Regenerator Placement strategies for Translucent OBS networks

Oscar Pedrola, *Student Member, IEEE,* Davide Careglio, *Member, IEEE,* Miroslaw Klinkowski, *Member, IEEE,* and Josep Solé-Pareta

***Abstract*—Most research works in optical burst switching (OBS) networks do not take into account the impact of physical layer impairments (PLIs) either by considering fully transparent (i.e., with optical 3R regeneration) or opaque (i.e., with electrical 3R regeneration) networks. However, both solutions are not feasible due to the technological requirements of the former and the high cost of the latter. In this paper, we deal with a translucent OBS (TL-OBS) network architecture which aims at bridging the gap between the transparent and opaque solutions. In order to evaluate its performance, a formulation of the routing and regenerator placement and dimensioning problem (RRPD) is presented. Since such formulation results in a complex problem, we also propose two alternative strategies. In particular, we evaluate the trade-off between optimality and execution times provided by these methods. Finally, we conduct a series of simulation experiments that prove that the TL-OBS network model proposed effectively deals with burst losses caused by the impact of PLIs and ensures that the overall network performance remains unaffected.**

I. Introduction

WITH the advent of ultra high bandwidth access systems such as the passive optical networks (xPON) and the next generation mobile networks (i.e., long term evolution (LTE) and 4G), we are forced to move into the next phase of broadband backbone technologies. Indeed, multi-industry initiatives have already started the definition of new business models with the aim of accelerating mass adoption of new devices and services such us video streaming/conferencing, HDTV, VoIP and VoD.

After becoming a real networking layer, optical technology and optical core transport networks in particular are the unrivalled candidates to meet the demands of such applications. Recent advances in optical technologies are fostering the deployment of fully transparent optical networks which involves all-optical switching and provisioning of end-to-end optical paths. Nevertheless, the Physical Layer Impairments (PLIs) of the optical domain and, concurrently, the lack of effective all-optical regeneration devices prevent it from taking place, at least, in the short-medium term [3]. For that very reason, translucent optical networks are the ideal yet feasible candidates for bridging the gap between opaque (i.e., with Optical-Electrical-Optical (O/E/O) conversion at each node) and transparent networks. Indeed, translucent networks combine features of both opaque and transparent networks allowing signal regeneration only at selected nodes in the network [4]. Note that, if not specifically given differently, in this paper the term regenerator implicitly refers to electrical 3R regenerator, that is, the optical signal undergoes an O/E/O conversion in order to be regenerated.

However, for translucent optical networks to be a competitive solution, they should be designed in such a way that both the cost and power consumption is minimized. Both constraints are clearly related to the number of regenerators deployed across the network, and therefore, their number must be reduced as much as possible. For this very reason, the definition of algorithms either for Regenerator Placement (RP) [5] or for Routing and Regenerator Placement (RRP) (see e.g., [6], [7]), if routing constraints are added to the problem, is essential to the problem's success. These techniques are aimed at minimizing the number of regenerators deployed in the network by finding their optimal location.

Due to the maturity of the technology that Wavelength Switched Optical Networks (WSONs) require (e.g., Reconfigurable Optical Add-Drop Multiplexers (ROADMs) and Optical Cross-Connects (OXCs)), translucent WSONs have been the first to receive the attention from the research community. Indeed, protocol extensions and requirements to take into account the presence of PLIs in WSONs are currently under development within IETF [8]. However, and in light of recent measurements, network operators now forecast a highly dynamic data traffic scenario characterized by short-lived, small granularity (i.e., occupying small portions of a wavelength) flows [9]. In this context, and due to their inflexibility and coarse granularity, WSONs would result in a bandwidth-inefficient approach. Hence, the development of sub-wavelength switching technology is nowadays strongly motivated. Indeed, technologies like Optical Packet Switching (OPS) and Optical Burst Switching (OBS) [10], which were initially proposed ten years ago, are re-gaining much of the research interest together with more recent proposals such as Optical Data-unit Switching (ODS) [11] and Optical Flow Switching (OFS) [12].

Among these sub-wavelength solutions, in this paper, we focus on the OBS switching paradigm. In OBS, edge nodes are in charge of assembling client input packets coming from different sources (e.g., IP packet, Ethernet or SDH frames)

Parts of this work have appeared in Globecom 2010 [1] and HPSR 2011 [2].
Oscar Pedrola, Davide Careglio and Josep Solé-Pareta are with the Department of Computer Architecture, Universitat Politècnica de Catalunya, Barcelona, Catalunya 08034, Spain (e-mail: opedrola@ac.upc.edu; careglio@ac.upc.edu; pareta@ac.upc.edu).
M. Klinkowski is with the National Institute of Telecommunications, 04-894 Warsaw, Poland (e-mail: M.Klinkowski@itl.waw.pl).

into outgoing bigger data containers called bursts which, once ready, are launched optically into the network. Similarly, edge nodes are also responsible for disassembling incoming bursts into original client packets. For each outgoing burst, an edge node emits a separate control information called Burst Control Packet (BCP) which is transmitted out-of-band and delivered to the core nodes with some offset-time prior to the burst. The offset-time provides the necessary time budget to reserve resources along the way from the ingress node to an egress node. Such reservation consists of a wavelength which is booked on-the-fly and can be reused afterwards by any other burst (i.e., the resources are therefore shared among all nodes and subject to statistical multiplexing). Core nodes and their corresponding control units are responsible for reading, processing, and updating the BCPs and for switching individual bursts accordingly. In OBS, core nodes are generally assumed to be wavelength conversion capable.

In this paper, we deal with the RRP problem in a translucent OBS (TL-OBS) network. In [1], we proposed a novel TL-OBS node architecture which relies on Semiconductor Optical Amplifier (SOA) technology to perform all-optical switching operation and full wavelength conversion and is equipped with a limited number of electrical regenerators. Such architecture (shown in Fig. 1) provides a fair access to the regenerator pool since all wavelengths from any input port have the same privileges when requesting a regenerator. However, in contrast to the classical RRP problem found in WSON, where there exists a one-to-one correspondence between optical path/connection and electrical regenerator, in TL-OBS the access to the signal regenerators is, like any other resource, subject to statistical multiplexing. Therefore, it is required the introduction of an additional dimensioning phase which eventually extends the problem to the joint Routing and Regenerator Placement and Dimensioning (RRPD) problem. Since the RRPD problem leads to an extremely complex formulation, we simplify it by decoupling RRPD into the routing and the regenerator placement and dimensioning (RPD) subproblems, and thus, we eventually provide a formal model to solve the so-called R+RPD problem by means of Mixed Integer Linear Programming (MILP) formulations. Since the resulting relaxation is still difficult when large problem instances are considered, we also propose two alternative RPD methods and evaluate their performance by considering the trade-off between optimality and complexity they provide. Finally, we study the performance of the proposed TL-OBS network under the considered R+RPD strategies by means of network simulation.

The rest of the paper is organized as follows. In Section II, we survey the previous work in this topic and highlight the main contributions of this paper. In Section III, first we define the RRPD problem, and then, we present a MILP model to solve it. In Section IV, two relaxed MILP-based resolution methods to solve the RRPD problem are proposed. All strategies proposed are compared and evaluated in Section V. Finally, the conclusions of this study are given in Section VI.

II. Related work and contributions

The evolution of optical networks from traditional opaque towards transparent network architectures has brought to light the serious impact that PLIs have on the optical end-to-end signal quality. In fact, due to these physical constraints and the lack of optical 3R regeneration, the deployment of a fully transparent long-haul network is still not viable. Hence, for the sake of scientific progress, the consideration of PLIs in the design and development of next-generation optical core transport networks has become unavoidable. As a matter of fact, the study and evaluation of translucent WSONs, which rely on already mature technology, has recently received increasing attention from the research community. Such an infrastructure makes use of a limited set of 3R regenerators which are strategically deployed across the network for signal regeneration purposes [13]. Since the research interest on translucent architectures lies in the trade-off between network cost (i.e., O/E/O devices are expensive) and service provisioning performance, both the routing and RP issues must be carefully engineered. However, the RP problem is known to be N-complete [14], and hence, heuristic approaches are generally employed [5]. In addition, recent studies in WSONs (see e.g., [6], [7]) show that by combining the RP problem with the routing problem in the so-called RRP problem, an improvement in the network performance can be achieved. However, in light of the foreseen highly dynamic data traffic scenario, fine-grained and flexible technologies such as the sub-wavelength paradigms (e.g., OPS, OBS, OFS) have emerged as potential candidates to cope with the needs of next-generation optical networks. In this work, we focus on OBS networks, a technology which, in essence, overcomes the technological constraints of OPS and the bandwidth inefficiency of WSONs. In the case of OBS, however, research has been mainly geared towards evaluating the opaque and transparent network architectures. Consequently, the vast majority of the works consider that either an ideal physical layer or signal regenerators at every channel, port and switching node of the network are available (i.e., OBS is either fully transparent or opaque). Recently, however, owing to the increasing interest on assessing the effect of the PLIs in the optical networks field, we find few interesting works that involve the PLI constraint in the evaluation of the OBS network performance.

