Quality of Service Aware Call Admission Control in Cell Based Multi-Service Photonic Network

Marko Lacković

Original scientific paper

Abstract: **The article tackles the problem of quality of service assurance in photonic networks. The idea of multi-service photonic network model with the coexistence of optical circuit and packet switching mechanisms and cell communication is used as a basis for service differentiation in the optical domain. Cell loss ratio as a key performance indicator determines the required optical switching mechanism. Service provisioning is performed using call admission control mechanism with real-time cell loss ratio estimation procedure. Service blocking probability calculation utilizes discrete event simulation of service provision and teardown requests applied to core network topology from COST 266 project. Three simulation scenarios are included in the analysis – pure optical packet switching network, and coexistence of optical packet and circuit switching with and without possibility of communication redirection between the switching mechanisms. Simulation scenarios are additionally altered with the cell loss ratio constraint and number of delay lines.**

Index Terms**:** multi-service photonic network, service provisioning, call admission control, quality of service, performance analysis

I. INTRODUCTION

Service differentiation in the optical domain is interesting both from the client and the provider perspective. A client can demand a dynamic communication provisioning through the photonic network with certain quality of service (QoS) parameters or attributes (e.g. cell loss ratio, service availability, delay, delay variation, etc.). This class of networks will be denoted as multi-service photonic networks (MSPN) [1], as they enable different services on the photonic level.

The idea of the multi-service photonic network is based on the complete automatization of the optical transport network functionality, with a simple IP-over-DWDM protocol stack, intelligent photonic layer with resource discovery, like the available bandwidth or switching capabilities. These components should enable transparent communication provisioning.

The optimal photonic network resource utilization and quality of service assurance are achieved in the MSPN model through the coexistence of optical switching mechanisms, including optical circuit (OCS) and optical packet (OPS) switching [2].

Different optical switching mechanisms can inherently support different quality of service parameter values. Optical packet switching can provide high capacity utilization and flexible communication, but poor guarantees on the delay variation. Optical circuit switching on the other hand enables communication without delay variation, but with limited reconfigurability and capacity granularity. This is an obvious distinction from the approaches based on a single switching mechanism. The existence of several optical switching mechanisms with dynamic resource reservation enabling QoS guaranteed transport has been demonstrated in laboratory conditions [3]. The precondition to such implementation is the hybrid optical switching that can support a wide range of traffic patterns and is composed of slow and fast switching fabrics [4].

We will assume in this work a centralized control plane with call admission control functionality. It selects a switching mechanism which will be utilized for specific service provision request within the defined QoS constraints. This decision is based on the current state of the network and functionality which determines if the operator can satisfy service QoS requirements. Performance evaluation method is hence one of the important components of the call admission control. The focus will be set to the cell loss ratio, as the quality of service parameter and, procedures for its quick estimation.

The second section of the article tackles the related work, particularly in the area of hybrid photonic networks, dynamic provisioning of communication in the optical domain, including call access control features, and quality of service on the optical layer. A multi-service photonic network model has been introduced in the third section. Call admission control mechanism is a crucial part of the network management and control model, while the service requests model serves as the simulation model for the service provision and teardown requests. The fourth section introduces a real-time estimation of the cell loss ratio, being the focal point of the article. Cell loss ratio is the key performance indicator which serves as the precondition for the call admission control mechanism. The cell loss ratio estimation is based on parameter extraction from measurements. The concept of effective bandwidth has been introduced and used for the service blocking probability analysis. The fifth section contains service blocking probability results calculated using discrete event simulation. Three simulation scenarios have been suggested – pure optical packet switching network, coexistence of optical packet and circuit switching with and without possibility of communication redirection between the switching mechanisms.

II. RELATED WORK

One of the research directions in the field of photonic networks are network architectures which employ several network technologies simultaneously. The aim is to improve the network design and communication performance by combining the advantages of different technologies. Some researchers denote such networks as hybrid, having several approaches. One of them employs a hierarchy of optical layer networks with different network technologies, meaning that the lower of two adjacent layer networks functions as a server layer setting up a virtual topology for the upper client-layer. This is a common approach for electro-optical transport network architectures, such as IP-(SDH)-DWDM [5]. The modification of this approach is the replacement of the electronic layer with photonic switching layer, such as optical burst switching (OBS) [6]. An overview of optical switching mechanisms can be found in [2].

The hybrid approach can be extended by employing several optical switching mechanisms in parallel. This approach has been described in literature, such as in polymorphic multi-service optical networks (PMON) [7], and in integrated switch for hybrid OBS and wavelength-switching [8].

The MSPN model takes the hybrid approach one step further by employing the optical switching mechanisms in parallel, sharing the same bandwidth resources in the same network simultaneously. Some

Manuscript received July 31, 2009; and revised February 6, 2010.
Author is with the Faculty of Electrical Engineering and Computing, University of Zagreb, (marko.lackovic@ieee.org).

1845-6421/10/8179@2010 CCIS

of the approaches in this category are OpMiGua [9] and Overspill Routing in Optical Networks (ORION) [10].

Dynamic service provisioning in photonic network has been implemented in several testbeds. However, these approaches implement single switching mechanism, optical circuits, either in the form of wavelength channels (such as networks for scientific purposes DRAGON [11], UCLP/CA*net4 [12] or OptiPuter [13]) or Ethernet connections (such as USN [14], CHEETAH [15] or OMNInet [16]). Employed control planes include distributed solutions, such as generalized multiprotocol label switching (GMPLS) [11][15], or through centralized Web service platforms [13].

The concept of optical services has been introduced in previous works, such as the concept of differentiated optical QoS-service classes [17], or classes of services defined by traffic and performance parameters like the connection setup time, service availability, resilience and routing constraints [18]. Apart from services, some approaches use other ways to differentiate traffic, such as priority index [19], or quality of protection [20].

