested in the optical test-bed and its performance has been evaluated also by simulations. Figure 7 shows the simple network topology of the comparison. Each link represents two uni-directional connections each carrying two different wavelengths. The capability of one wavelength is STM-16. The control plane advertises this capacity in granularities of STM-4. The traffic demand in the network is uniformly distributed. One demand requests a connection of STM-4 granularity.

The measurements in the test-bed were taken in five batch runs, each including 5000 connection requests. The simulation considered more than a million connection requests. Figure 7 depicts the comparison on the connection blocking probability for both set-ups, the test-bed and the simulation. Over the considered load range, measurement and simulation fit well and the testbed implementation and simulation model correspond within the confidence intervals.

The outcome of this activitiy showed that the ML routing algorithm is in principle realizable in a real network. Although some limitations were either solved by proprietary protocol extensions or skipped in the implementation. In particular the implementation required the extention of the GMPLS control plane. The performance of the implementation were tested against simulations showing a very good agreement.

D. Dynamic Optical Circuit-Switched Transport Network

This activity aims at investigating the effectiveness to serve a dynamic traffic while switching on different granularities in the optical layer. The idea is to consider and compare Optical circuit switched (OCS) networks switching on wavelength (WDM channel) granularity with multi-granular optical circuit switching networks that switch on different granularities (e.g. waveband and wavelength).

Switching on waveband (group of contiguous wavelengths) granularity allows reducing the number of ports of the optical cross-connect (OXC). Multi-granular OXCs (MG-OXCs) combine both wavelength cross-connects (WXCs) and waveband cross-connects (BXCs). MG-OXCs are classified as hierarchical or non-hierarchical ones as shown in Figure 8. In a hierarchical MG-OXC input/output wavebands are applied to a BXC and can likely pass through the WXC. However, in

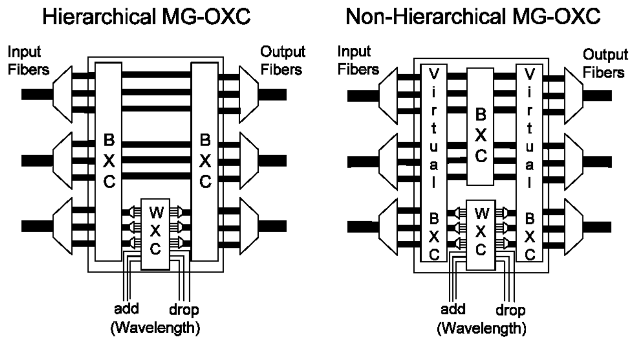


Fig. 8. Multi-granular optical cross-connects

a non-hierarchical MG-OXC Virtual BXCs represent the a priori configuration of input/output wavebands that are, once for all (by hardware), either applied to a BXC or demultiplexed and applied to a WXC. Although hierarchical MG-OXCs need more input/output ports than non-hierarchical MG-OXCs, they still can reduce the number of ports while enhancing the lack of flexibility resulting from bulk switching and hence balancing scalability and flexibility.

Introducing multi-granularity adds to routing and wavelength assignment (RWA) the problem of controlling the MG-OXC that consists in choosing at which switching granularity a new dynamically arriving demand must be served at each intermediate node [23]. We intend to study the blocking probability when proposing different controlling schemes and different architectures having compared complexity reduction (i.e. reduction in the number of input/output ports). We can then, for instance, compare a network using multi-granular but non-hierarchical cross-connects, where the lack of flexibility is balanced by alternating the use of waveband and wavelength cross-connects all over each path, to a network with hierarchical cross-connects.

In this activity was developed a framework to analytically model dynamic waveband switching in a multi-granular optical network. The solution consists in modelling each potential carrier of waveband tunnels independently by a Markov chain while modulating the rate of critical transitions, i.e. reserving a new waveband tunnel, by the waveband setup availability computed from the solution of other potential carriers. An iterative procedure is repeated to obtain a consistent numerical solution all over the network [24]. Using the analytical model [24], the dynamic waveband switching problem is reduced to a classical circuit switching problem (i.e. computing blocking probability in a loss network) where the main difficulty is the link load correlation as confirmed by the obtained results, for instance the example in Figure 9. This difficulty is shown to be surmountable in [25] using the object independence assumption and the model is currently under extension in this direction.

IV. Network Resilience

Network resilience in OCNs has been widely investigated in recent years, nonetheless, some of the recent progresses in optical networking pose new challenges.

Future optical networks will interconnect several optical domains with different switching granularities, where each domain is controlled by an independent and autonomous service provider. In this case the end-users' expectation is that they get the same or near the same reliability for long, inter-domain connections as for the short, intra-domain connections. This is set back by several difficulties, the most obvious are:

- physically longer connections may fail with higher probability and the desired grade of availability is not met by using merely the traditional protection schemes, that do not take this into account;
- the different network domains are typically run by different operators that are reluctant to share information with their competitors;
- broadcasting every minor change of the network state would unnecessarily overload the network with state messages causing scalability issues.

Presently responsibility of rerouting inter-domain resources is left to the BGP protocol that has been rather successful in the current Internet but has a few drawbacks in this new perspective. For scalability reasons, the internal topology of the domains is not known and the amount of paths offered between two given domains is limited to one. Moreover BGP is slow in establishing a new route and can not be used for real time protection. Therefore "ad hoc" protection strategies and algorithm must be designed.

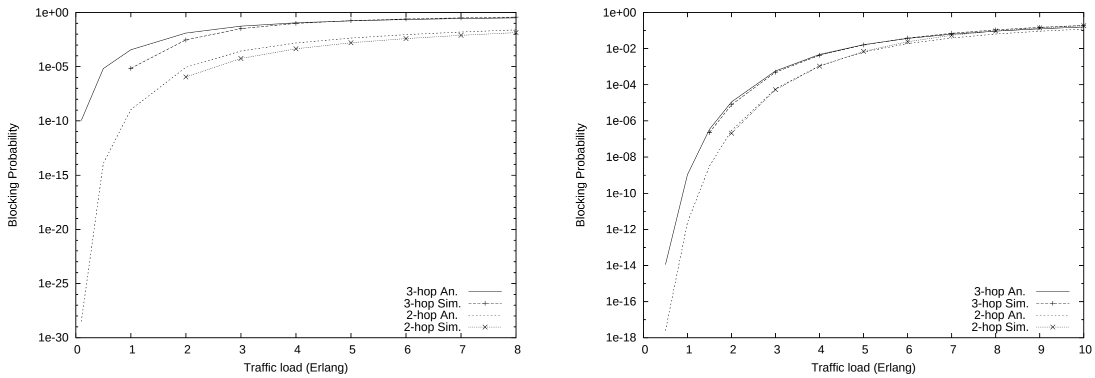
Furthermore, the emerging applications call for an intelligent optical network control plane which is aware and can compute end-to-end paths based on several domain-specific and technology-specific parameters. The GMPLS protocol suite provides control mechanism to dynamically setup and release connections in WDM networks. Signaling is essential in this kind of control-plane enabled networks, but most of the studies on dynamic traffic routing tend to neglect the effect of control plane on routing performance. Similarly while propagating through the optical network, a signal may degrade in quality as it encounters physical impairments, and this may in turn make the Bit Error Rate (BER) at the destination unacceptably high. Thus, Quality of Transmission (QoT) parameters have to be introduced in the GMPLS protocol suite to provide reliable lightpath provisioning.