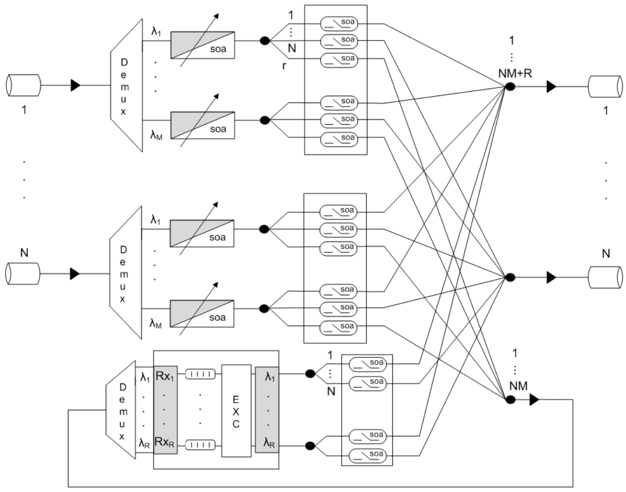


Fig. 1. The TL-OBS node architecture.

For example, some impairment-aware scheduling policies with the aim of minimizing the burst loss probability are presented in [15]. Another interesting study that incorporates PLIs in the definition of an algorithm for distributing manycasting services over an OBS network can be found in [16]. An extensive study that evaluates the design and maximum size and throughput for OBS core nodes considering the effects of a range of PLIs such as amplifier noise, crosstalk of WDM channels, gain saturation and dynamics can be found in [17]. However, in [17], all nodes are equipped with a full set of regenerators (i.e., one per each wavelength) also performing wavelength conversion, and thus, an opaque OBS network is being considered.

Our preliminary work in [1] tackled the issue of designing a complete TL-OBS network architecture. To be precise, we first presented a feasible TL-OBS network architecture and a model to capture the impact of the main PLIs. Then, we evaluated the performance of the TL-OBS network by means of two simple RRPD heuristics. In this paper, we complete such work by presenting a novel MILP formulation and two relaxed MILP-based algorithms to solve the RRPD problem, assessing their performance and comparing it with that of the transparent and opaque reference scenarios. The study here presented follows a static/off-line approach since RRPD decisions are taken during the network planning stage. Note, however, that the routing problem in OBS networks has been extensively studied under both static and dynamic traffic scenarios, and consequently, several techniques to reduce contentions have been proposed (see e.g., [18], [19]). For a recent and comprehensive survey on routing strategies for OBS networks we refer the reader to [20] and to its references. Note that this paper is, by contrast, focused on the mandatory stage of network planning/dimensioning, that is, given a network topology and a prediction of traffic demands, we first compute an optimal routing, and then, perform the placement and dimensioning of regenerators. After this planning stage, the consideration of a dynamic scenario would result in an on-line routing and regenerator allocation problem, an issue which is left out of the scope of this paper.

Here it is worth pointing out that the proposed algorithms require a Quality of Transmission (QoT) estimator to account for the accumulation of the PLIs along the path and by these means determining the feasibility of the path. In the literature, there are two main QoT estimators [21] based on the numerical calculation of the Optical Signal to Noise Ratio (OSNR) [22] or on the computation of the Q-factor value either by means of analytical formulas [23] or numerical interpolation and laboratory measurements [24]. Note that the Q-factor of a path is in direct relation to its signal Bit Error Rate (BER) performance [25]. Although in this work any QoT estimator can be used, in this paper we adopt the OSNR model proposed in [1] and extended in [26]. This OSNR model considers the Amplified Spontaneous Emission (ASE) noise introduced by both the Erbium-Doped Fiber Amplifier (EDFA) and SOA amplifiers as well as the splitting and attenuation losses as the significant signal impairment factors that have to be taken into account [22]. Accordingly, OSNR is defined as the ratio between the signal channel power and the power of the ASE noise in a specified bandwidth (e.g., $0.1nm$ are usually taken by convention). The OSNR model relies on two main components, namely the link and node OSNR, to quantify the OSNR degradation along an optical path traversing n TL-OBS nodes, and therefore, it enables us to determine which bursts will require regeneration when sent into the network. To be precise, we assume that all bursts arriving at the destination node with an accumulated OSNR value lower than a predefined quality threshold (T_{osnr}) cannot be read correctly, and thus, are discarded. Finally, to compute the power and noise values we consider performance parameter values obtained from datasheets of commercially available devices or at most lab trial devices (see e.g., [27]-[29]).

However, it should be noted that in the context of OBS networks, non-linear impairments mainly arise due to the fast ON-OFF switching nature of bursty traffic, which causes the signal power of every single channel to constantly vary. These power variations strongly impact system performances. For example, on the one hand, signal degradations in a burst caused by neighboring bursts which co-propagate simultaneously over several common links (e.g., Cross-Phase Modulation (XPM)-induced crosstalk) and, on the other, OSNR degradation due to dynamic power fluctuations generated by gain changes in amplifiers. Indeed, WDM burst channels randomly switched ON and OFF may be a problem when considering amplifier dynamics. This problem was studied in [30], and it was shown that EDFA amplifiers implemented in a simple and all-optical configuration known as optical gain-clamped can reduce output power excursions by effectively limiting gain ripples. In this paper, non-linear impairments are taken into account by adding an OSNR penalty to T_{osnr}. To be precise, we consider that the OSNR threshold is determined by [22]:

$$T_{osnr} = T_{osnr-min} + T_{osnr-pen} \tag{1}$$

where $T_{osnr-min}$ represents the OSNR tolerance of the receiver, and $T_{osnr-pen}$ accounts for the OSNR penalties due to maximum tolerable Polarization Mode Dispersion (PMD), residual Chromatic Dispersion (CD), and all the other non-linearities. We consider that the $T_{osnr-pen}$ margin is configured by the network operator according to the transmitted signal bitrate, modulation format, etc. [22]. For the systems for which the impact of non-linear impairments is dominant, either larger values of $T_{osnr-pen}$ should be setup (with a possible impact on the network performance) or more accurate and computationally efficient analytical models to capture dynamic PLIs have to be developed.

III. MILP formulation of the RRPD problem

In this Section, we focus on the modeling of the RRPD problem in a TL-OBS network. We begin by presenting the problem definition and its particular design assumptions. In general, our approach to RRPD concerns, respectively, the design of explicit paths to be used to route bursts through the network, and the placement and dimensioning of regenerators at selected nodes on those paths. The result of this design procedure is a set of routing paths and a subset of regenerative nodes which is specified for each individual path that does not comply with the QoT requirements. It is essential to our approach that a burst, whenever sent on a path, will be regenerated only at the nodes that are specified as regenerative nodes for this path. It is worth pointing out that since we are addressing an off-line design problem, we can assume that BCPs are provided at their respective source node with the information on the set of nodes where their corresponding data burst will be regenerated. We also assume that the signal quality of the BCPs is always satisfactory as they undergo an O/E/O conversion at each node for processing purposes and a successful transmission must be assured between at least two adjacent nodes. Finally, it is worth recalling that we assume core nodes with full wavelength conversion capability.

A. RRPD Problem definition

We address the RRPD problem by uncoupling the routing formulation from that of the RPD issue, and therefore, we provide a model to tackle the R+RPD problem. Two main reasons support this modeling decision. First, treating both problems together and at a time would definitely make of the problem an extremely complex undertaking, particularly in terms of computation times or even of solving feasibility. Second, and most compelling, is the fact that in OBS networks, routing must be carefully engineered as the main source of performance degradation is the contention between bursts that arise due to both the lack of optical buffering and the generally considered one-way resource reservation scheme. In fact, in [1], we showed that if routing decisions are biased towards minimizing the number of regenerators deployed, burst losses in network links become uncontrollable, thereby further justifying the decoupling of the problems.

Hence, given a set of traffic demands, we first find a proper routing that minimizes burst losses due to congestion in bottleneck network links. Then, this routing solution is used as input information to solve the RPD problem. Since in the TL-OBS network the access to the regenerator pools is based on statistical multiplexing, the RPD method must deal with both the selection of regeneration nodes and the dimensioning of regenerator pools so that a given target burst loss rate due to QoT non-compliant bursts is satisfied. Thus, the aim of the RPD formulation here proposed is the minimization of the number of regenerators deployed in the network whilst at the same time guaranteeing that losses caused by QoT signal degradation are kept well below those caused by contentions in network links.

B. Global notation

We use $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ to denote the graph of an OBS network; the set of nodes is denoted as $\mathcal{V}$, and the set of unidirectional links is denoted as $\mathcal{E}$. Let $\mathcal{P}$ denote the set of predefined candidate paths between source s and termination t nodes, $s, t \in \mathcal{V}$, and $s \neq t$. Each path $p \in \mathcal{P}$ is identified with a subset of network links, that is, $p \subseteq \mathcal{E}$. Adequately, subset $\mathcal{P}_e \subseteq \mathcal{P}$ denotes all paths that go through link e. Let s_p and t_p denote the source and termination nodes of p. Let $\mathcal{D}$ denote the set of demands, where each demand corresponds to a pair of source-termination nodes. Let $h_d = \lambda_d/\mu$ denote the average offered burst traffic load for demand $d \in \mathcal{D}$, where λ_d is the average burst arrival rate and μ is the average burst service rate; Let $\mathcal{N}_p$ be the set of all nodes constituting path p. Finally, let $\mathcal{V}_p$ denote the set of intermediate nodes on path p such that $\mathcal{V}_p = \mathcal{N}_p \setminus \{s_p, t_p\}$.