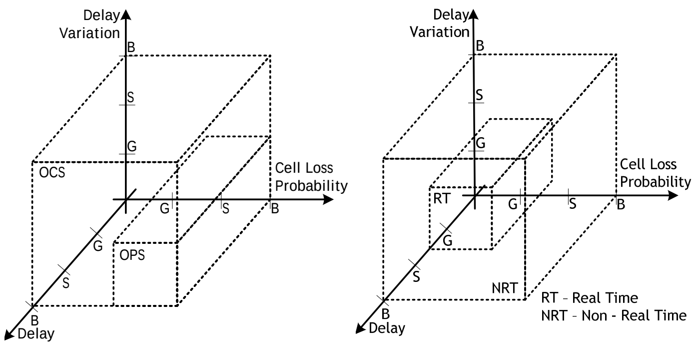
III. MULTI-SERVICE PHOTONIC NETWORK MODEL

A. Network Management and Control Model

The complete network management, rather than separate equipment item management, is becoming essential for the service providers as it enables simple reconfiguration, monitoring, failure detection and recovery and dynamic provisioning.

The implementation of network management and control mechanisms changes the focus from a user perspective to a provider perspective. For example, a provider has to select the appropriate switching mechanism to support requested QoS parameters. If the QoS parameters can be satisfied with the connectionless switching mechanism, like OPS, the OCS mechanism can potentially satisfy those parameters. However, the idea is to choose the optimal solution in terms of solution cost minimization.

Fig. 1. graphically depicts a range of performance values that can be assured by the OCS and OPS switching mechanisms (Fig. 1.*a*), and performance values required by real time and non real time services (Fig. 1.*b*). Values are expressed as performance classes (*G* - gold, *S* - silver and *B* - bronze) by performance category.



a) OCS and OPS Switching Mechanism

b) Non-real Time and Real Time Services

Fig. 1. Range of Performance Values

Fig. 2. will be used to explain the role of performance calculation algorithms in the network management and control in more details. A service request contains the following information:

- *A communication source and destination pair* which serves as the crucial information for the path calculation, and
- *A set of requested QoS parameters* which are posing constraints on communication performance.

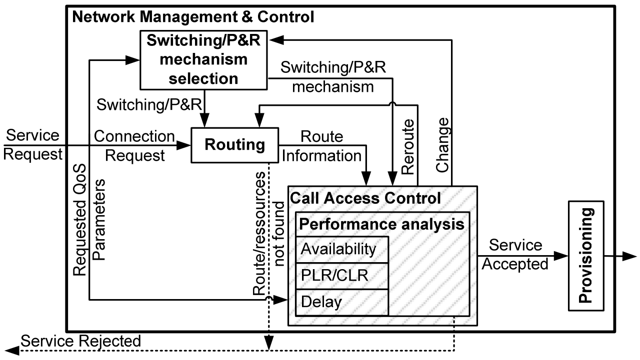


Fig. 2. Role of Performance Calculation Algorithms in the Network Management & Control

A service request triggers the connection provisioning procedure. In the simplest case where a service does not have any QoS requirements, a connection request provision corresponds to the selection of a physical path between the defined source and destination which has enough available resources. The information on the path is sent to the call admission control (CAC) module which automatically accepts the path, communication and the service request if no additional QoS requests are defined. Additionally the CAC module can check existing connections to determine how a new connection influences their performances. Although this is a very important feature, it will be left as a design option for the operator as it is usually a time demanding procedure in the case of a larger network with many connections.

In the case of a service request with defined QoS parameters the call admission control has to evaluate the performance of the potential path and compare it to the required QoS parameters. It is assumed that the developed network and performance calculation models are accurate enough to provide results matching the real performances that are to be measured after the connection provisioning.

If the calculated performance values are not in accordance with the requirements, a new path search is started, until the potential paths are exhausted, or the calculated performance values reach required parameters. The calculation of a new path can be preceded by the change of the switching mechanism. The search for an alternative path will probably result in a path with worse communication performances as the first selection of the path was optimal assuming the defined switching requirements. The change of the switching mechanism can potentially lead to the communication with better performances. A good example is the case where optical packet switching cannot satisfy the delay variation requirements due to the high network load, and is being replaced by the optical circuit switching. The optical circuit switching was not selected in the first case as it represents the most expensive solution, and the cheaper solutions that could have provided performances good enough to satisfy requested QoS parameters were examined first.

B. Service Requests

In this work communication demands are characterized by the required communication capacity. It is assumed that service requests in node i for node j are arriving with the intensity $\lambda_{i,j}$, inversely proportional to the interarrival service time $IAT_{i,j}$. Service teardown requests arrive with the teardown intensity $\mu_{i,j}$, inversely proportional to the mean service hold (duration) time $HTM_{i,j}$.

Service provision and teardown requests can be modelled using the Markov chain, as depicted in Fig. 3.. Each state is defined by the number i of active (concurrent) services. The chain has unlimited number of states. Each state can have transitions with non-zero probabilities only to adjacent states. This chain therefore corresponds to a birth and death process.

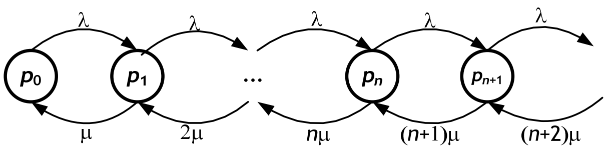


Fig. 3. Markov Chain Description of Required Service Capacity

The birth rate corresponds to the service provisioning request intensity, while the death rate corresponds to the service teardown request intensity. To simplify the analysis it will be assumed that the network has unlimited resources, so that the arrival intensity remains constant over time, and connection acceptance does not depend on the network load. This assumption is valid if the objective of dimensioning procedure is to assure very small blocking probability. The dying intensity is directly proportional to the number of active services in the network.

Transition probabilities p_{ij} from state i in time point t to state j in time point $t + \Delta t$ for the birth and death process can be defined using the Poisson process. Stationary probabilities with the assumption $\lambda < \mu$ can be calculated [21]:

$$\hat{p}_0 = e^{-\frac{\lambda}{\mu}}, \ \hat{p}_n = \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n e^{-\frac{\lambda}{\mu}}. \qquad (1)$$

It is clear that the occurrence of higher number of active services will have very small probability considering normal link loads. It will be assumed that a network will be dimensioned so that it can support a number of concurrent active services that occur in the network with probability larger than 1% (as shown in TABLE 1).