In the four activities described in the following, these issues are investigated:

- the first two contributions address protection in a multi-domain environment;
- the last two activities address the interactions among the control-plane capabilities and the quality of protection achievable.

A. *p-Cycle Protection in Multi-Domain Environment*

Providing resilient inter-domain connections in multi-domain optical GMPLS networks is a challenge. On the one hand, the integration of different GMPLS domains to run traffic engineering operations requires the development of a framework for inter-domain routing and control of connections, while keeping the internal structure and available resources of the domains undisclosed to the other operators. On



(a) (b)

Fig. 9. Blocking probability for connection chosen arbitrarly (a) and avoiding link load lorrelation (b): analytical and simulation results

Fig. 10. Considered network architecture

Fig. 11. Considered network architecture

the other hand, the definition of mechanisms to take advantage of such automatically switched inter-domain connectivity are still an open issue. This paper focuses on the analysis of applicability of one of these mechanisms to the new scenario: *p*-cycle-based protection.

The proposed solution is based on the decomposition of the multi-domain resilience problem into two sub-problems, namely the higher level inter-domain protection and the lower level intra-domain protection. Building a *p*-cycle at the higher level is accomplished by certain tasks at the lower level, including straddling link connection, capacity allocation and path selection.

In [**?**] we present several methods to realize inter-domain *p*-cycle protection at both levels and we evaluate their performance in terms of availability and spent resources. A discussion on a proposal of implementation of signalling based on extensions of existing protocols such as RSVP-TE and the PCE architecture illustrates the practical viability of the approach.

B. Survivability in Multidomain networks: applicability and signaling

The implementation of fast mechanisms for protection and restoration usually requires pre-computed backup paths. This is not easily applicable to the inter-domain case because operators prefer not to share topology information and wavelength usage in their domains. A domain is a set of routers under a single technical administration, using some interior gateway protocol(s) (IGPs) and common metrics to route packets within the domain, and using an exterior gateway protocol to route packets to other domains.

This work starts from the assumption that every domain has an optical transport layer supporting IP traffic and traffic engineering capabilities (e.g. G/MPLS) in each of its network layers. The network scenario is shown in Figure 11.

The proposed multidomain solution is based on a per-domain computation (using the Path Computation Element framework), where each domain has two ingress points and two egress points. The solution consists of a survivable upper layer connection over multiple domains. In the IP/MPLS layer, this is implemented using a path-protected connection (working connection and backup connection). The implementation of the connections in the optical layer in each domain again uses path protection for the working connection (called the Primary Path P_x and its backup path PB_x and no protection for the backup connection (B_x.

The research performed has two major results. First, we show that our proposed solution can be computed in any 2-connected topology [48]. This is important, because it shows that it is possible to set up this solution in a simple sequential way, and we effectively avoid trap topologies.

Second, we have proposed a signaling scheme using RSVP-TE, where we have proposed a new routing object, Gateway Specification Routing Object (GSRO) to allow signaling the backup gateways while setting up the working connection [49].

C. Effects of Outdated Information on Protected Routing in WDM Networks

The effects of the control plane constraints on the routing of dedicated and shared protected connections in an all-optical network has never been investigated so far. In this JA, we have analyzed the general effects of outdated information in a control-plane enabled optical network. on protected routing performance.

In distributed-GMPLS networks, each node builds a network image in its routing performer to identify the best path to route a connection (source routing). As an effect of control delays, source nodes databases are not updated and routing may be not optimized, since it is carried on according to an outdated image of the network. The information to build this image depends on various factors, e.g., which protection is applied and if wavelength conversion is enabled. In short, the information that it is usually distributed allows to discover the network topology and evaluate the amount of free capacity available on each link. Moreover, in case of shared protection, the state of shareable backup resources has to be disseminated and ad-hoc routing protocol (such as OSPF-TE) extensions are needed, as in [31] where an actual realization of SPP over the CTTC test-bed has been demonstrated.

The contributions to the control delay can be resumed as: information propagation delay, set-up delay, switching delay, processing delay, periodical database update. To generically include all these contributes, we propose a simple delay model based on the following assumptions:

1) Constant Control Delay: as a sum of many delay contributions, we consider the a fixed control delay τ;
2) Negligible Set-Up Time: once routing has been identified, provisioning occurs without no delays;
3) Identical network vision: we assume that all the nodes share the same network vision, referred to the instant $t - \tau$.

Our simplified approach allows to quantify the effect of a wide range of control delay values, independent of the specific routing/control algorithms and parameters (update frequency, amount of information), while still allowing to draw general considerations on the effect of outdated information.

In summary, in the JA we have analyzed the control-delay effects on routing performance [32] using realistic case-study network topologies in a dynamic network environment. We have considered the unprotected case, the DPP and the SPP case, in both the cases of presence (VWP) and absence (WP) of wavelength conversion. The considered metric is the blocking probability (BP). The influence of the various BP components and of the dynamism of the traffic on routing performance are discussed. Finally, simulative results are compared to those obtained in the CTTC testbed.

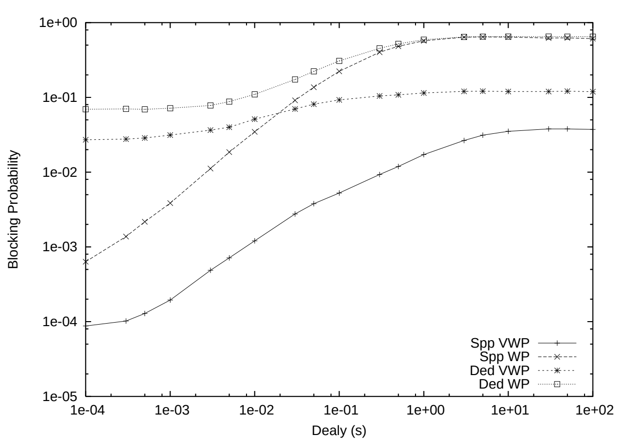


Fig. 12. Total BP for DPP and SPP routing in VWP and WP case

Figure IV-C shows the BP for the DPP and SPP case (under WP and VWP assumption). Curves are plot for increasing value of delay, considering 100 arrivals per second (which leads to a network load of around 0.55). The delay on the x-axis is a relative measure, expressed as the ratio between the absolute delay D and the average holding time HT. All these curves can be described by three well distinct phases:

- I phase - Not influential delay: in this first phase the BP is constant and it is not influenced by the delay.
- II phase - Linear increase: outdated information starts affecting the quality of source-routing, causing a significant and linear increase of the BP
- III phase - saturation: the BP is not affected by increases in the control delay, since the network image at the source node is now uncorrelated to the actual network state.