C. Routing problem

1) Model assumptions: The routing model that we consider and the routing algorithm that we apply are similar to the Linear Programming (LP) approach presented in [20]. To be precise, the authors consider a multi-path routing approach (i.e., splittable routing) to solve the routing problem. The objective of this method is to distribute traffic over a set of candidate paths so that to reduce congestion in network bottleneck links. To this end, the network is assumed to apply source based routing, and hence, the source node is able to determine the path that a burst entering the network must follow. Although we take the same routing objective, in our study we consider unsplittable (non-bifurcated) routing and, accordingly, all the traffic offered to demand $d \in D$ is carried over a single path in the network. Note that this approach can be easily converted into a multi-path (i.e., splittable) routing problem by relaxing the routing variables. Nonetheless, we use the unsplittable solution to avoid the problem of out-of-order burst arrival which is inherent in any splittable routing algorithm.

It is also worth noticing that the average burst traffic load (h_d) offered by all path $p \in \mathcal{P}_e$ to a particular link $e \in \mathcal{E}$, will decrease due to both contentions at output ports and at regenerator pools in the preceding links on those paths. This problem was studied in [31], where authors present a reduced link load loss-model for OBS networks based on the *Erlang fixed-point approximation*. This model was later compared with a simplified non-reduced link load loss-model in [32], and it was shown that the accuracy of the non-reduced link load model is very strict for values of the BLP lower than 10^{-2}. Thus, we can assume a non-reduced link load model since this requirement is to be largely met in a properly dimensioned network.

Let $\mathcal{P}_d \subseteq \mathcal{P}$ denote the set of candidate paths supporting demand d; $\mathcal{P} = \bigcup_{d \in D} \mathcal{P}_d$. Each subset $\mathcal{P}_d$ comprises a (small) number of paths, for example, k shortest paths. The selection of path p from set $\mathcal{P}_d$ is performed according to a decision variable x_p. In this study, on the contrary to the assumption taken in [20], variables x_p are forced to be binary. Strictly speaking, a burst flow is routed over path p iff $x_p = 1$. Moreover, there is only one path $p \in \mathcal{P}_d$ such that $x_p = 1$. Hence, these routing constraints can be expressed as:

$$\sum\nolimits_{p \in \mathcal{P}_d} x_p = 1, \quad \forall d \in \mathcal{D}, \tag{2a}$$

$$x_p \in \{0, 1\}, \quad \forall p \in \mathcal{P}, \tag{2b}$$

and the traffic ρ_p to path $p \in \mathcal{P}_d$ can be calculated as:

$$\rho_p = x_p h_d = \left\{ \begin{array}{ll} h_d & \text{if } x_p = 1, \\ 0 & \text{otherwise.} \end{array} \right\} \tag{3}$$

As a consequence, the problem formulations presented in the next subsection are MILP formulations. Notice that the set of variables x_p (i.e., vector $\mathbf{x} = (x_1, ..., x_{|\mathcal{P}|})$) determines the distribution of the traffic over the network. This vector has to be optimized in order to reduce link congestion and to improve the overall network performance.

2) Problem formulation: Following the LP algorithm presented in [20], the next two MILP models are sequentially solved to find a solution to the routing problem. First, let variable y represent the average traffic load on the bottleneck link. Then, the first MILP formulation, which aims at minimizing the load on such particular link of the network, can be written as follows:

$$\begin{aligned} & \textit{minimize} \quad y && \text{(RMILP1)} \\ & \textit{subject to} \\ & \sum\nolimits_{p\in\mathcal{P}_e} x_p h_d - y \leq 0, \ \forall e \in \mathcal{E} && (4) \end{aligned}$$

and subject to the routing constraints given by (2a) and (2b).

Despite minimizing the average traffic load on the bottleneck link, many solutions to this problem may exist and most of them exploit unnecessary resources in the network (i.e., solutions that select longer paths). Therefore, the next MILP is solved in order to obtain, between the solutions of RMILP1, the one that entails the minimum increase of the average traffic load offered to the remaining network links. For this purpose, let us denote y^* as an optimal solution of RMILP1, then we solve the following problem:

$$\begin{aligned} & \textit{minimize} \quad \sum\nolimits_{e\in\mathcal{E}} \sum\nolimits_{p\in\mathcal{P}_e} x_p h_d && \text{(RMILP2)} \\ & \textit{subject to} \\ & \sum\nolimits_{p\in\mathcal{P}_e} x_p h_d \leq y^*, \ \forall e \in \mathcal{E} && (5) \end{aligned}$$

and subject to the routing constraints given by (2a) and (2b). Note that, in constraint (5), we ensure that the maximum average traffic load on the bottleneck link is bounded by the solution of RMILP1.

These MILP models, if sequentially solved, determine the path p that will be in charge of carrying the traffic for each demand d. Hence, only one path $p_d \in \mathcal{P}_d$ is selected as the valid path to be followed by all bursts belonging to demand d. Thus, we can now denote $\mathcal{Q}$ as the set of valid paths, $\mathcal{Q} = \{p_d, d \in \mathcal{D}\}$. In the next Section, we use $\mathcal{Q}$ as input information to solve the RPD problem.

D. RPD problem

1) Model assumptions: Let $\mathcal{P}^o \subseteq \mathcal{Q}$ denote the subset of paths for which the QoT level at receiver t is non-compliant with the quality of signal requirements, and thus, paths $p \in \mathcal{Q}$ requiring regeneration at some node $v \in \mathcal{V}_p$. For each $p \in \mathcal{P}^o$ there may exist many different options on how to build an end-to-end QoT compliant path, composed by its transparent segments, since the node or group of nodes where the regeneration has to be performed might not be a unique solution. Thus, let $\mathcal{S}_p = \{s_1, \ldots, s_{|\mathcal{S}_p|}\}$ denote the set of different options to establish a QoT compliant path for each path $p \in \mathcal{P}^o$, where $s_i \subseteq \mathcal{V}$, $i = 1 \ldots |\mathcal{S}_p|$ and size $|\mathcal{S}_p|$ depends on the length of the transparent segments in path p. Figure 2 illustrates this concept by means of an optical path between a source-termination pair ($s - t$) with two different options to establish a QoT compliant path. To be precise, if s_1 is selected, the optical signal only undergoes regeneration at node v_y, whereas if s_2 is the choice, it is converted to the electrical domain twice (i.e., at nodes v_x and v_z). Hence, $s_1 = \{v_y\}$ and $s_2 = \{v_x, v_z\}$. In this particular case, the transparent segments that make it possible to use both regeneration solutions are segments $[s - v_y]$-$[v_y - t]$ and $[s - v_x]$-$[v_x - v_z]$-$[v_z - t]$. Notice that we could also consider other cases like $s_3 = \{v_x, v_y, v_z\}$, however, we have not depicted all of the options for the sake of clarity. In order to obtain $\mathcal{S}_p$, $p \in \mathcal{P}^o$ (i.e., all possible regeneration options) we make use of the OSNR model presented in [1] as commented in Section II. However, it could also be done considering any other valid QoT estimator. In order to find an upper bound on the size of set $\mathcal{S}_p$ we must focus on the number of nodes constituting the largest path in $\mathcal{P}^o$. To this end, let us denote such a number by $\delta = \max\{|\mathcal{N}_p| : p \in \mathcal{P}^o\}$. Hence, an upper bound on the maximum size of set $\mathcal{S}_p, p \in \mathcal{P}^o$ can be written as,

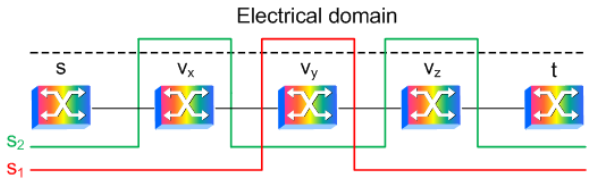


Fig. 2. Two different valid options to perform the regeneration for a particular source-termination pair.

$$\Theta = 2^{(\delta-2)} - 1. \tag{6}$$

We assume that for each path $p \in \mathcal{P}^o$, the selection of the regeneration option s from set $\mathcal{S}_p$ is performed according to a decision variable z_{ps}, which later is referred to as regenerator placement variable, such that the following constraints are fulfilled:

$$\sum\nolimits_{s\in S_p} z_{ps} = 1, \quad \forall p \in \mathcal{P}^o, \tag{7a}$$

$$z_{ps} \in \{0, 1\}, \quad \forall s \in \mathcal{S}_p, \forall p \in \mathcal{P}^o. \tag{7b}$$

Let ρ_v^o denote the offered traffic load requiring regeneration at node v. To estimate ρ_v^o (approximately) we add up the traffic load ρ_p offered to each path $p \in \mathcal{P}^o$ that both crosses and undergoes regeneration at node v:

$$\rho_v^o = \sum\nolimits_{p\in\mathcal{P}^o:\mathcal{V}_p\ni v} \sum\nolimits_{s\in\mathcal{S}_p:s\ni v} z_{ps}\rho_p. \tag{8}$$

Similarly,

$$\rho_v = \sum\nolimits_{p\in\mathcal{P}^o:\mathcal{V}_p\ni v} \rho_p, \tag{9}$$

denotes an estimation of the maximal traffic load that is subject to regeneration at node $v \in \mathcal{V}$.