TABLE 1 MAXIMUM NUMBER OF ACTIVE SERVICES WITH PROBABILITY < 1%

λ/μ	0.1	0.3	0.5	0.7	0.9
Service Number (Probability)	2 (0.0045)	3 (0.0014)	4 (0.0032)	4 (0.0070)	5 (0.0044)

From TABLE 1, we see that if the λ/μ ratio is limited with 0.9 one can expect no more than 5 concurrent active services with the probability 99%.

Up to now only the number of concurrent active services has been analyzed, but not their capacity. If it is assumed that services have uniformly distributed capacity requests with the lower bound c_l, and the upper bound c_h, it can be concluded that the relation between the static capacity demand between nodes i and j (usually defined by a traffic matrix), denoted as $C(i, j)$, and the dynamic service request parameters from node i to node j is given as

$$C(i,j) = n_s \frac{\lambda_{i,j}}{\mu_{i,j}} \frac{c_l + c_h}{2} = n_s \frac{HTM_{i,j}}{IAT_{i,j}} \frac{c_l + c_h}{2}, \qquad (2)$$

where n_s denotes the maximum number of concurrent services, usually determined by the dimensioning of the access cable, $\lambda_{i,j}$ service arrival intensity in the node i for node j, and $\mu_{i,j}$ teardown intensity for services between nodes i and j. It will be assumed that the same maximum number of concurrent services is the same for the whole network, as well as the lower and upper service capacity constraints. Arrival and teardown intensities can be expressed as the interarrival time *IAT* and holding time mean *HTM*.

Values n_s, c_l, c_h and $C(i, j)$ are usually defined, what enables us to calculate the λ/μ ratio. The same mean service hold time can be defined for the whole network, what enables us to only vary the service arrival intensity depending on the static capacity. Assuming that the upper capacity constraint is equal to the channel capacity C_{ch}, and that *HTM* is equal to the mean service hold time for all services in the network, a following equation can be written for the interarrival time of services at node i for node j:

$$IAT_{i,j} = n_s \frac{HTM_{i,j}}{2} \frac{c_l + c_h}{C(i,j)}. \qquad (3)$$

IV. CALL ADMISSION CONTROL BASED ON REAL-TIME ESTIMATION

The network communication is based on the MSPN model with optical packet switching functionality implemented in each MSPN node [1]. We assume that the communication model is synchronous with fixed length packets (cells). A slotted photonic network model enables use of some results for the ATM networks, as discussed in the following paragraphs. An extensive overview of Call Admission Control Schemes for ATM networks is given in [22].

Cell loss ratio calculation as a part of call admission control should be performed in real-time. This often implies introduction of simplifications which reduce the accuracy of the calculations. On the other hand accuracy can be improved by using traffic measurements and performance monitoring, carried out by the network operator.

For the purposes of call admission control we will introduce a quick evaluation of the cell loss ratio of the incoming request in the case it would be provisioned in the network with some current state (already provisioned communications). This approach calculates required capacity for a new connection that would assure CLR below some defined constraint. The procedure is based on the calculation of the effective capacity and has its roots in the ATM communication.

This approach requires only asymptotic behaviour of the CLR in the regime of large buffers, avoiding other traffic statistics. This fact makes it suitable for on line execution, as it does not require extensive measurements of traffic parameters.

C. Cell Loss Ratio Model

The goal is to keep the cell loss ratio model as simple as possible to allow fast estimation. Large number of models fail to do so, mostly due to complex descriptions of traffic sources, and large number of parameters. Application of these models in the real-time connection admission control is almost impossible.

The cell loss ratio is defined as the ratio of lost cells and all cells. The term lost cells includes all cells that were discarded because of contention that could not be solved, i.e. the insufficient buffer capacities. All physical impairments introduced by the components (such as power losses and noise additions) are not taken into consideration. This model includes only queuing aspects of the cell transmission.

It is rather difficult to accurately evaluate the packet interarrival time correlation from measurements. The solution is to make assumptions on this correlation without real measurement data. The examples from the ATM theory include burst overflow rate [23] and the packet-loss-ratio upper bound [24]. These approaches are based on evaluation of the performance in the limit of long-range correlation; i.e. the burst duration time is infinite. Long-range-correlation limit based on performance evaluation gives the worst case performance, which makes it attractive for conservative traffic management. The alternatives mostly use real-time measurements, compensating for the above long-range-correlation assumption [24][25].

An example of a single server queue that has an infinite buffer with stationary and ergodic packet arrivals from a single source will be constructed. A_t denotes the number of arrivals during period $\langle 0, t]$, and Q the number of packets in the queue in the stationary state. The rate function of the arrival process is defined as

$$I(x) \overset{def}{=} \lim_{t\to\infty}\left[\log P(A_t > xt)\right]. \quad (4)$$

The logarithm of the queue length distribution has the following asymptotic according to the large deviation principle [26]:

$$\lim_{x\to\infty}\frac{1}{k}\log P(Q\geq k) = -\eta. \quad (5)$$

The asymptotic decay rate can be expressed in the fixed length packet case as

$$\eta = \inf_{x>0}\frac{I(x)}{x - N_{wl}}. \quad (6)$$

N_{wl} in the upper expression denotes the number of packets that can be served in unit time. This implies the following asymptotic tail distribution of queue length:

$$p_k = P(Q\geq k) \approx \beta e^{-\eta k}. \quad (7)$$

The results from the large deviation principle will be further used to draw conclusions on the asymptotic of packet-loss ratio in the regime of large buffers. This principle theoretically states that the CLR exponentially decays as the buffer size increases for a wide-range of traffic inputs. In other words, the CLR can be estimated between the real CLR (lower bound) and the CLR in the long-range correlation limit (upper bound). According to [27] by using the Gärtner-Ellis theorem, the CLR of a single-server queue with a finite buffer size, B, is less than p_B:

$$CLR \leq P(Q\geq B) = \beta e^{-B\eta}, \quad (8)$$

where Q denotes the number of buffered packets. Now the asymptotic constant β can be interpreted as the constant which produces the CLR in the long-range-correlation limit, and the asymptotic decay rate η which determines how the actual CLR becomes smaller than the long-range correlation limit CLR due to the effect of packet buffering.