The blocking values are higher in the WP case. As a matter of fact, the provisioning of a connection over a given path in the VWP case fails only in the case all the channels on a link are saturated, while in the WP case only the chosen channel has to be free to allow a successful provisioning of the connection over the chosen path.

D. QoT-aware Control Plane

Making the control plane QoT-aware poses two main challenges:

- define a QoT model to estimate the expected QoT and thus the lightpath feasibility;
- implement the protocol extensions that allows the application of the QoT model in the distributed control plane [33] [**?**].

In the literature are proposed many QoT models that capture certain physical impairments, such as the Polarization-Mode Dispersion (PMD), the Amplified Spontaneous Emission (ASE) noise, cross-talk or non-linear effects. Although some models are based on impairments (e.g., Optical Signal to Noise Ratio (OSNR)), most of them are based on constraints related to the impairments (e.g., Q-factor) [34]. The more accurate the QoT model, the better is the impairment-aware connection provisioning. Therefore, all the QoT parameters should be ideally obtained from the optical transport network in a rapid way to allow real-time QoT-aware connection provisioning. However, in practice this is not feasible due to the following main reasons:

- Cost: Capex and Opex limit the amount of monitoring points in the network.
- Difficulty of obtaining reliable QoT parameters: some impairments can not be measured, but only estimated or accounted for by increasing system margins to cope with their effects (e.g., non-linear parameters)
- QoT estimation: To evaluate the lightpath feasibility. It can be done at the receiving end of a candidate lightpath through a single parameter, e.g. the Q factor, or using an analytical model of the physical layer.

To set up transparent connections (T, i.e. lightpaths) or non-fully transparent connections (NoT, i.e. lightpaths with some intermediate nodes performing opto-electronic regeneration), it is required to extend the GMPLS protocol suite to include information related to QoT and to the presence of shared-per-node regenerators. Four approaches can be considered:

- Routing-based Approach (RA). In RA, additional extensions are introduced in the routing protocol to flood QoT parameters and the availability of shared-per-node regenerators. The main advantage of RA is that it is fully distributed and potentially able to compute effective TE solutions (i.e., engineering lightpath provisioning from

both QoS and QoT viewpoints). However RA may heavily suffer from QoT information inconsistency, scalability and convergence problems particularly in case of frequent link parameter changes or upon failure occurrence.
- Signaling Approach (SA) [35]. In SA, no extensions are introduced in the routing protocol which elaborates routes ignoring QoT. Then, SA performs the dynamic estimation of the QoT during the signaling phase by collecting QoT parameters from intermediate nodes. At the destination node, if the accumulated parameters are within an acceptable range, the lightpath set up request is accepted. Otherwise the lightpath request is rejected and further set up attempts following possibly link-disjoint routes are triggered. The main advantage of SA is that it avoids the flooding of QoT parameters and regenerator availability and preserves control plane scalability. However, it may increase the amount of control plane packets and delay the lightpath establishment process.
- PCE-based Approach [35]. The PCE performs path computations considering both QoT and QoS constraints. QoT and regenerator information are collected from the Management System or through a performance monitoring system. In this approach, neither the routing nor the signaling protocols are enhanced with further extensions. Thus control plane stability is preserved. The centralized nature of the PCE allows to achieve optimal TE solutions however it may introduce scalability issues, delay in setting up the lightpath and PCE failures may affect the overall network performance.
- Probe-based Approach [34]. Probe traffic is sent over the lightpath, and if the BER level measured at the receiver by the FEC is sufficient, then the lightpath is considered suitable and client traffic can start to be transmitted. This approach allows to effectively evaluate the lightpath QoT however it may delay the lightpath set up time and unnecessarily reserve network resource in case of unacceptable lightpath QoT.

The performance of the Signaling Approach is evaluated by means of a custom built C++ event-driven network simulator. A Pan-European topology with 17 nodes and 32 links is considered [36]. Each link carries 40 wavelengths. Each network node is equipped with N shared-per-node regenerators. Connection requests are dynamically generated with uniform distribution among all node pairs. Network load is kept limited in order to experience connection blocking due mainly to unacceptable QoT or lack of regenerators.

Figure 13 shows the blocking probability (BP) of T connections within k set up attempts as a function of the generated unidirectional connection requests (k equal to 1, 2 and 3 in Fig. 13a, b and c respectively). Fig. 13 also shows the BP of NoT connections due to lack of intermediate regenerators.

Figure 13 refers to the presence of $N = 1$ and $N = 2$ shared-per-node regenerators. Results show that, by exploiting successive set up attempts, the overall BP of transparent connections decreases. In addition, also the non-fully transparent connection blocking probability due to the lack of regenerators decreases with the increase of the number k of transparent connection set up attempts. Indeed, a higher number of explored routes and nodes allows to save regenerators and collect information that improves the likelihood of establishing NoT connections. However, while the increase from $k = 1$ to $k = 2$ leads to significant BP reductions, increasing k from 2 to 3 provides negligible reductions. Thus, just two set up attempts before resorting to regenerators guarantee the best performance.

V. Contention resolution strategies in Optical Packet Switching

Contention resolution is a critical issue in OPS. The focus of the research developed in e-Photon-ONe+ was on solving contention by scheduling algorithms exploiting wavelength conversion and delay lines in a combined way.

The former activity studies the issue of choosing an optimal combination of the numerous set of parameters that play a role in determining the performance of a scheduling algorithm. A new comparison metric is defined that allows an a complexity/performance trade-off of different alternatives under a new perspective. The latter activity explores the issue of designing scheduling algorithms that are able to maintain the packet sequence and therefore may be suitable for QoS sensitive traffic.

Both activities refer to an OPS switching system able to emulate output queuing with delay lines and converters shared per output port [37]. The N input/output ports are equipped with F fibers each, carrying W wavelengths and with B delay lines. As a consequence the number of input/output channels per node is $N \cdot F \cdot W$. As usually in these studies the focus is on the packet loss probability (PLP) given that the delay experienced in the nodes should be negligible or at most comparable with the propagation delay.

A. Key Parameters for Congestion Resolution in Optical Burst/Packet Switching

This activity focused on contention resolution algorithm in OPS switches exploiting the wavelength and time domains in a coordinated manner. This is called *Channel and Delay Selection* (CDS) scheduling, requiring CDS algorithms. Work on this topic is not new but a general framework to compare different scheduling alternatives was not yet presented in the literature. In this activity we defined the concept of scheduling space and proposed to use it as a basis of comparison to assess which are the best complxity/performance trade-offs.This activity propose to compare CDS algorithms on the basis of the dimensions of their scheduling space.