Eventually, we define a regenerator pool dimensioning function $F_v(\cdot)$, which for a given traffic load ρ_v^o, determines the minimum number of regenerators to be allocated in node v. This number must ensure that a given B^{QoT} is met. Assuming Poisson arrivals and fairness in the access to regenerator pools among bursts, such a function is given by the following discontinuous, step-increasing function,

$$F_v(\rho_v^o) = \lceil B^{-1}(\rho_v^o, B^{QoT}) \rceil \,, \tag{10}$$

where B corresponds to the Erlang B-loss formula which for a given number of regenerators $r \in \mathbb{N}$ available at node v can be calculated as,

$$B(\rho_v^o, r) = \frac{(\rho_v^o)^r / r!}{\sum_{k=0}^{r} (\rho_v^o)^k / k!}, \tag{11}$$

and where $B^{-1}(\rho_v^o, B^{QoT})$ is the inverse function of (11) extended to the real domain [33], and $\lceil \cdot \rceil$ is the ceiling function. It is worth noticing that the Poisson arrivals which lead to an Erlang formula for the dimensioning of regenerator pools can be replaced with another distribution for which the blocking probability is attainable. Because B^{QoT} is a predetermined parameter, for simplicity of presentation we skipped it from the list of arguments of function $F_v(\cdot)$.

For the purpose of problem formulation, it is convenient to define a_r as the maximal load supported by r regenerators given a B^{QoT}, that is, $a_r = B^{-1}(r, B^{QoT})$. Note that the inverse function $B^{-1}(r, B^{QoT})$ is expressed with respect to r and B^{QoT}, which is not the same as in function (10).

Although there is no close formula to compute the inverse of (11), we can make use of a line search method (see e.g., [34]) to find the root ρ^* of the function $f(\rho) = B^{QoT} - B(\rho, r)$ so that the value of a_r is approximated by $a_r = \rho^*$ for any index r. Finally, let R denote the number of regenerators required in the most loaded node, that is, $R = \max\{F_v(\rho_v) : v \in \mathcal{V}\}$. Note that we can make use of vector $\mathbf{a} = (a_1, ..., a_R)$ to obtain the Piecewise Linear Approximation (PLA) of $F_v(\cdot)$, which for a single node $v \in \mathcal{V}$, can be expressed as $F_v(\rho_v^o) = \min\{r : a_r > \rho_v^o\}$. The PLA will be of practical interest in the next subsection, where it will allow us to better deal with function $F_v(\cdot)$, and consequently with $B^{-1}(\cdot)$. For the sake of clarity, function $F_v(\rho_{vo})$ is depicted in Fig. 3 for some exemplary B^{QoT} values. Note that $B^{-1}(\cdot)$ is a real-valued concave function. Moreover, we also provide points (a_r, r) (represented by circles in the plot) which will eventually help us generate the different $\mathbf{a}$ vectors. The accuracy of the PLA of $F_v(\cdot)$ depends on the precision of the line search algorithm that is used to generate vector $\mathbf{a}$. In our implementation of the line search algorithm the termination criteria is set to $\xi = 10^{-6}$, which guarantees a fine approximation (i.e., $B^{QoT} - B(\rho, r) \leq \xi$).

Eventually, vector $\mathbf{a}$ will also be used in subsection IV-C to determine $F_v(\rho_v^o)$ according to Procedure 1. Note that Procedure 1 is a polynomial time algorithm of complexity $O(R)$.

2) Problem formulation: Taking into account the network modeling assumptions previously presented, here we present a mathematical formulation for the RPD part of the problem.

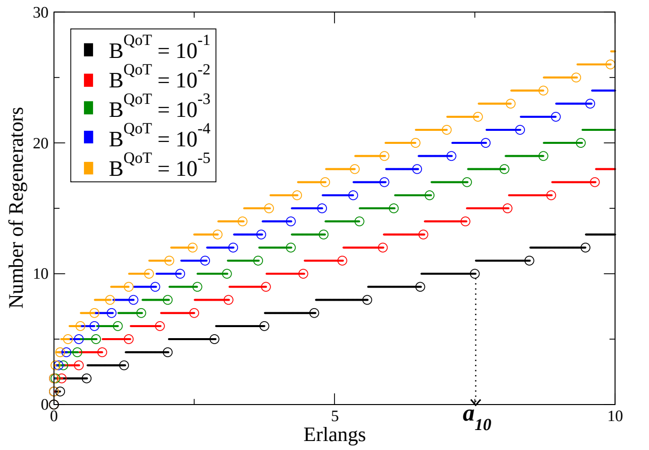


Fig. 3. Discontinuous step-increasing regenerator pool dimensioning function and (a_r, r) points for some exemplary target burst loss probabilities.

Procedure 1 Regenerator Pool Dimensioning

1: $r \leftarrow 0$
2: **while** $\rho_v^o > a_r$ **do**
3: $\quad r \leftarrow r + 1$
4: **end while**
5: $F_v \leftarrow r$

The RPD problem can be formulated as a non-convex optimization problem:

$$\underset{\mathbf{z}}{minimize} \quad F = \sum\nolimits_v F_v(\rho_v^o(\mathbf{z})) \tag{NLP1}$$
$$subject\ to \quad \text{(7a) and (7b)} \tag{12a}$$

where $F_v(\cdot)$ is the step-increasing regenerator pool dimensioning function defined by (10) and $\rho_v^o(\mathbf{z})$ is the function representing the traffic load offered to a regenerator node defined by (8). The optimization objective of NLP1 is to minimize the sum of regenerators installed in network nodes. Constraints (12a) represent the selection of a QoT compliant path from the options provided for each path requiring regeneration. Eventually, the regenerator placement decision vector $\mathbf{z}$ is defined as $\mathbf{z} = (z_{11} \ldots z_{1|\mathcal{S}_p|}, \ \ldots, \ z_{|\mathcal{P}^o|1} \ldots z_{|\mathcal{P}^o||\mathcal{S}_{|\mathcal{P}^o|}|})$.

The difficulty of formulation NLP1 lays in the fact that there is no close formula to express $F_v(\cdot)$ since no such formula exists for the inverse of the Erlang function $B^{-1}(\cdot)$. A way to solve the problem is to substitute function $F_v(\cdot)$, $v \in \mathcal{V}$ with its piecewise linear approximation and reformulate NLP1 as a MILP problem.

For a single node $v \in \mathcal{V}$, the PLA of $F_v(\cdot)$ can also be expressed by means of the following 0-1 integer programming (IP) formulation:

$$\underset{\mathbf{u}}{minimize} \quad F_v = \sum\nolimits_r u_v^r r \tag{IP1}$$
$$subject\ to \quad u_v^r(a_r - \rho_v^o) \geq 0, \ \forall r \in [1, R], \tag{13a}$$
$$\sum\nolimits_r u_v^r = 1, \tag{13b}$$
$$u_v^r \in \{0, 1\}, \ \forall r \in [1, R]. \tag{13c}$$

In IP1, decision variables u_v^r have been introduced in order to represent the number of regenerators required in node v. Due to constraint (13b), in each node only one variable u_v^r is active (i.e., equal to 1), and the one with minimum r satisfying $a_r \geq \rho_v^o$ is found when solving the problem. Notice that formulation IP1, when solved, gives the same solution as Procedure 1. The shortcoming of IP1 is that since ρ_v^o is dependent on vector $\mathbf{z}$ (i.e., ρ_v^o is a function of $\mathbf{z}$), constraints (13a) have quadratic form. To overcome this difficulty, we can consider the following alternative formulation:

$$\begin{aligned} \underset{\mathbf{u}}{minimize} \quad & F_v = \sum\nolimits_r u_v^r r && \text{(ILP1)} \\ subject\ to \quad & \sum\nolimits_r u_v^r a_r \geq \rho_v^o, && \text{(14a)} \\ & \sum\nolimits_r u_v^r = 1, && \text{(14b)} \\ & u_v^r \in \{0,1\},\ \forall r. && \text{(14c)} \end{aligned}$$

It is easy to note that formulation of ILP1 results directly from IP1; it is enough to add up constraints (13a) and use (13b) for substituting $\rho_v^o \sum\nolimits_r u_v^r$ for ρ_v^o.