Following [27][28], η can be calculated as

$$\eta = \sup\left\{\theta : \frac{M(\theta)}{\theta} \leq N_{wl}\right\},$$
$$M(\theta) = \lim_{t\to\infty}\frac{1}{t}\ln E\left(e^{\theta A_t}\right) = \sup_x\left\{x\theta - I(x)\right\}. \quad (9)$$

Effective bandwidth is the bandwidth required by a new connection request to maintain the cell loss ratio below defined level. This analysis is based on results in CLR asymptotic in the regime of large buffers [27] - [29].

The QoS objective of the CLR less than the threshold CLR_{obj}, is satisfied if

$$\eta \geq \eta_{obj} \overset{def}{=} \frac{\log\beta - \log CLR_{obj}}{B}. \quad (10)$$

The effective bandwidth required to satisfy the CLR_{obj} constraint (i.e. the condition that the tail distribution of the queue length has the decay rate η_{obj}) can be expressed as

$$\alpha(\eta_{obj}) = \frac{M(\eta_{obj})}{\eta_{obj}} = \lim_{t\to\infty}\frac{1}{t\eta_{obj}}\ln E\left(e^{\eta_{obj}A_t}\right) \leq N_{wl}. \quad (11)$$

It is clear that the $\alpha(\eta)$, $\eta\to\infty$, corresponds to the peak traffic rate, while $\alpha(0)$ corresponds to the average traffic rate.

Generally speaking the effective bandwidth satisfies the additive property [28] - the effective bandwidth of the aggregation of independent sources is equal to the sum of the effective bandwidths of aggregated sources:

$$\sum_{n=1}^{N}\alpha_n(\eta_{obj}) \leq N_{wl}, \quad (12)$$

when the single buffer is shared by N independent sources, and $\alpha_n(\eta_n)$ is the effective bandwidth function of the n-th source.

To calculate the asymptotic constraint we have to introduce more information on the input traffic characteristics. Assuming the input flow in multiplexed of N on-off sources it can be calculated [29]:

$$\beta = \sum_{n=1, M\geq N_{wl}}^{N}\binom{N}{n}\left(\frac{A}{M}\right)^n\left(1-\frac{A}{M}\right)^{N-n} \approx$$
$$\exp\left[-N\left(c\log\frac{c}{p} + (1-c)\log\frac{1-c}{1-p}\right)\right], \quad (13)$$

where

$$c = \frac{N_{wl}}{NM}, \; p = \frac{A}{M}. \quad (14)$$

A is the average rate and M is the peak rate of each on-off source.

The η and β can be expressed as follows [30]:

$$\eta = \sum_{k=1}^{\infty}c_k^{(2)}\left(\frac{1-\rho}{\rho}\right)^k, \; \log\beta = \sum_{k=1}^{\infty}c_k^{(1)}\left(\frac{1-\rho}{\rho}\right)^k, \quad (15)$$

where ρ denotes capacity utilization, and $\{c_i^{(1)}\}$, $\{c_i^{(2)}\}$ can be expressed using $\{d_i^{(1)}\}$, $\{d_i^{(2)}\}$:

$$c_1^{(1)} = \frac{d_1^{(1)}}{d_1^{(2)}}, c_2^{(1)} = \frac{d_1^{(2)}d_2^{(1)} - d_1^{(1)}d_2^{(2)}}{\left(d_1^{(2)}\right)^3}, c_1^{(2)} = \frac{1}{d_1^{(2)}}, c_2^{(2)} = -\frac{d_2^{(1)}}{\left(d_1^{(2)}\right)^3}. \quad (16)$$

D. Parameter Extraction from Measurements

In order to calculate the parameter η, which is essential for the cell loss ratio calculation, parameters $\{d_i^{(1)}\}$, $\{d_i^{(2)}\}$ have to be estimated from the buffer occupancy measurements. If we denote buffer occupancy probabilities and link utilization in the n-th measurement as $\{p_k(n)\}$ and $\{\rho(n)\}$ respectively, we can produce the following equations for η and β:

$$\eta(n) = \frac{\log p_{k_1}(n) - \log p_{k_2}(n)}{k_2 - k_1},$$
$$\log\beta(n) = \frac{k_2}{k_2-k_1}\log p_{k_1}(n) - \frac{k_1}{k_2-k_1}\log p_{k_2}(n). \quad (17)$$

In the upper equation $p_k(n)$ denotes the probability (frequency) that the buffer length in the n-th measurement will be larger than k packets.

The CLR model parameters can be approximated in real time from the buffer measurement. However, only the results presented in [30] will be used, and it will be assumed that the process definition is sufficient to calculate CLR, i.e. that the real traffic conforms to the parameters defined by the service provision request. Taking into account the upper assumptions, following relations can be derived from (15) - (16):

$$d_1^{(1)}(n) = \frac{\log\beta(n)}{\eta(n)}, d_1^{(2)}(n) = \frac{1}{\eta(n)}\frac{1-\rho(n)}{\rho(n)}. \quad (18)$$

Now we have all the parameters needed for the effective bandwidth calculation.

E. Effective Bandwidth

The effective bandwidth is suitable for estimating whether a network can satisfy the parameters defined by the service provision request. The demand acceptance can be judged simply by comparing the link capacity and the sum of effective bandwidths of individual demands.

The drawback is the conservative approach due to the neglect of statistical multiplexing gains. This problem could be overcome by estimating the effective bandwidth of the aggregated connections based on real-time measurement.

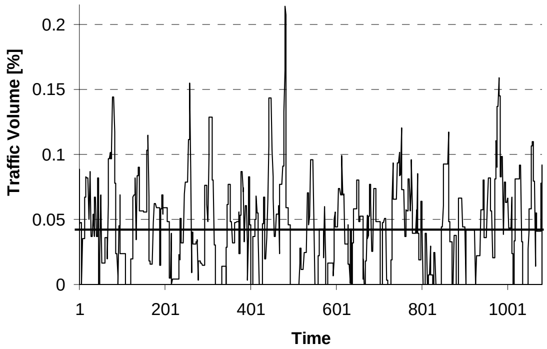
The estimation of the effective bandwidth uses expression (11) and the estimation of parameters by some statistical method or their definition by the service provision request. At the connection-setup phase, one evaluates the effective bandwidth required for the multiplexed demands (i.e. flows on links). The residual free capacity on the link is therefore given by the link capacity minus the estimated effective bandwidth of the multiplexed demands. The connection admission can be judged by comparing the residual capacity of the demand and the effective bandwidth required for a newly requested demand.