If $C \leq F \cdot W$ is the number of channels available to transmit a packet on a given output port, the *scheduling space* is the *set* of possible scheduling choices available for the CDS algorithm

$$s(i, j, k) = (t_i, j, k) \in \mathcal{S}$$

where i gives the time the packet will be transmitted (delayed if necessary) that will be $t_i = t_0 + d_i$, j gives the wavelength and k the fiber for transmission. The number of elements (i.e. the cardinality) in $\mathcal{S}$ is $|\mathcal{S}| = B \cdot C$.

In an ideal switching matrix with full range wavelength conversion (FWC), where information units can be freely

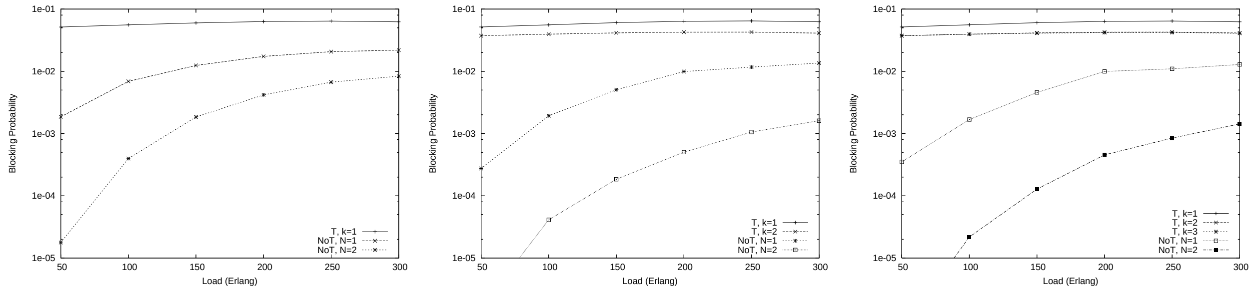


Fig. 13. Blocking probability within k (transparent) set up attempts and after the successive set up attempt exploiting N regenerators per node. a) $k = 1$; b) $k = 2$; c) $k = 3$.

switched in space and wavelength (i.e. $C = F \cdot W$), the number of elements (i.e. the cardinality) in $\mathcal{S}$ is $|\mathcal{S}_{\mathcal{FWC}}| = B \cdot W \cdot F$. However in real systems $|\mathcal{S}| leq |\mathcal{S}_{\mathcal{FWC}}|$ as a result of hardware or software limitations. For instance if the switching matrix is equipped with limited range wavelength converters (LWC) the pool of W wavelengths available on each of the F fibers is divided in wavebands of L wavelengths and conversion may happen only within the same waveband.

We believe $|\mathcal{S}|$ is a measure of the "cost" of the CDS algorithm, since it is correlated to the complexity and amount of devices needed to implement the switching matrix. So a "fair" comparison between scheduling algorithms must be performed keeping with the same values of $|\mathcal{S}|$. This leaves a lot of freedom to the engineer that has to dimension the B, W or L, F parameters, answering the questions:

1) *time vs. channel multiplexing*: is it better to have more delays or more channels (i.e. wavelengths and fibers) on the output links?
2) *space vs. wavelength multiplexing*: is it better to have more fibers or more wavelengths on the output link?

Numerical results provided an answer to these questions, showing that it is more profitable to invest in channels rather than in delays. That increasing the number of wavelengths/channels per interface improves performance is a known results [**?**]. But when we compare solutions with a given scheduling space we discover that performance are greatly improved if we halves the number of wavelengths and provide at least 1 delay per interface, as shown in figure 14. Here the curve B=1 (0 delays and only cut through path) is by far not optimal, while the choice of $B = 2$ outperforms significantly the other cases.

Another non intuitive result is shown in Figure 15, showing that LWC is not necessary worse than FWC. Comparing the the two solutions with the same scheduling space it happens that a little increase in delays may well compensate the limited range conversion, also improving the overall performance.

B. Comparison of end-to-end packet ordering issues in synchronous and asynchronous OPS networks

The aim of this activity is to evaluate the performance improvement due to optical packet alignment, with and without the packet sequence preservation constraint. We assume packets of fixed duration. It has been shown that in this

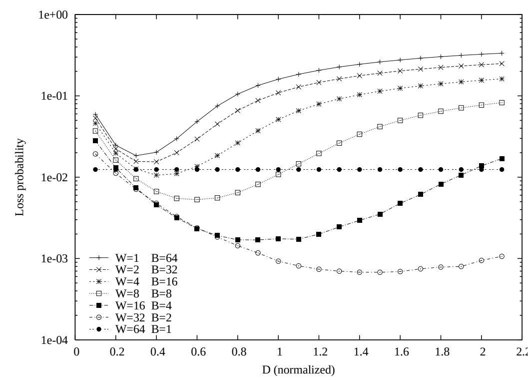


Fig. 14. Trade off between W and B: PLP as a function of D for a switch with FWC, $F = 1$ and CDS algorithms working on a scheduling set having cardinality $|\mathcal{S}| = 64$

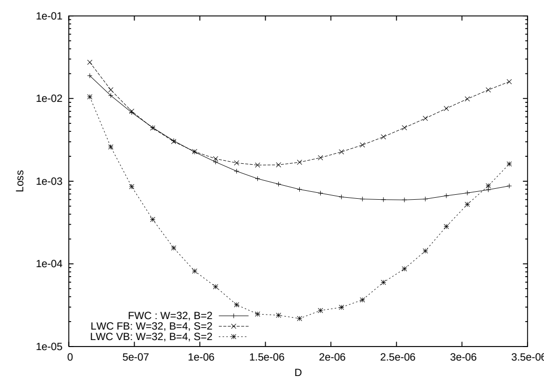


Fig. 15. PLP as a function of D comparing FWC and LWC. For a switch with a switch with $|\mathcal{S}'| = 64$.

situation, if packets are aligned, packet order can be guaranteed with no performance penalty, still providing the optimum throughput/delay performance [43]. Furthermore, this can be accomplished with a simple scheduler, which makes a round-robin (RR) distribution of the packets across the output wavelengths. We are interested in assessing the merits of the RR approach to design schedulers operating in the asynchronous case, when packets are not aligned at switch inputs. The scheduler under test employs one RR pointer per output fibre. For an arriving packet targeted to output fibre $f \in [1 : F]$:

- its transmission wavelength is given by RR[f]. After scheduling the packet, the RR[f] pointer is incremented by 1, modulo W, being W the number of wavelengths in fibre f [3];

- the delay d assigned to each packet is the minimum delay that fulfills two constraints: (i) the packet must not overlap with other packets in the same output wavelength, (ii) the packet should not overtake previous packets that arrived from the same input fibre and are targeted to the same output fibre, irrespective to its transmission wavelength.