Eventually, taking into account all network nodes and introducing the regenerator placement decision variables, problem NLP1 can be reformulated as the following MILP problem:

$$\begin{aligned} & \underset{\mathbf{u},\boldsymbol{\rho}^o,\mathbf{z}}{minimize} \quad F = \sum\nolimits_v \sum\nolimits_r u_v^r r && \text{(MILP1)} \\ & subject\ to \\ & \sum\nolimits_r u_v^r a_r - \rho_v^o \geq 0, \quad \forall v \in \mathcal{V}, && \text{(15a)} \\ & \sum\nolimits_r u_v^r = 1, \quad \forall v \in \mathcal{V}, && \text{(15b)} \\ & \sum\nolimits_{s \in S_p} z_{ps} = 1, \quad \forall p \in \mathcal{P}^o, && \text{(15c)} \\ & \sum\nolimits_{p \in \mathcal{P}^o : \mathcal{V}_p \ni v} \sum\nolimits_{s \in \mathcal{S}_p : s \ni v} z_{ps} \rho_p - \rho_v^o = 0,\ \forall v \in \mathcal{V}, && \text{(15d)} \\ & u_v^r \in \{0,1\}, \quad \forall r \in [1,R], \forall v \in \mathcal{V}, && \text{(15e)} \\ & z_{ps} \in \{0,1\}, \quad \forall p \in \mathcal{P}^o, \forall s \in \mathcal{S}_p, && \text{(15f)} \\ & \rho_v^o \in \mathbb{R}^+, \quad \forall v \in \mathcal{V}. && \text{(15g)} \end{aligned}$$

where we consider ρ_v^o to be an auxiliary variable representing the traffic load requiring regeneration offered to node $v \in \mathcal{V}$.

The objective of the optimization problem MILP1 is to minimize the total number of regenerators that have to be placed in the network. Constraints (15a) and (15b) result from the 0-1 representation of the dimensioning function and from the reformulation of IP1 as mentioned before. In particular, the number of regenerators in node $v \in \mathcal{V}$ should be such that the maximum traffic load (given a B^{QoT}) is greater or equal to the offered traffic load ρ_v^o. Constraints (15c) are the QoT compliant path selection constraints. Constraints(15d) are the traffic load offered to a regenerator node calculation constraints. Eventually, (15e), (15f), and (15g) are the variable range constraints.

MILP1 is a well-known Discrete Cost Multicommodity Flow (DCMCF) problem [35]. DCMCF was shown to be an extremely difficult combinatorial problem for which only fairly small instances (in our case, situations where $\mathcal{P}^o$ has a rather small size) can be solved exactly with currently available techniques. Indeed, considering the problem in hand, the total amount of variables can be approximated by $|\mathcal{V}| \cdot R + |\mathcal{P}^o| \cdot \Theta$, where the first term represents the amount of u_v^r variables and the second term is an upper bound on the size of variable vector $\mathbf{z}$. Similarly, the size of the constraint set is $3 \cdot |\mathcal{V}| + |\mathcal{P}^o|$. For example, if the Large network (see Section V and Appendix A for network details) is considered, then $\delta = 12$. Now assume that R is set to 100. Hence, the problem size increases to approximately $8 \cdot 10^5$ variables and $9 \cdot 10^2$ constraints, thereby making highly difficult its exact solution. It must also be noted that, as shown in (6), the size of set $\mathcal{S}_p$ increases exponentially to the size of the problem instance. In order to limit the problem size, we only consider the K smallest options (with respect to the number of regenerations along the path) to fill $\mathcal{S}_p$, that is, $\mathcal{S}_p = \{s_1, ..., s_K\}$. In the next Section, we propose two relaxed MILP-based methods to solve the RPD problem.

IV. MILP-based RPD resolution methods

To overcome the difficulty imposed by the resolution of MILP1, in this Section, we propose two heuristic methods that provide near-optimal solutions to the RPD problem within acceptable computational times. The main idea behind both strategies is to decouple the RPD problem into the RP problem, which is solved first, and the dimensioning phase. Hence, we derive models to solve the so-called RP+D problem. The performance of these methods is later discussed in Section V.

A. Load-based MILP formulation

The MILP formulation here proposed is focused on the distribution of the traffic load requiring regeneration (i.e., ρ_v^o, $\forall v \in \mathcal{V}$). Hence, this load must be aggregated in such a way that the number of regenerators to be deployed is minimized. After a ρ_v^o solution is obtained for each node $v \in \mathcal{V}$, we take advantage of the regenerator pool dimensioning function detailed in Section IV-C to obtain the number of regenerators required.

Owing to the concave character of the dimensioning function (10), it must be noted that it is of our interest to aggregate the traffic requiring regeneration in as few nodes as possible rather than spreading out such load in little amounts over a large number of nodes. Hence, we propose to solve the problem by making use of two MILP models, namely MILP2 and MILP3. These models can be sequentially solved to obtain a sub-optimal solution of MILP1.

First, MILP2 aims at minimizing the number of nodes where the regenerators must be installed (i.e., nodes such that $\rho_v^o > 0$), and thus, groups as much as possible the load that requires regeneration. Let $\mathbf{y} = (y_1, ..., y_{|\mathcal{V}|})$ denote a vector of binary decision variables. Each value corresponds to one node and determines if this node is used as regeneration point by some path $p \in \mathcal{P}^o$ ($y_v = 1$) or not ($y_v = 0$).

Then, we solve the following problem:

$$\underset{\boldsymbol{\rho}^o, \mathbf{z}, \mathbf{y}}{minimize} \quad \sum\nolimits_v y_v \qquad \text{(MILP2)}$$

$$subject\ to \quad \rho_v y_v \geq \rho_v^o, \quad \forall v \in \mathcal{V}, \qquad (16a)$$

$$y_v \in \{0, 1\}, \ \forall v \in \mathcal{V}. \qquad (16b)$$

and subject to constraints (7a), (7b), (15d) and (15g).

Although ILP1 minimizes the number of nodes where the regenerations are performed, multiple solutions to this problem may exist and some of them may exploit more regenerations than required, increasing unnecessarily ρ_v^o at some nodes. Therefore, a second MILP model, that is, MILP3, needs to be formulated with the objective of minimizing the total network load requiring regeneration.

Therefore, let k^* denote an optimal solution of MILP2. Second, we solve the following problem:

$$\underset{\boldsymbol{\rho}^o, \mathbf{z}, \mathbf{y}}{minimize} \quad \sum\nolimits_v \rho_v^o \qquad \text{(MILP3)}$$

$$subject\ to \quad \sum\nolimits_v y_v \leq k^*, \qquad (17a)$$

and subject to constraints (7a), (7b), (15d), (15g), (16a) and (16b).

Due to the simplicity of both formulations, both models are expected to be promptly solved even for large-sized problem instances. Here it is worth mentioning that problems MILP2 and MILP3 as well as routing problems RMILP1 and RMILP2 could have been solved by using a single weighted multi-objective MILP formulation. However, we have considered the sequential approach for both the sake of clarity and to avoid dealing with the weights used in the resulting multi-objective cost functions.

It is also important to notice that the sequential resolution of both MILP2 and MILP3, which will hereinafter be cited within the text as MILP2/3, provides an optimal solution in terms of the distribution of the traffic and not with respect to the number of regenerators (which is precisely the case of MILP1). This being said, the last step in this method is the dimensioning of regenerator pools as detailed in Section IV-C.

B. Reduced MILP1 (MILP1*)

This method aims at reducing the complexity of MILP1 by introducing new constraints to its definition. Specifically, these constraints are the sequentially obtained solutions of both MILP2 and MILP3 as detailed previously in subsection IV-A. Although these new constraints are not valid in that they may exclude the optimal solution of MILP1, they can be used to achieve good near-optimal solutions within reasonable time limits.

Therefore, let us denote g^*, and again k^*, as the optimal sequentially solved solutions of MILP3 and MILP2 respectively. Then, we reformulate MILP1 as follows,

$$\underset{\mathbf{u}, \boldsymbol{\rho}^o, \mathbf{z}}{minimize} \quad F = \sum\nolimits_v \sum\nolimits_r u_v^r r \qquad \text{(MILP1*)}$$

$$subject\ to \quad \sum\nolimits_v y_v \leq k^*, \qquad (18a)$$

$$\sum\nolimits_v \rho_v^o \leq g^*, \qquad (18b)$$

and subject to constraints (15a), (15b), (15c), (15d), (15e), (15f), (15g), (16a) and (16b).

In fact, we sequentially solve all three models in order, that is, first MILP2, second MILP3 and finally MILP1 including all solutions obtained as constraints for the subsequent problem.

It is worth pointing out that, as long as the scenario considered does not involve optical paths that require a large number of regenerations, constraint (18a) is very unlikely to exclude the optimal solution of MILP1. Basically, it is due to the fact that the dimensioning function of our problem is (10), which favours, to some degree, the grouping-like behaviour. Constraint (18b), by contrast, is just an heuristic approach to help solve the problem. Notice that (18b) does not deal with the distribution of the load but with its minimization, and thus, the optimal solution in terms of the number of regenerators is generally excluded.