It is assumed that the optical node capable of optical packet switching contains a buffer section consisted of fibre delay lines (FDLs). Taking into account that the network model implies cell based communication, it can be stated that each fibre delay line can store a number of cells limited by the number of wavelengths in the system. One of the delay lines has a capacity 0 (no delay), so the total number of cells that can be stored in the buffer with N_{FDL} fibre delay lines and the network with N_{wl} wavelengths can be expressed as $(N_{FDL} - 1) * N_{wl}$.

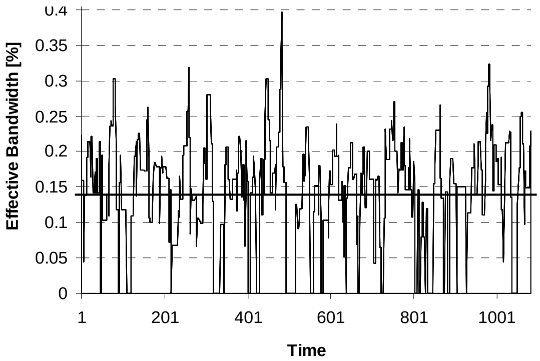
Fig. 4.*a* depicts traffic volume while Fig. 4.*b* depicts effective bandwidth estimation on a link for the same communication setup. The link has been randomly chosen in the core topology (as depicted in Fig. 5.).

Service requests are all classified as OPS communication requests. The buffer has 32 FDLs. The CLR objective has been set to 10^{-5}. Traffic volumes and the effective bandwidth have been expressed in percentages relative to the capacity allocated for the OPS communication (4 channels). The calculation times correspond to the changes in the network (i.e. to each service provisioning and teardown). The mean values in both cases have been depicted with a thick line.

The upper calculations have been performed on the total traffic on one link what is in accordance with the fact that the OPS flows share all wavelengths in the full wavelength conversion scenario. The CLR objective is set for a single link meaning that the flows on that link will not have the CLR higher than the CLR objective. However, the CLR of the whole demands can be potentially higher than the CLR objective what will be analyzed in the subsequent sections.



a) Traffic Volume



b) Effective Bandwidth

Fig. 4. Traffic Volume and Effective Bandwidth

F. Service Blocking Probability Analysis

The effective bandwidth can be used to determine which services can be accepted regarding two conditions:

1. Expected cell loss ratio of already provisioned services will remain lower than the requested cell loss ratio constraint, and
2. Expected cell loss ratio of a new service to be provisioned will be lower than the requested cell loss ratio constraint.

The effective bandwidth in this kind of calculations which are related to the call admission control has two components:

1. Effective bandwidth of already provisioned demands (services), and
2. Effective bandwidth of a new demand (service).

The effective bandwidth of a service corresponds to the effective bandwidth of a demand which supports it.

The sum of effective bandwidth of a new demand to be provisioned and the already provisioned demands on some link has to be less than the capacity of that link. This condition has to be satisfied for all links that are to be used by a new demand. The above figure depicts the effective bandwidth estimation on a single link.

The implementation of the CLR constrained service provisioning is straightforward if it is assumed that the previously described analytical procedure for estimation of $d_1^{(1)}$ and $d_1^{(2)}$ parameters. In that case the effective bandwidth that should have been achieved if the requested service has been already provisioned can be estimated.

If the set of already provisioned services using link l is denoted as $\Omega(l)$, and the set of new services on the link l after the arrival of a new service request S that uses the link l as $\bar{\Omega}(l) = \Omega(l) \cup S$, then the effective bandwidth of the aggregated services (demands) on the link l equals to

$$\alpha_{agg}(l) = \sum_{s\in\bar{\Omega}} A_s \left(1 + d_1^{(2)} \frac{\log PLR_{obj}}{d_1^{(1)} - B}\right), \qquad (19)$$

where A_s denotes the mean packet rate (capacity) of the service S. A new service S can be provisioned if and only if the following statement holds

$$\alpha_{agg}(l) \le C_{OPS}(l)\ \forall l \in \pi(S), \qquad (20)$$

where $C_{OPS}(l)$ denotes the capacity reserved for the OPS communication on the link l, and $\pi(S)$ the set of links (a physical path with selected links on each cable) that is chosen to be used for provisioning communication for service S.

Obviously this kind of calculation is possible only if the effective bandwidth of the aggregated demands including the demand supporting a new service request can be estimated. If we rely only on the simulation procedure, then the aggregated effective bandwidth on the link l can be estimated as

$$\alpha_{agg}(l)=\sum_{s\in\Omega} A_s\left(1+d_1^{(2)}\frac{\log CLR_{obj}}{d_1^{(1)}-B}\right)+C(S)_{peak}, \tag{21}$$

where $C(S)_{peak}$ denotes the peak capacity (packet rate) of the service S. This equation follows the proof that the effective bandwidth is always smaller than the peak packet rate stated in [31].

V. CALCULATIONS

Calculations will be focused on the service blocking probability depending on CLR constraint and the number of FDLs. Different combinations of switching mechanisms will be used, with respect to their influence on the service blocking probability. Similar approach has been employed in [32], analyzing the dependence of blocking probability on the existence of information on physical impairments in optical control plane. Blocking probabilities in the multi-domain environment has been analyzed in [33].

All calculations will be performed using the core topology network depicted in Fig. 5. [34]. The network has been dimensioned with a single link per cable and 4 wavelengths per fibre. The channel capacity corresponds to 40 Gbit/s, and the maximum service capacity is equal to 8 Gbit/s. Requests for service provisioning and teardown will be simulated using a discrete event simulation based on the Poisson model described in section B. The capacity $C(i, j)$ between nodes i and j has been defined using population distance model based on year 2008 data and parameters defined in [34].