Contraint (ii) enforces the packet sequence between packets in the same input fibre/output fibre pair. We evaluate the performance of a switch fabric with an infinite number of FDLs of lengths 0,1,2,... times a FDL granularity value. Our objective is to evaluate the buffering requirements to achieve a packet loss probability (PLP) below 10^{-6}. We estimate this, as the delay D that satisfies that the probability of a packet to be assigned an FDL larger than D is below 10^{-6}. Note that this is a pessimistic approach for the buffering requirements. The granularities of the FDLs evaluated, normalized to packet duration, are $\{0.2, 0.5, 0.8, 1, 1.5, 2\}$. Our scheduler does not implement void filling techniques for those cases where voids can be created (as packets are of fixed duration, this can only happen for granularities 1.5 and 2). Exponential interarrival time of input packets is considered. Input loads tested are $\rho = \{0.3, 0.5, 0.7, 0.9\}$. The results shown in this paper are for switch sizes N=4 input and output fibres and $W = \{8, 32\}$ wavelengths per fibre. More complete results obtained for switch sizes $N = \{2, 4, 8\}$ and $W = \{2, 4, 8, 16, 32, 64\}$ are not included, but confirm the conclusions drawn.

The columns named 'Async. order' and 'Async. no order' in Table II collect the results with and without the packet sequence constraint (ii). Columns 'Sync' provide the optimum performance bound, achieved when packets are aligned, and FDL granularity is set to packet duration. For the sake of comparison, the row named '1 (No RR)' illustrates the results of a non-RR scheduler which assigns to each packet the output wavelength which allows the lowest delay, assumming a packet granularity of 1.

The conclusions of our study are that if optical packets are not aligned and packet sequence is enforced, RR and non-RR schedulers become unstable at medium loads (denoted by * in the table), which yields to a notable performance degradation. This effect is related to the moment in which an incoming packet finds a delay D available attending to constraint (i), but is assigned delay $D + 1$ because of constraint (ii) (packet order). After this moment, subsequent packets in the same input fibre/output fibre pair, can not be scheduled to overtake the packet in delay $D + 1$, forbidding delay D in all the up to W output wavelengths of the same output fibre, and thus shrinking the system throughput. This suggests that if packet order is to be preserved, packet alignment in the nodes should be considered. However, if packet sequence is not enforced, system unstability appears only at high loads, and the RR strategy yields to a low performance degradation compared to the synchronous case. Also, a narrow performance improvement margin exists for more complex non-RR schedulers which jointly decide on packet output wavelength and packet delay.

[3]For the sake of simplicity we assume all fibers carry the same number of wavelength but the discussion can be easily generalized to the case of different number of wavelengths.

VI. Service Oriented OCNs

To provide advanced QoS-enabled connectivity service, e.g., VPN, to a set of different applications running on national or world-wide scale, such as Global Grid Computing or Service Delivery Platform, the OCNs must be enhanced with the capability to interact with diverse qualified applications and consistently perform the network resource allocation among such applications. The problems to be addressed are:

- support for direct invocation and fulfillment of QoS-enabled connectivity services;
- a decision algorithm to share the electronic and optical resources among the incoming service requests thus properly mapping the application traffic flows on the overall network resources.

To solve the first problem a service architecture, namely Service Oriented Optical Network (SOON) architecture, based on distributed signalling among designated service nodes, has been designed to fulfill service requests issued by applications while masking the transport related implementation details from the abstract request of the service. To solve the second problem, a techno-economic algorithm is proposed to help MPLS routers take the decision whether to switch traffic flows (Label Switched Paths or LSPs) optically or electronically. The mathematical foundations for such techno-economic analysis with a Bayesian decisor are defined and the decisor properties are evaluated in a realistic scenario.

A. Advanced connectivity service provisioning in GMPLS networks

The SOON architecture [**?**] consists of a GMPLS-enabled transport network enhanced with new functional layer, named Service Plane (SP), acting on top of the GMPLS CP (see Fig. 16). Several application-to-network mediation layers have been proposed in order to address the application requirements on the network service provisioning in transport networks [**?**]. These approaches generally either do not exploits the CP capabilities [**?**] [**?**], or are based on centralized resource broker [[**?**]].

In SOON architecture, the SP exploits the source-initiated GMPLS CP signaling and adopt a cooperative approach for providing on-demand transport network services with a level of service abstraction and resource virtualization suitable for being invoked by applications. In fact, the SP translates a network service request issued by an Application Entity (AE), i.e. an OSI Layer 7 (L7) entity, into a set of technology-dependent directives to the network devices based on a distributed signaling among designated "service nodes". On the one hand, the SP allows AEs to issue network services in terms of application parameters without any reference to the network technology and topology details. On the other hand, it allows network technologies to be independent from future evolution of the network services and the CP to be unburdened of the service-related functionality. In order to

FDLS	INPUT LOAD											
	0.3			0.5			0.7			0.9		
	Async order	Async no order	Sync	Async order	Async no order	Sync	Async order	Async no order	Sync	Async order	Async no order	Sync
0.2	5 ; 1	5 ; 1		9 ; 3	8 ; 3		20 ; 19	15 ; 4		* ; *	124 ; 18	
0.5	3 ; 1	3 ; 1		8 ; 8	4 ; 2		* ; *	9 ; 3		* ; *	* ; 12	
0.8	4 ; 2	2 ; 1		19 ; *	4 ; 2		* ; *	7 ; 2		* ; *	* ; 10	
1	4 ; 2	2 ; 1	2 ; 1	* ; *	3 ; 2	2 ; 2	* ; *	7 ; 2	3 ; 2	* ; *	* ; 15	9 ; 3
1 (No RR)	3 ; 1	2 ; 1		5 ; 3	3 ; 2		22 ; *	7 ; 2		* ; *	* ; 15	
1.5	6 ; 2	2 ; 1		* ; *	3 ; 2		* ; *	14 ; 2		* ; *	* ; *	
2.0	9 ; 2	2 ; 1		* ; *	3 ; 2		* ; *	* ; 2		* ; *	* ; *	

TABLE II

BUFFER REQUIREMENTS TO ACHIEVE A PLP BELOW 10^{-6}. ASYNCHRONOUS TRAFFIC WITH ORDER, WITHOUT ORDER AND SYNCHRONOUS TRAFFIC COMPARISON. NODE SIZE $N = 4$, $W = 8; 32$

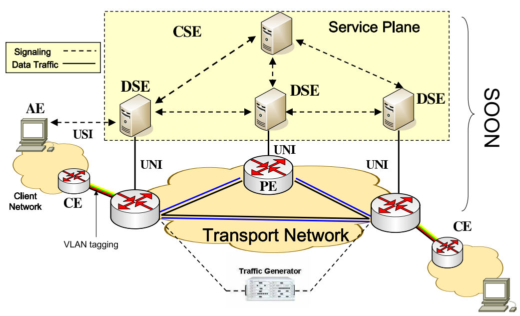


Fig. 16. Testbed for on-demand VPN set-up and QoS validation using SOON architecture

perform the mapping, the SP has been conceived as composed by one Centralized Service Element (CSE) and a number of Distributed Service Elements (DSEs). The CSE performs AE identification and authorizes the relevant service requests using the information stored in its Service Level Agreement (SLA) database. The DSEs process network service requests, issued by applications and received via a User to Service Interface (USI), and interact with the other DSEs to perform the necessary technology-specific network setting via UNI to the controlled edge network nodes.