C. Regenerator dimensioning phase

The load of burst traffic requiring regeneration at any node $v \in \mathcal{V}$ is (approximately) given by (8). In order to determine the number of regenerators required in node v, we define a dimensioning function $f(\rho_v^o, B^{QoT}) : (R^+, R^+) \mapsto \mathbb{Z}^+$. Under the assumption that any burst may access any regenerator in a node (as shown in [1], the considered architecture allows a fair access to the regenerator pool), we make use of the inverse of the Erlang B-loss function as the dimensioning function f. An straightforward way to implement this dimensioning function is to make use of vector $\mathbf{a}$ and Procedure 1, which have been both detailed in Section III-D.

V. Results and discussion

In this Section, we first present and compare the performance results of all the RRPD resolution methods presented in Section III and Section IV. Then, we study the performance of the TL-OBS network architecture under the MILP2/3 method in order to prove that it is effective at satisfying the QoT requirements. As a QoT estimator, we use the OSNR model proposed in [1] and define that all bursts arriving at the destination node with an accumulated OSNR value lower than the predefined quality threshold T_{osnr} cannot be read correctly, and consequently, are dropped.

A. Resolution methods comparison

The evaluation has been performed by considering four different network topologies that are detailed in Appendix A. For this experiment and hereinafter in this paper, we consider a maximum of $K = 1000$ regeneration options to fill set $\mathcal{S}_p, p \in \mathcal{P}^o$, that is, for the network instances considered in this paper, all possible regeneration options are added to the problem. Besides, the T_{osnr} values evaluated are provided in Table I. In this work, we assume bidirectional network links with 32 wavelengths of $10Gbps$ each. We consider $19dB$ as the $T_{osnr-min}$ threshold, a value which is commonly used for the experimental assessment of translucent optical networks with such network links [22]. Moreover, we consider $1dB$ and $2dB$ as additional OSNR penalties ($T_{osnr-pen}$) to account

$T_{osnr-min}$	$T_{osnr-pen}$	Usa-Can	Core	Basic	Large
$19dB$	$1dB$	421	18	349	746
$19dB$	$2dB$	657	55	462	919

TABLE I
NUMBER OF PATHS THAT REQUIRE REGENERATION ($|\mathcal{P}^o|$) AND OSNR THRESHOLD VALUES

Scenario	Usa-Can	Core	Basic	Large
OPAQUE	3904	1472	2624	3648
$\mathbf{T_{osnr} = 20dB}$				
MILP1	355	55	497	854
MILP2/3	351	56	502	866
MILP1*	344	55	496	860
$\mathbf{T_{osnr} = 21dB}$				
MILP1	634	146	752	1231
MILP2/3	652	147	757	1238
MILP1*	646	146	751	1225

TABLE II
RPD RESULTS COMPARISON

Scenario	Usa-Can	Core	Basic	Large
$\mathbf{T_{osnr} = 20dB}$				
MILP1	(7.61%)	2	(2.91%)	(3.4%)
MILP2/3	89	3	9	28
MILP1*	313	3	152	427
$\mathbf{T_{osnr} = 21dB}$				
MILP1	(9.43%)	2	(0.66%)	(2.68%)
MILP2/3	57	3	12	36
MILP1*	224	3	751	280

TABLE III
RPD EXECUTION TIMES (SECONDS) AND OPTIMALITY GAPS (%)

# **Regenerators**	0	250	400	415	426
OSNR (%)	100	99.99	99.72	91.51	12
contention (%)	0	0.01	0.28	8.49	88

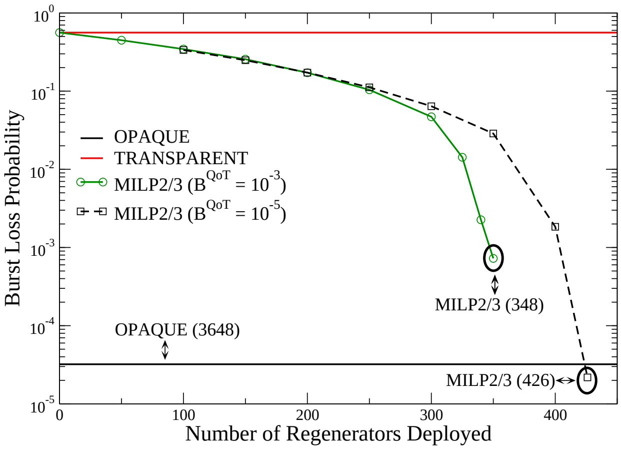
TABLE IV
SHARE OF BURST LOSSES FOR THE CURVE WITH $B^{QoT} = 10^{-5}$ PRESENTED IN FIG 4(A)

for the signal degradation caused by non-linear impairments. Hence, we evaluate our algorithms considering $20dB$ and $21dB$ as the system T_{osnr} thresholds. Note also that T_{osnr} determines the number of paths that require regeneration (i.e., $|\mathcal{P}_o|$), and hence, the level of complexity that is given to the problem. $|\mathcal{P}_o|$ values are also given in Table I for each considered network.

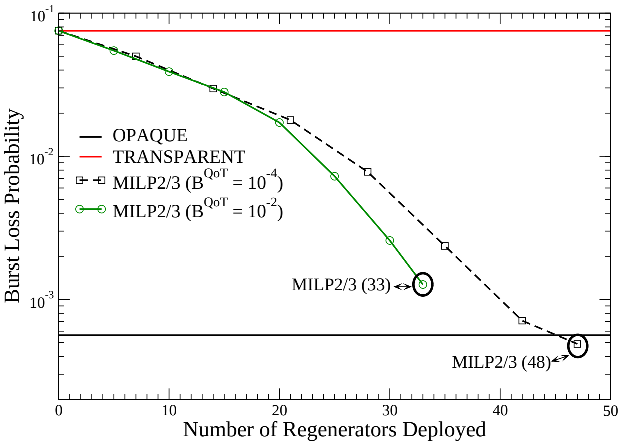
We use CPLEX to solve, for each network and scenario, the three MILP RPD models presented, namely MILP1 (optimal), MILP2/3, and MILP1*. Table II reports the minimum number of regenerators to be deployed considering $B^{QoT} = 10^{-3}$ and that each node injects 20.8 Erlangs into the network. CPLEX is run with the time limit set to 1 hour. Note that Table II also provides the number of regenerators required when an opaque network architecture is considered. Finally, Table III reports the computation times for all the algorithms as well as the optimality gaps (%) for those cases in which optimality is not reached after 1 hour. One can note that MILP1 is solved very effectively when small problem instances are considered (i.e., Core). However, and due to its computational complexity, MILP1 reports optimality gaps in all the other cases. In contrast, MILP1* is always solved to optimality and is able to substantially improve the trade-off provided by MILP1 for some of the scenarios evaluated. Finally, MILP2/3 also reports an overall good trade-off performance, as it is solved very quickly and with an average deviation to the best solution of 1.71%. From the results obtained, it can be concluded that both of the heuristic MILP formulations proposed, that is, MILP1* and MILP2/3, provide satisfactory near-optimal solutions within short running times. In the rest of our experiments, we consider the MILP2/3 algorithm with $T_{osnr} = 20dB$.

B. Impact on the OBS network performance

In order to evaluate the effectiveness of the MILP2/3 method, in this section, we conduct extensive simulations on the TL-OBS network. In this study we consider the BLP as the metric of interest. In Fig. 4(a), we show the results obtained in the Large topology when the number of Erlangs offered per node is equal to 6.4. In this experiment, two different B^{QoT} targets are considered, namely 10^{-3} and 10^{-5}. In addition, the opaque and transparent scenarios are plot and used as benchmark indicators. It is easy to observe that the progressive and even placement of regenerators (i.e., the amount of regenerators to be placed is fairly distributed among all selected nodes) reduces the overall BLP until both B^{QoT} targets are reached (i.e., the required number of regenerators has been deployed). In the $B^{QoT} = 10^{-3}$ case, the BLP is dominated by OSNR losses, and consequently, when all the regenerators have been deployed $BLP \approx B^{QoT}$. On the other hand, if B^{QoT} is set to 10^{-5}, contention losses become predominant, and therefore, $BLP \approx BLP_{OPAQUE}$. For this case ($B^{QoT} = 10^{-5}$), Table IV shows the percentage of losses resulted from contentions in network links and unacceptable OSNR signal levels. One can note that, as expected, OSNR-based losses are progressively reduced with the deployment of regenerators. Similarly, Fig. 4(b), shows the same experiment performed in the Core topology. However, this time each edge node offers 12.8 Erlangs and B^{QoT} targets are set to 10^{-2} and 10^{-4}. It is worth pointing out that both the load and B^{QoT} values were selected in order to illustrate two different and representative situations in both figures. For the sake of illustration, in Table V, the locations and number of regenerators deployed in both experiments are shown. Note that whilst in the Core network the minimum amount of nodes equipped with regenerators is 3, in the Large network it increases up to 13.



(a)



(b)

Fig. 4. MILP2/3 performance in a) the Large topology, and b) the Core topology.

Network	B^{QoT}	Node (Regenerators)
Core	10^{-2}	$0(18), 10(9), 14(6)$
	10^{-4}	$1(19), 3(13), 14(16)$
Large	10^{-3}	$0(26), 6(18), 8(19), 10(34), 12(59), 14(30), 17(13), 18(21), 22(42), 23(19), 27(11), 32(33), 33(33)$
	10^{-5}	$0(32), 6(22), 8(24), 10(41), 12(68), 14(37), 17(18), 18(27), 22(50), 23(24), 27(15), 32(28), 33(40)$

TABLE V
LOCATION AND NUMBER OF REGENERATORS FOR THE EXPERIMENTS PRESENTED IN FIG. 4(A) AND FIG 4(B).