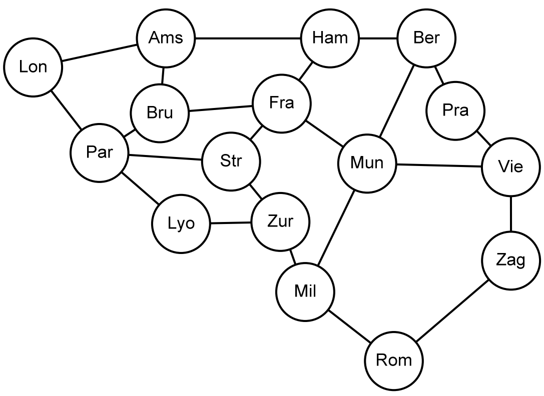
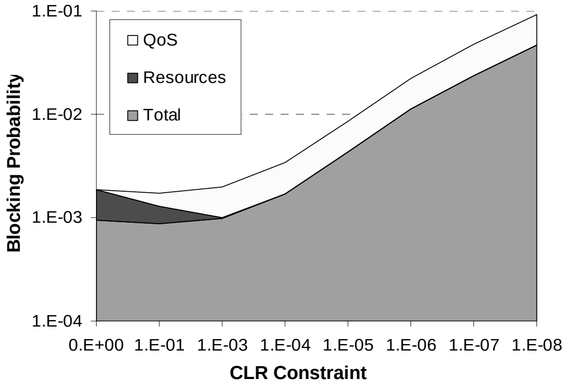


Fig. 5. Network Topology

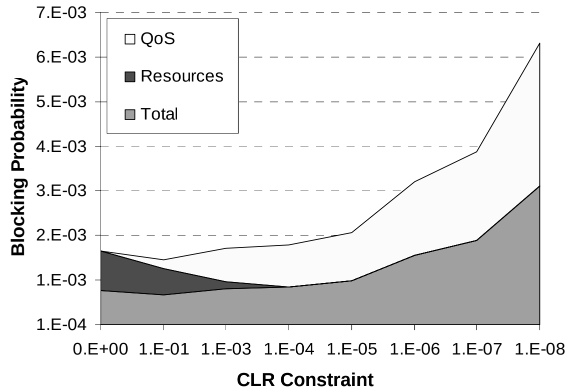
Fig. 6. depicts the service blocking probability with the cell loss ratio constraint for network nodes with optical buffers with 16 FDLs (part *a*, logarithmic axes *y*) and with 32 FDLs (part *b*). The network model in this case implies OPS communication only.

Both graphs have three blocking probability components – total, resources and QoS. Total blocking probability is the sum of the blocking probability caused by the lack of resources and the blocking probability caused by the QoS (cell loss ratio) constraint. The resource blocking probability degrades to 0 relatively fast (with CLR = 10^{-3} for 16 FDLs and CLR = 10^{-4} for 32 FDLs). The blocking probability increases with the decrease of the CLR constraint which causes more service provision requests to be rejected. Service rejection reduces the load of the network and reduces the resource blocking probability.

The number of FDLs in the buffer influences the blocking probability values although the trend remains the same. 16 FDLs case yields total blocking probability higher than 0.01 while the 32 FDLs case yields the total blocking probability as high as 0.003.



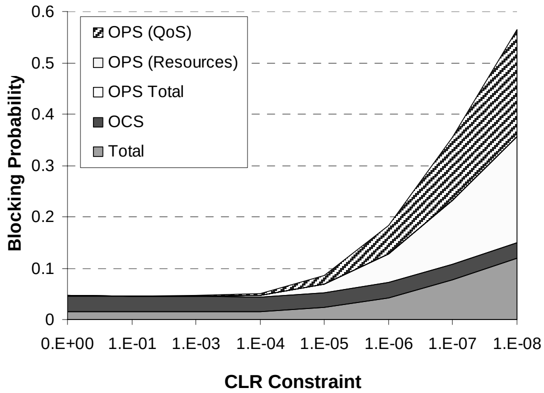
a) 16 FDLs



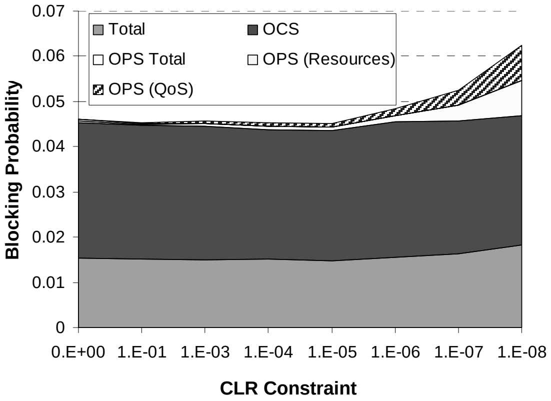
b) 32 FDLs

Fig. 6. Service Blocking Probability with CLR Constraint (OPS Communication Only)

Fig. 7. shows the same analysis as in the previous case, but this time with the service provision requests divided between those requiring the OCS and those requiring the OPS communication (50-50% split of the total number of services). Both service types have the same mean value of the required service capacity. The graph includes total blocking probability, OCS blocking probability and OPS blocking probability (classified as total, caused by the lack of resources and caused by the QoS constraint). Fig. 7.*a* analyzes the case with 8 FDLs, while Fig. 7.*b* analyzes the case with 16 FDLs. If the OPS communication cannot be provisioned, network management and control plane do not analyze the OCS resources that are not used, even if it is possible to provision the service requiring OPS as an OCS communication (denoted as *no redirection to OCS* case). It is visible that the blocking probability is considerably higher than in the previous case of OPS communication.



a) 8 FDLs



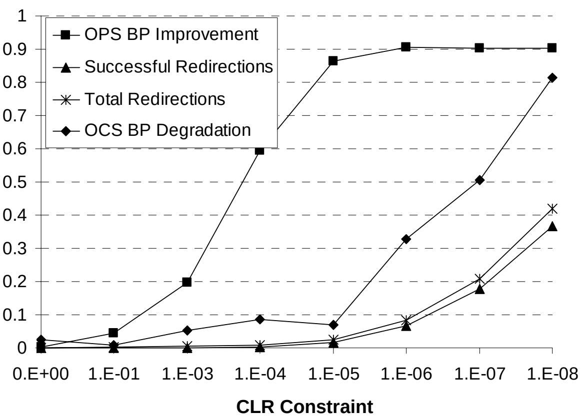
b) 16 FDLs

Fig. 7. Service Blocking Probability with CLR Constraint (Hybrid Communication, No Redirection to OCS)