An implementation of SP based on Java technology has been realized and used to validate the SOON architecture

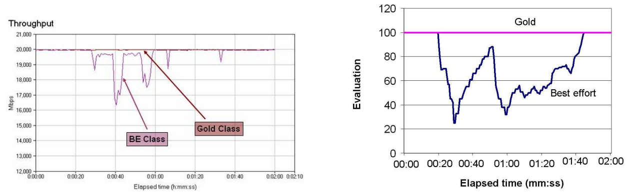
1) for GMPLS integration within ITU-T Next Generation Network (NGN) architecture
2) for the automatic VPN set-up across MPLS network and with QoS support.

The first activity regards the NGN harmonization with circuit-switched transport network provided with GMPLS CP functionalities. The network scenario we considered includes a multi-layer transport network composed of commercial IP/MPLS routers and WDM equipments (e.g. OXC) manufactured by different vendors. This work led to a novel GMPLS-enabled Resource Admission Control Function definition and validation according to NGN directives [44].

The second activity led to the experimental validation of the on-demand set-up of L2/L3 VPN with QoS assurance across a MPLS network comprised of metro/core routers (see Fig. 16)using SP. Within this activity the performance of SP has been evaluated and the QoS guarantees have been established and validated.

Concerning SP performance, the timing of the signalling has been evaluated using two PCc connected to MPLS network through routers configured as Customer Edge (CE). Each PC runs an instance of VLC media player, one configured as a Video Server transmitting DVD video and the other as Video Client. The messages exchanged among the SP entities and the network was captured and analyzed using the Wireshark probe software installed on PCs. From the experimental results, the time needed by the SP for fulfilling the service request (i.e. overall service provisioning time) is about 13 sec [45]. This time is not significantly affected by the number of routers involved since the SP configures them in parallel and approximately at the same time. In particular, the processing time of the SP is about 1,8 sec and the time needed by router to elaborate the UNI commands is about 2.5 sec.

Concerning the QoS implementation and validation, we created as many VPLS instances as the traffic class and performed consistent DiffServ [8] settings across network nodes. Specifically we identified two traffic categories. The first one for non-guaranteed traffic that was conventionally called Best Effort (BE) and mapped to DiffServ BE class. The second one for valuable traffic that was called Gold and mapped to the Expedited Forward (EF) class defined in the DiffServ approach. The network link under test was loaded with a High Definition (HD) Video Stream tagged in Gold Class (about 20 Mb/s). In addition a traffic congestion has been produced by using a traffic generator and applying a load equal to 980 Mbps marked in BE Class. Then we investigates that the QoS requirements of video traffic are addressed in case of traffic congestion. In Fig. 17 we report the results of objective throughput (in 17(a)) and of perceived quality measurements (in 17(b)) for both the Gold and BE classes in the presence of link congestion. As shown in Fig. 16, the Gold class steadily maintains the throughput, permitting and excellent quality for the service.



(a) Throughput (b) Perceived quality

Fig. 17. Throughput vs elapsed time for HD streaming video (a) and corresponding average perceived quality from a group of reviewers (b).

B. Multi-layer switching algorithm for an all-optical router

When the switching nodes have multiple switching alternatives (electronic, optical fiber based, optical wavelength based, optical packet based etc.) an important question to answer is how to map the traffic flows on the switching layer. This activity proposes a solution to this question in a multi-layer capable router, which can route the incoming LSPs in end-to-end lightpaths or in hop-by-hop connections [46].

Let N refer to the number of LSPs handled at a given random time by the multi-layer router, and let $L(d_i, x)$ refer to the loss function:

$$L(d_i, x) = (C_e(i) + C_o(N - i)) - U(x), \quad x > 0 \tag{1}$$

where d_i denotes the "decision" of routing i LSPs ($i = 1, 2, \ldots, N$) out of a total of N in the electronic domain, and is defined for some decision space $\Omega = \{d_1, \ldots, d_N\}$. $C_e(i)$ and $C_o(N-i)$ refer to the cost associated to routing i LSPs in the electronic domain and $U(x)$ refers to the utility associated to a queuing delay of x units of time, experienced by the electronically switched LSPs. Once we have defined the loss function, we can compute the Bayes risk, which is essentially the expectation of the loss function with respect to x [**?**]:

$$R(d_i) = \mathbb{E}_x L(d_i, x) = (C_e(i) + C_o(N - i)) - \mathbb{E}_x U(x) \tag{2}$$

The goal of our algorithm is to obtain the optimal decision d_N^* such that the Bayes risk $R(d_N^*)$ is minimum.

The next step is to define a set of utility functions, $U(x)$, that measures the QoS experienced (in terms of queuing delay) by the electronically-switched packets. Three utility functions are proposed:

- **Mean utility:** computes the mean delay of the LSPs in the electronic domain.
- **Hard-real time utility:** evaluates the probability that the delay in the router queue is lower than a given $T_{\max}$ threshold.
- **Elastic utility:** assesses the gradual degradation of elastic services.

To complete the model definition, a metric is introduced for quantifying the relative cost of optical switching with respect to electronic switching (R_{cost}). We have considered a *linear* cost approach, that evaluates the ratio at which the optical cost increases with respect to the electronic cost.

Fig. 18 shows the risk function and the utility function (dashed line) when QoS constraints changes assuming the

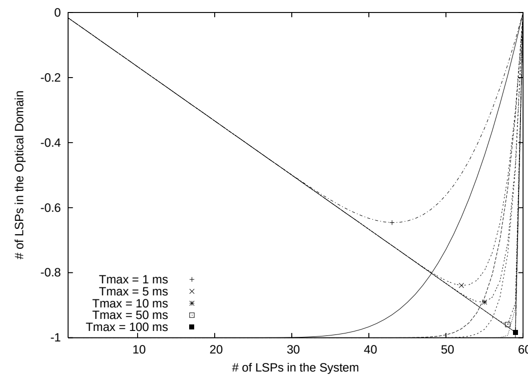


Fig. 18. Optimal decisions for several $T_{\max}$ values assuming hard real-time utility function (Dashed line = Utility).