Notice that in the $B^{QoT} = 10^{-2}$ case, although OSNR losses have a noticeable impact on the network performance, the BLP decreases up to nearly 10^{-3}. This is due to the fact that the percentage of the traffic requiring regeneration in the network is quite low, or in other words, $|\mathcal{P}^o|$ has a small size. If B^{QoT} is set to 10^{-4}, in contrast, we observe the same behaviour as in Fig. 4(a), that is, contention losses

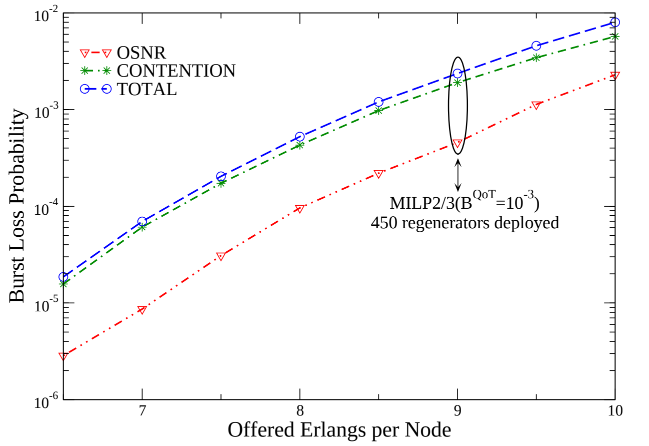


Fig. 5. TL-OBS network performance for a dimensioning assuming the Large topology, the MILP2/3 algorithm, a load of 9 erlangs and a target $B^{QoT} = 10^{-3}$, which results in the deployment of 450 regenerators.

are predominant, and hence, $BLP \approx BLP_{OPAQUE}$. Note that in both figures provided, the BLP found in the case where contention losses are predominant slightly improves that of the opaque case. This is due to the differences in node architectures between the opaque and translucent networks: whilst the opaque network relies on in-line regenerators as in [17], our translucent architecture operates in the feed-back mode (see Fig. 1), and hence, bursts remain in the electrical buffer until a free wavelength is found at the desired output link.

Eventually, we assess how effective at keeping OSNR losses under control the MILP2/3 method is. To this end, we study how both contention and OSNR losses contribute to the total BLP by performing a dimensioning of the TL-OBS network considering the Large topology, a load of 9 erlangs and a target $B^{QoT} = 10^{-3}$. In this scenario, MILP2/3 provides a solution requiring 450 regenerators to be deployed. Given this planning of the network, we evaluate the impact that load variations have on the overall TL-OBS network performance. One can note in Fig. 5, that OSNR losses are kept satisfactorily under control regardless of the load, and thus, that our approach guarantees that OSNR losses are well below those caused by burst contentions in network links.

VI. CONCLUSIONS AND FUTURE WORK

In this paper, we propose several methods for the sparse placement of regenerators in a translucent OBS network. Such methods are based either on an optimal MILP formulation or on heuristic MILP-based techniques. For this purpose, we have focused on the problem of PLIs in OBS networks and addressed the RRPD problem. Strictly speaking, we have uncoupled the routing issue from the RPD problem, and eventually solved the so-called R+RPD problem. We have presented a link congestion-reduction unsplittable routing strategy which is based on a MILP formulation aimed at reducing congestion in bottleneck network links. The routing solution obtained has then been used as input for the RPD problem. The RPD scheme presented relies on the piecewise linear approximations of the inverse of the Erlang B-loss formula.

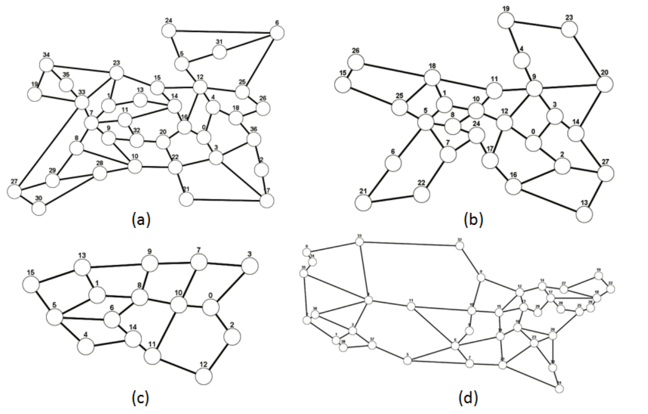


Fig. 6. a) Large (37 nodes), b) Base (28 nodes), c) Core (16 nodes), d) Usa-Can (39 nodes).

Since such formulation corresponds to the complex DCMCF problem, we have also developed two heuristic methods to help solve the RPD problem (i.e., RP+D heuristics). We have evaluated and compared these methods by considering the trade-off between optimality and complexity they provide. After assessing its performance over a range of network topologies, we have found that the heuristic RPD methods proposed, that is, MILP2/3 and MILP1*, provide the best trade-offs. Finally, we have conducted a series of exhaustive simulations in the TL-OBS network proposed considering the MILP2/3 method. From the results obtained, we have concluded that both the architecture and model proposed in this paper ensure that, according to a pre-specified target performance, losses caused by QoT signal degradation are kept satisfactorily under control and do not impact negatively the overall network performance.

In our future work, we plan to extend our model to consider the case of an on-line/dynamic scenario.

Appendix A
Simulation Scenario

In our simulation scenario, we consider several topologies (see Fig. 6), all of which being real network topologies: a set of Pan-European [37] networks known as: Large (a), Basic (b) and Core (c) with 37, 28 and 16 nodes and 57, 41 and 23 links respectively; the JANOS-US-CA [39] (d), a reference network that interconnects cities in the USA and Canada with 39 nodes and 61 links.

Network links are bidirectional and dimensioned with the same number of wavelengths $M = 32$. The transmission bitrate of both transmitters at edge nodes and regenerators at core nodes is set to $10Gbps$.

We assume that each node is both an edge and a core bufferless node capable of generating bursts destined to any other node. We consider the offset time emulated OBS network architecture (E-OBS) [40], and the Just-In-Time (JIT) [41] resources reservation protocol together with the Last Available Unscheduled Channel (LAUC) scheduling algorithm also known as Horizon [42]. For the sake of simplicity, the switching and processing times are neglected.

The traffic is uniformly distributed between nodes. We assume that each edge node offers the same amount of traffic to the network; this offered traffic is normalized to the transmission bitrate and expressed in Erlangs. In our context, an Erlang corresponds to the amount of traffic that occupies an entire wavelength (e.g., 20 Erlangs mean that each edge nodes generates $200Gbps$).

Bursts are generated according to a Poisson arrival process and have exponentially distributed lengths. The mean duration of a burst ($1/\mu$) is $100\mu s$ ($1Mb$). Note that due to both the Poisson assumption and the fact that we neglect both the switching and processing times of bursts, the burst size does not have any impact on the results obtained [31]. In obtaining the simulation results, we have estimated 99% confidence intervals. However, since the confidence intervals found are very narrow, we do not plot them in order to improve readability.

All simulations have been conducted on the JAVOBS [43] network simulator on an Intel Core 2 Quad 2.67 GHz with $4GB$ RAM.

The RMILP1, RMILP2, MILP1, MILP2, MILP3 and MILP1* problems have all been solved using the IBM ILOG CPLEX v.12.1 solver [44].

Acknowledgment

The work described in this paper was carried out with the support of the STRONGEST-project, an Integrated Project funded by the European Commission through the 7th ICT-Framework Programme, the Spanish Ministry of Science and Innovation under the ”DOMINO” project (Ref. TEC2010-18522), the Catalan Government under the contract SGR-1140, and the Polish Ministry of Science and Higher Education under the contract 643/N-COST/2010/0.

References

[1] O. Pedrola, D. Careglio, M. Klinkowski, and J. Solé-Pareta, “Modelling and performance evaluation of a translucent OBS network architecture,”in *Proc. IEEE GLOBECOM 2010*, Miami, FL, USA, Dec. 2010.

[2] O. Pedrola, D. Careglio, M. Klinkowski, and J. Solé-Pareta, “RRPD strategies for a T-OBS network architecture,”in *Proc. IEEE HPSR 2011*, Cartagena, Spain, Jul. 2011.

[3] S. Azodolmolky et al., “A survey on physical layer impairments aware routing and wavelength assignment algorithms in optical networks,”*Com. Netw.*, vol. 53, no. 7, May 2009.

[4] B. Ramamurthy, H. Feng, D. Datta, J.P. Heritage, and B. Mukherjee, “Transparent vs. opaque vs. translucent wavelength-routed optical networks,”in *Proc. OSA OFC 1999*, San Diego, CA, USA, Feb. 1999.

[5] X. Yang, and B. Ramamurthy, “Sparse regeneration in translucent wavelength-routed optical networks: architecture, network design and wavelength routing,”*Phot. Netw. Commun.*, vol. 10, no. 1, pp. 39-53, Jan. 2005.