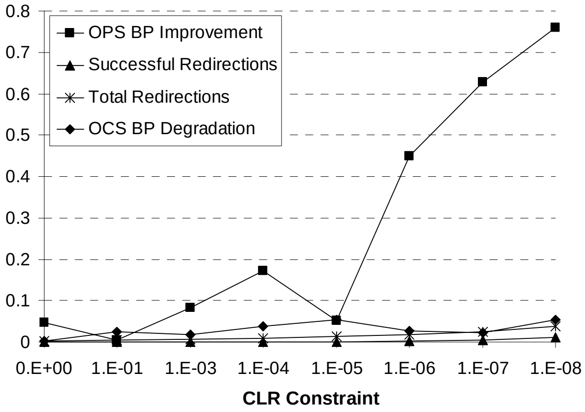
The OCS blocking probability does not depend on the imposed CLR constraint and remains constant throughout the simulation. The OPS blocking probability caused due to resources is too small to be depicted, as the total OPS blocking probability remains very small until the CLR = 10^{-4} constraint for 8 FDLs and CLR = 10^{-5} for 16 FDLs. The total CLR has a QoS-CLR as the main component after the threshold. The difference between 8 and 16 FDLs is in the blocking probability values which are higher for the 8 FDL case.

Fig. 8. analyzes the same communication setup as in the previous case, but with the redirection of unsuccessfully provisioned OPS communication to OCS communication if possible (if enough OCS resources). The *a* part depicts the 8 FDL case, while the *b* part shows the 16 FDL case. Both figures show the relative OPS blocking probability (OPS BP) improvement and the OCS blocking probability (OCS BP) degradation compared to the no redirection case and expressed relatively to the no redirection results. Apart from OPS blocking probability improvements and the OCS blocking probability degradation, the graphs depict the number of total redirections and the number of successfully redirected OPS services expressed relatively to the total number of requested OPS services.

The OPS blocking probability improvement increases with the decrease of the CLR constraint in both cases analyzed. On the other hand the OCS relative degradation increases rapidly in the 8 FDLs case as the initial OPS blocking probability is high and a lot of services are redirected to OCS resources. The OCS degradation is not clearly visible in the 16 FDLs case as the number of redirected OPS services is not very high.



a) 8 FDLs



b) 16 FDLs

Fig. 8. Service Blocking Probability with CLR Constraint (Hybrid Communication, Redirection to OCS)

VI. CONCLUSION

This work has been based on the proposed multi-service photonic network model, which allows coexistence of optical circuit and packet switching in the same network. This is a precondition for the service differentiation in the optical domain. Based on the assumption of fixed length packets and cell loss ratio as the key performance indicator, a call admission control mechanism for requests for service provisioning has been suggested.

Crucial part of the service provisioning procedure is the real-time estimation of cell loss ratio that can be achieved for the new service and its impact on already provisioned services. The proposed network model and call admission control mechanism have been applied to the core network topology proposed within COST 266 project. Using discrete event simulation of service arrival and teardown requests based on the defined Poisson model, the service blocking probability has been calculated with variations of number of fibre delay lines and switching mechanism selection process. Three simulation scenarios have been suggested – pure optical packet switching network, and coexistence of optical packet and circuit switching with and without possibility of communication redirection between the switching mechanisms. In all calculations blocking probability caused by the lack of resources and the blocking probability caused by the cell loss ratio constraint have been distinguished.

The blocking probability improvements justify the use of such call admission control mechanisms in the photonic network, along with the support for different switching mechanisms. The real-time estimation procedure is a step forward in practical implementation of such mechanisms.

VII. REFERENCES

[1] R. Inkret, M. Lackovic: *Multi-service photonic network*, in Proc. of 8th International Conference on Telecommunications 2005, Volume 2, pp. 493 – 500, June 2005.

[2] R. Ramaswami, K. Sivarajan, G. Sasaki: "Optical Networks: A Practical Perspective", Morgan Kaufmann, 3rd ed., pp. 619 – 660, 2009.

[3] M. Takagi et al.: *400Gb/s hybrid optical switching demonstration combining multi-wavelength OPS and OCS with dynamic resource allocation*, in Proc. of Conference on Optical Fiber Communication OFC 2009., pp. 1 – 3., 2009.

[4] M. De Leenheer et al.: *Performance analysis of a hybrid optical switch*, in Proc. of International Conference on Optical Network Design and Modeling ONDM 2008., pp.1 –6, 2008.

[5] R. Izmailov et al.: *Hybrid hierarchical optical networks*, IEEE

Communications Magazine, Vol.40, No.11, pp 88 –94, Nov. 2002

[6] C. M. Gauger, B. Mukherjee: *Optical burst transport network (OBTN) - a novel architecture for efficient transport of optical burst data over lambda grids*, IEEE High Performance Switching and Routing, Hong Kong, May 2005.

[7] I. de Miguel et al.: *Polymorphic architectures for optical networks and their seamless evolution towards next generation networks*, Photonic Network Communications, Vol.8, No.2, pp. 177 – 189, Sept. 2004.

[8] C. Xin, C. Qiao, Y. Ye, S. Dixit: *A hybrid optical switching approach*, Proceedings of IEEE Globecom, San Francisco/CA, pp. 3808 – 3812, Dec. 2003.

[9] S. Bjornstad, M. Nord, D. R. Hjelme: *QoS differentiation and header/payload separation in optical packet switching using polarisation multiplexing*, in Proc. of ECOC 2003, Mo3.4.6, pp. 28 – 29, 2003.

[10] E. Van Breusegem et al.: *A broad view on overspill routing in optical networks: a real synthesis of packet and circuit switching?*, Optical Switching and Networking, Vol.1, No.1, pp. 51 – 64, Jan.2004.