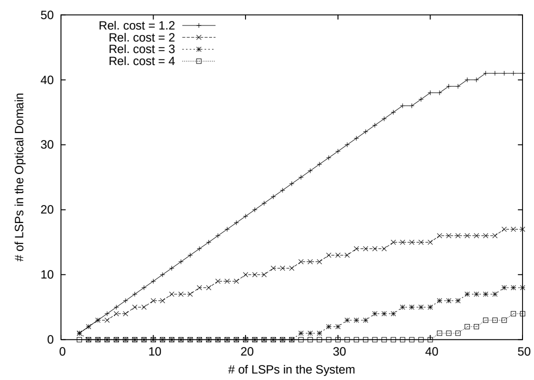


Fig. 19. Optimal decisions when R_{cost} varies using the mean utility function.

hard-real time utility. The optimal decision (d_N^*) is plot with a symbol in all curves. The optimal decision when $T_{\max} = 1ms$ is $d_{60}^* = d_{43}$. This means that 43 LSPs from a total number of 60 LSPs are switched using the electronic layer. However, when QoS constraints are tighter ($T_{\max} = 100ms$) the decisor uses more the electronic resources ($d_{60}^* = d_{59}$).

The second experiment (Fig. 19) shows the impact of R_{cost} using mean utility function, when the number of incoming LSPs in the system increases. R_{cost} refers to the relative cost of optical switching with respect to electronic switching. According to the Fig. 19, the more expensive optical switching is (large values of R_{cost}), the less number of LSPs is switched optically. In other words, for high R_{cost} values, only a small portion of LSPs is switched in the optical domain.

This activity has also evaluated [46]:

1) the self-similar characteristics of the incoming flows, represented by the Hurst parameter H;
2) the variation of the mean and variance of the incoming LSPs;
3) the dynamic behavior of the decisor.

Main conclusions are: $T_{\max}$, H, mean or variance of the incoming flows are parameters that influence the decisor behavior and helps to change the decision based on the traffic features. However, when optical switching becomes too expensive, the R_{cost} is critical in the optimal decision, thus canceling any influence of the other parameters. In this light, the network operator has to decide where the optimal decision lies, trading off the R_{cost} parameter and the incoming traffic parameters.

VII. Conclusion

This paper reported a summary of the joint research activities on core optical networks within the e-Photon-ONe+ project. This experience showed that it is possible to leverage on the integration of different expertises to tackle new problems that could hardly be addressed by a single research group while, at the same, building a shared view of the solutions to such problems within a wide community. The research approach was always quiet pragmatic, aiming at providing and/or validating guidelines and solutions that may be of help to the network engineers in the medium/long term.

For space reasons technical details are not given here, the interested readers may refer to the references. The focus is that of providing an overview of the research topics considered important by the e-Photon-ONe+ community in this technical field, together with a summary of results and technical conclusions.

We believe the content of this paper shows not only that optical core networks are still a lively research topic, but most of all that solutions and evolutionary paths towards their full implementation exist.

Acknowledgment

The work described in this paper was carried out with the support of the e-Photon/ONe+ and BONE ('Building the Future Optical Network in Europe") projects, Networks of Excellence funded by the European Commission through the 6th and 7th ICT-Framework Programme. The authors would like to thank all the colleagues that collaborated at the joint activities here described.

References

[1] FP6 Instruments Task Force, "Provisions for Implementing Networks of Excellence", European Commission, Edition: 12 May 2003, http://cordis.europa.eu/fp6/instr_noe.htm

[2] F. Callegati, F. Neri, P. Poggiolini, "The Network of Excellence e-Photon/ONe on Optical Networks", OFC 2006, Anaheim, CA, March 2006.

[3] http://networks.cs.ucdavis.edu/~mukherje/US-EU-wksp-June05.html

[4] F. Callegati et al., "Research on Optical Core Networks in the e-Photon/ONe Network of Excellence", Proc. of IEEE Infocom 2006, Barcelona, Spain, April 2006.

[5] J. Aracil et al., "Research in optical burst switching within the e-Photon/ONe network of excellence", Optical Switching and Networking, Vol. 4, No. 1, 2007.

[6] M. Poikselka, et al., "The IMS: IP Multimedia Concepts and Services", 2nd ed., Wiley, 2006.

[7] Keith Knightson et al. "NGN Architecture: Generic Principles Functional Architecture, and Implementation", IEEE Communications Magazine, Vol. 43, No. 10, 2005.

[8] L. Wu, B. Davie, S. Davari, P. Vaananen, R. Krishnan, P. Cheval, J. Heinanen "Multi-Protocol Label Switching (MPLS) Support of Differentiated Services", IETF RFC 3270, May 2002

[9] B. Quinn and K. Almeroth, "IP multicast applications: Challenges and solutions", IETF RFC 3170, Sep. 2001

[10] P. Soproni, M. Perenyi, T. Cinkler, "Grooming-Enhanced Multicast in Multilayer Networks", ONDM 2007, Athens, May 2007

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

[12] Thomas H. Cormen et al., ÂŞIntroduction to AlgorithmsÂŤ, Second Edition, MIT Press and McGraw-Hill, 2001, Section 23.2: ÂŞThe algorithms of Kruskal and PrimÂŤ, pp.567ÂŰ574.

[13] R. Inkret et al., "Advanced Infrastructure for Photonic Networks: Extended Final Report of COST Action 266", Faculty of Electrical Engineering and Computing, University of Zagreb, 2003.

[14] ILOG CPLEX Mathematical Programming Optimizer (www.ilog.com/products/cplex).

[15] M. Perenyi, P. Soproni, T. Cinkler, D. Larrabeiti, "Regular Reconfiguration of Light-Trees in Multilayer Optical Networks", ONDM 2008, Athens, March 2008.

[16] N. K. Singhal, L. H. Sahasrabuddhe, B. Mukherjee, "Optimal multicasting of multiple light-trees of different bandwidth granularities in a WDM mesh network with sparse splitting capabilities", IEEE/ACM Transactions on Networking, Vol. 14, Iss. 5, pages 1104-1117, 2006

[17] N. Sengezer, E. Karasan, "An Efficient Virtual Topology Design and Traffic Engineering Scheme for IP/WDM Networks", ONDM 2007, Athens, Greece, May 2007.

[18] B. Puype, Q. Yan, D. Colle, S. De Maesschalck, I. Lievens, M. Pickavet, P. Demeester "Multi-layer traffic engineering in data-centric optical networks", ONDM 2003, Budapest, Hungary, February 2003.

[19] M.C. Necker, "Improving performance of sdh/sonet-wdm multilayer networks using weighted integrated routing", BeitrÃd'ge zur 13. GI/ITG Fachtagung Kommunikation in Verteilten Systemen (KiVS 2003), pages 421–432. VDE, 2003.