[11] T. Lenhman, J. Sobieski, B. Jabbari: *DRAGON: A Framework for Service Provisioning in Heterogeneous Grid Networks*, IEEE Communications Magazine, Vol.44, No.3, pp. 84 – 90, Mar. 2006.

[12] B. St Arnaud et al.: *Web Services Architecture for User Control and Management of Optical Internet Networks*, Proceedings of the IEEE, Vol.92, No.9, pp.1490 – 1500, Sept. 2004.

[13] N. Taesombot et al.: *The OptIPuter: High-Performance, QoS-Guaranteed Network Service For Emerging E-Science Applications*, IEEE Communications Magazine, Vol.44, No.5, pp. 38 – 45, May 2006.

[14] N. S. V. Rao et al.: *Ultrascience Net: Network Testbed for Large-Scale Science Applications*, IEEE Communications Magazine, Vol.43, No.11, pp. s12 – s17, Nov. 2005.

[15] X. Zheng et al.: CHEETAH: *Circuit-Switched High-Speed End-to-End Transport Architecture Testbed*, IEEE Communications Magazine, Vol.43, No.8, pp. s11 – s17, Aug. 2005.

[16] J. Mambretti et al.: *Optical Dynamic Intelligent Network Services (ODIN): An Experimental Control-Plane Architecture for High-Performance Distributed Environments Based on Dynamic Lightpath Provisioning*, IEEE Communications Magazine, Vol.44, No.3, pp. 92 – 99, Mar. 2006.

[17] M.R. Yoon et al.: *Optical-LSP Establishment and a QoS Maintenance Scheme Based on Differentiated Optical QoS Classes in OVPNs*, Photonic Network Communications, Vol.7, No.2, pp. 161 – 178, Mar. 2004.

[18] W. Fawaz et al.: *Service level agreement and provisioning in optical networks*, IEEE Communications Magazine, Vol.42, No.1, pp. 36 – 43, Jan. 2004.

[19] M. Xia, M. Tornatore, C.U. Martel, B. Mukherjee: *Service-Centric Provisioning in WDM Backbone Networks for the Future Internet*, Journal of Lightwave Technology, Vol. 27, No.12, pp.1856 – 1865, June 2009.

[20] O. Gerstel, G. Sasaki: *Quality of protection (QoP): A quantitative unifying paradigm to protection service grades*, in Proc. SPIE, Vol. 4599, 12, pp. 12 – 23, 2001.

[21] G. Bolch et al.: "Queuing Networks and Markov Chains", Wiley – InterScience, pp. 217, 1998.

[22] H.G. Perros, K.M. Elsayed: *Call admission control schemes: a review*, IEEE Communications Magazine, Vol.34, No.11, pp. 82 – 91, Nov. 1996.

[23] X. Yu, J. W. Modestino, X. Tian: *The Accuracy of Markov Chain Models in Predicting Packet-Loss Statistics for a Single Multiplexer*, IEEE Transactions on Information Theory, Vol.54, No.1, pp. 489-501, Jan. 2008.

[24] H. Saito, K. Shiomoto: *Dynamic call admission control in ATM networks*, IEEE Journal of Selected Areas in Communications, Vol.9, No.7, pp. 982 – 989, July 1991.

[25] R. J. Gibbens, F. P. Kelly, P. B. Key: *A decision theoretic approach to call admission control in ATM networks*, IEEE Journal of Selected Areas in Communications, Vol.13, No.6, pp. 1101 – 1114, June 1995.

[26] J.A. Bucklew: "Introduction to Sequencing and Large Deviation Techniques in Decision, Simulation, and Estimation", John Wiley and Sons, 1990.

[27] N.G. Duffield: *Exponential bounds for queues with Markovian arrivals*, Queueing Systems, vol. 17, pp. 413 – 430, 1994.

[28] C. S. Chang: *Stability, queue length, and delay of deterministic and stochastic queueing networks*, IEEE Trans. Auto. Cont., Vol.39, No.5, pp. 913 – 931, May 1994.

[29] A. Elwalid, et al.: *Fundamental bounds and approximations for ATM multiplexer with applications to video teleconferencing*, IEEE Journal of Selected Areas in Communications, Vol.13, No.6, pp. 1004 – 1016, June 1995.

[30] G. L. Choudhury, W. Whitt: *Heavy - traffic asymptotic expansions for the asymptotic decay rates in the BMAP/G/l queue*, Commun. Statistics and Stochastic Models, Vol. 10, No.2, pp. 453 – 498, Feb. 1994.

[31] S. Shioda, H. Saito: *Real-Time Cell Loss Ratio Estimation and its Applications to ATM Traffic Controls*, in Proceedings of IEEE INFOCOM `97, pp. 1072 – 1081, 1997.

[32] S. Azodolmolky et al.: *A dynamic impairment-aware networking solution for transparent mesh optical networks*, IEEE Communications Magazine, Vol. 47, No.5, pp. 38-47, May 2009.

[33] X.J. Zhang, S.I. Kim, S.S. Lumetta: *Resource provisioning for dynamic multi-domain WDM networks: Effectiveness and fairness*, in Proceedings Of the International Conference on Optical Network Design and Modeling ONDM 2009, pp.1-6, 18-20, Feb. 2009.

[34] R. Inkret, A. Kuchar, B. Mikac (editors): "Advanced Infrastructure for Photonic Networks", extended final report of COST 266 action, Faculty of Electrical Engineering and Computing, Croatia, 2003.

Marko Lacković has obtained PhD degree from the Faculty of Electrical Engineering and computing in 2005 respectively, with the focus on telecommunication networks. His technical education has been extended by the MBA degree from the Faculty of Economics and Business, Zagreb in 2007. He has started his professional career as a research assistant working in Zagreb and Lausanne, Switzerland. Further professional development has been continued in Ericsson R&D in Zagreb on the project manager position, and Huawei Technologies Co. Ltd., Representative office in Zagreb, Croatia, on the key account manager position. Marko Lacković currently works as a sales director in Atento d.o.o. He has participated in several European Union projects, and has authored and co-authored numerous scientific publications. Research emphasis is on core optical network, with interests in dynamic service provisioning, quality of service and network resilience. He is Executive Board Member at IEEE Croatia Section, responsible for industry relations.