[20] M.C. Necker, C.M. Gauger, and S. Bodamer, "A new efficient integrated routing scheme for sdh/sonet-wdm multilayer networks", Proceedings of the Optical Fiber Communication Conference (OFC 2003), OSA, 2003.

[21] S. Spadaro et al., "The CARISMA ASON/GMPLS Network: Overview and Open Issues", Proceedings of 9th International Conference on Transparent Optical Networks 2007, 2007.

[22] J. Comellas et al. "Integrated IP/WDM Routing in GMPLS-Based Optical Networks", *IEEE Network*, pages 22–27, 2003.

[23] Paul Ghobril and Samir TohmÃľ, "Dynamic traffic in Multi-Granular Optical Networks" Proceeding of the ONDM2006, Copenhagen, May 22-24 2006.

[24] Paul Ghobril, âĂIJAnalytical Model for Dynamic Waveband Switching,"Proceedings of the 12th IFIP/IEEE Conference on Optical NetworkDesign and Modelling ONDM 2008, March 2008.

[25] Raul C. Almeida Jr., Divanilson R. Campelo, Helio Waldman and Ken Guild, "Accounting for Link Load Correlation in the Estimation of Blocking Probabilities in Arbitrary Network Topologies", IEEE Communications Letters, Vol. 11, No. 7, July 2007.

[26] A. Farkas, J. Szigeti, T. Cinkler: "P-Cycle Based Protection Schemes for Multi-Domain Networks", DRCN 2005, Ischia, Italy, Oct. 2005

[27] R. Romeral, M. Yannuzzi, D. Larrabeiti, X. Masip, M. UrueÅĎa: "Multi-domain G/MPLS recovery paths using PCE," In 10th European Conference on Networks & Optical Communications, NOC 2005, UK, July 2005.

[28] B. Quoitin, C. Pelsser, L. Swinnen, O. Bonaventure, S. Uhlig: "Interdomain traffic engineering with BGP," Communications Magazine, IEEE, Volume: 41, Issue: 5, May 2003.

[29] A. Farrel, J.-P. Vasseur, J. Ash: "A Path Computation Element (PCE)-Based Architecture", IETF RFC 4655 (Informational), August 2006.

[30] D. Meskó, G. Viola, T. Cinkler, "A Hierarchical and a Non-Hierarchical European Multi-Domain Reference network: Routing and Protection", Networks2006, Nov 6-9, 2006, NewDelhi, India, (www.networks2006.de)

[31] R. Martinez, R. Munoz, R. Casellas, J. Comellas, and G. Junyent, "Experimental shared path protection in distributed all-optical GMPLS-based networks," in *DRCN2007*, 2007.

[32] M. Tornatore, R. Munoz, R. Casellas, R. Martinez, A. Pattavina, "Effects of Outdated Control Information in WDM Networks," submitted for publication, August. 2008.

[33] R. Martinez, C. Pinart, F. Cugini, N. Andriolli, L. Valcarenghi, P. Castoldi, L. Wosinska, "Challenges and requirements for introducing impairment-awareness into the management and control planes of ASON/GMPLS WDM networks", IEEE Comm. Mag., Dec. 06.

[34] C. Pinart, E. Le Rouzic, I. Martinez, "Physical-Layer Considerations for The Realistic Deployment of Impairment-Aware Connection Provisioning", ICTON 2007, June 2007

[35] P. Castoldi, F. Cugini, L. Valcarenghi, N. Sambo, E. Le Rouzic, M. J. Poirrier, N. Andriolli, F. Paolucci, A. Giorgetti, "Centralized vs. Distributed Approaches for Encompassing Physical Impairments in Transparent Optical Networks", ONDM 2007, May 2007

[36] F. Cugini, N. Sambo, N. Andriolli, A. Giorgetti, L. Valcarenghi, P. Castoldi, E. Le Rouzic, J. Poirrier, "GMPLS extensions to Encompass

Shared Regenerators in Transparent Optical Networks", ECOC 2007, Sept. 2007
[37] C. Guillemot et al. "Transparent optical packet switching: the European ACTS KEOPS project approach". IEEE J. Lightwave Technol. 1998; 16(12): 2117ÃŘ2134.
[38] F. Callegati, W. Cerroni, G. Corazza, "Optimization of Wavelength Allocation in WDM Optical Buffers", Optical Networks Magazine, Vol. 2, No. 6, pp. 66-72, 2001.
[39] S. Yao et al. "Unified study of contention-resolution schemes in optical packet-switched networks", IEEE J. Lightwave Technology, Vol. 21, No 3, March 2003, Page(s):672 - 683.
[40] F. Callegati, W. Cerroni , L. H. Bonani, F. R. Barbosa, E. Moschim, G. Pavani, "Congestion Resolution in Optical Burst/Packet Switching with Limited Wavelength Conversion", Proc. of IEEE Globecom 2006, San Francisco, CA, USA, November 2006.
[41] F. Callegati, W. Cerroni , "Which trade-off between time and wavelength congestion resolution in Optical Burst or Packet Switching?", Proc. of Photonics in Switching 2006, Iraklion (Crete), Greece, October 2006.
[42] F. Callegati, W. Cerroni, and G. S. Pavani, "Key parameters for contention resolution in multi-fiber optical burst/packet switching nodes broadnets 2007", 4th International Conference on Broadband Communications, Networks and Systems (BROADNETS 2007), Raleigh, USA, pp. 217-223, September 2007.
[43] Pavon-Marino P., Gonzalez-Castano F.J., Garcia-Haro J. "Round-Robin wavelength assignment: A new packet sequence criterion in Optical Packet Switching SCWP networks". European Transactions on Telecommunications 2006; 17(4): 451-459.
[44] F. Baroncelli, B. Martini, V. Martini, and P. Castoldi " Supporting Control Plane-enabled Transport Networks within ITU-T Next Generation Network (NGN) architecture" Network Operations and Management Symposium (NOMS) 2008, April 2008
[45] F. Baroncelli, B. Martini, V. Martini, and P. Castoldi "A distributed signaling for the provisioning of on-demand VPN services in transport networks", Integrated Network Management (IM) 2007, Germany, Munich , May 2007.
[46] V. López, J.A. Hernández, J. Aracil, J. P. Fernández Palacios and O. González de Dios, "A Bayesian Decision Theory Approach for the Techno-Economic Analysis of an All-Optical Router (extended version)", Computer Networks, Vol. 52, No. 1, 2008.
[47] J. Szigeti, R. Romeral, T. Cinkler, D. Larrabeiti, David, "p-Cycle Protection in Multi-Domain Environment," in: Photonic Network Communications, Sept. 2008
[48] D. Staessens et. al, "Computation of high availability connections in multidomain IP-over-WDM networks", submitted to Infocom 2009.
[49] D. Staessens et al., "Enabling high availability over multiple optical networks", IEEE Communications Magazine, June 2008, Vol. 46, pp. 120-126