[6] S. Pachnique, T. Paschenda, and P.M. Krummrich, “Physical impairment based regenerator placement and routing in translucent optical networks,”in *Proc. OSA OFC/NFOEC 2008*, San Diego, CA, USA, Feb. 2008.

[7] W. Zhang, J. Tang, K. Nygard, C. Wang, “REPARE: regenerator placement and routing establishment in translucent networks,”in *Proc. IEEE Globecom 2009*, Honolulu, HA, USA, Nov. 2009.

[8] Y. Lee, G. Bernstein, D. Li, and G. Martinelli, “A framework for the control of wavelength switched optical networks with impairments,”*IETF draft*, draft-ietf-ccamp-wson-impairments-04.txt, Oct. 2010.

[9] O. González de Dios, G. Bernini, G. Zervas, and M. Basham, “Framework for GMPLS and path computation support of sub-wavelength switching optical networks,” *IETF draft*, draft-gonzalezdedios-subwavelength-framework-00, Mar. 2011.

[10] C. Qiao, and M. Yoo, “Optical burst switching (OBS) - a new paradigm for an optical Internet,”*J. High Speed Netw.*, vol. 8, no. 1, pp. 69-84, Jan. 1999.

[11] T. Rasmussen, "Next generation packet-optical systems,"*OSA OFC 2010*, market watch, panel IV, San Diego, CA, USA, Mar. 2010.
[12] V. Chan, "Optical Flow Switching,"in *Proc. OSA OFC 2010*, San Diego, CA, USA, Mar. 2010.
[13] G. Shen, and R. S. Tucker, "Translucent optical networks: the way forward,"*IEEE Commun. Mag.*, vol. 45, no. 2, pp. 48-54, Feb. 2007.
[14] A. Sen, S. Murthy, and S. Bandyopadhyay, "On sparse placement of regenerator nodes in translucent optical network,"in *Proc. IEEE Globecom 2008*, New Orleans, LA, USA, Nov. 2008.
[15] Y. Fan, and B. Wang, "Physical impairment aware scheduling in optical burst switched networks,"*Phot. Netw. Commun.*, vol. 18, no. 2, pp. 244-254, Oct. 2009.
[16] B.G. Bathula, V.M. Vokkarane, R.R.C. Bikram, "Impairment-aware manycasting over optical burst-switched networks,"in *Proc. IEEE ICC 2008*, Beijing, China, May 2008.
[17] H. Buchta, and E. Patzak, "Analysis of the physical impairments on maximum size and throughput of SOA-based optical burst switching nodes,"*IEEE/OSA J. Lightwave Technol.*, vol. 26, no. 16, pp. 2919-2927, Aug. 2008.
[18] S. Kim, N. Kim, M. Kang, "Contention resolution for optical burst switching networks using alternative routing,"in *Proc. IEEE ICC 2002*, New York, NY, USA, Apr. 2002.
[19] L. Yang, G.N. Rouskas, "Adaptive path selection in optical burst switched networks,"*IEEE/OSA J. Lightwave Technol.*, vol. 24, no.8, pp. 3002-3011, Aug. 2006.
[20] M. Klinkowski et al., "An overview of routing methods in optical burst switching networks,"*Opt. Swit. Netw.*, vol. 7, no. 2, Apr. 2010.
[21] D.C Kilper et al., "Optical performance monitoring,"*IEEE/OSA J. Lightwave Technol.*, vol. 22, no. 1, pp. 293-304, Jan. 2004.
[22] R. Martínez, R.Casellas, R. Muñoz, and T. Tsuritani, "Experimental translucent-oriented routing for dynamic lightpath provisioning in GMPLS-enabled wavelength switched optical networks,"*IEEE/OSA J. Lightw. Technol.*, vol. 28, no. 8, pp. 1241-1255, Apr. 2010.
[23] P. Pavon-Mariño et al., "Offline impairment aware RWA algorithms for cross-layer planning of optical networks,"*IEEE/OSA J. Lightwave Technol.*, vol. 27, no. 12, pp. 1763-1775, Jun. 2009.
[24] A. Morea et al.,"QoT function and A* routing: an optimized combination for connection search in translucent networks,"*OSA J. Opt. Netw.*, vol. 7, no. 1, pp. 42-61, Jan. 2008.
[25] G.P. Agrawal, "Fiber-optic communications systems,"*3rd ed. New York, John Wiley & Sons*, pp.490-497, 2002.
[26] O. Pedrola, D. Careglio, M. Klinkowski and J. Solé-Pareta, "An analysis of regenerator placement strategies for a translucent OBS network architecture,"*Tech. Rep. UPC-DAC-RR-2010-38*, Sep. 2010.
[27] INPHENIX: *http://www.inphenix.com/soa_devices.html*, accessed in Mar. 2010.
[28] MRV: *http://www.mrv.com/product/MRV-LD-OAB*, accessed in Mar. 2010.
[29] M. Mestre et al., "Tuning characteristics and switching speed of a modulated grating Y structure laser for wavelength routed PONs,"in *Proc. ANIC 2010*, Karlsrhue, Germany, Jun. 2010.
[30] M. Zannin et al., "On the benefits of optical gain-clamped amplification in optical burst switching networks, ", *IEEE/OSA J. Lightwave Technol.*, vol. 27, no. 23, pp. 5475-5482, Dec. 2009.
[31] Z. Rosberg, Hai Le Vu, M. Zukerman, J. White,"Performance analyses of optical burst-switching networks,"*IEEE J. Sel. Area. Comm.*, vol. 21, no. 7, pp. 1187-1197, Sep. 2003.
[32] M. Klinkowski, M. Pióro, D. Careglio, M. Marciniak, and J. Solé Pareta, "Non-Linear optimization for Multi-Path source routing in OBS networks,"*IEEE Commun. Lett.*, vol. 11, no. 12, Dec. 2017.
[33] R. Syski, "Introduction to Congestion Theory in Telephone Systems, "*North-Holland*, 1960.
[34] M. Minoux, "Mathematical Programming: Theory and Algorithms,"*John Wiley & Sons*, 1986.
[35] M. Minoux, "Discrete cost multicommodity network optimization problems and exact solution methods,"*Annals of Operations Research*, vol. 106, no. 1, 2001.
[36] E. Salvadori et al., "Distributed optical control plane architectures for handling transmission impairments in transparent optical networks,"*IEEE/OSA J. Lightwave Technol*, vol. 27, no. 13, pp. 2224-2239, 2009.
[37] S. De Maesschalck et al., "Pan-Europeanoptical transport networks: an availability-based comparison,"*Phot. Netw. Commun.*, vol. 5, no. 3, pp. 203-225, May 2003.
[38] K.C. Claffy, G.C. Polyzos, and H.W. Braun, "Traffic characteristics of the T1 NSFNET backbone,"in *Proc. IEEE Infocom 93*, San Francisco, CA, USA, Apr. 1993.
[39] S. Orlowski, M. Pióro, A. Tomaszewski, and R. Wessly, "SNDlib 1.0 Survivable Network Design Library," in *Proc. INOC 2007*, Spa, Belgium, 2007.
[40] M. Klinkowski, D. Careglio, J. Solé-Pareta, and M. Marciniak, "Performance overview of the offset time emulated OBS network architecture,"*IEEE/OSA J. Lightwave Technol.*, vol. 27, no. 14, pp. 2751-2764, Jul. 2009.
[41] J. Y. Wei, and R. I. McFarland, "Just-in-time signaling for WDM optical burst switching networks,"*IEEE/OSA J. Lightwave Technol.*, vol. 18, no. 12, pp. 2019-2037, Dec. 2000.
[42] J. S. Turner, "Terabit burst switching,"*J. High Speed Netw.*, vol. 8, no. 1, pp. 316, Mar. 1999.
[43] O. Pedrola et al., "JAVOBS: a flexible simulator for OBS network architectures,"*J. Netw.*, vol. 5, no. 2, pp. 256-264, Feb. 2010.
[44] IBM ILOG CPLEX: *http://www-01.ibm.com/software/integration/optimization/cplex/*, accessed in May 2010.

Oscar Pedrola received the M.Sc. degree in telecommunications engineering and the M.Sc. degree in information and communication technologies both from the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain in July 2008. He is a PhD student at the UPC, in the Broadband Communications Research Group (CBA). At present, he is involved in both the FP7 Network of Excellence BONE and the FP7 ICT STRONGEST project. His current research interests are in the field of optical networks with emphasis on burst/packet based switching technologies, network modeling, design and performance analysis.

Davide Careglio received the M.Sc. and Ph.D. degrees in telecommunications engineering from the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, in 2000 and 2005, respectively, and the Dr.Ing. degree in electrical engineering from Politecnico di Torino, Turin, Italy, in 2001. He is an Associate Professor in the Department of Computer Architecture at UPC. Since 2000, he is a staff member of the CCABA (www.ccaba.upc.edu) and of the Broadband Communications Research (CBA) group (www.cba.upc.edu). Together with his colleagues and students he has authored more than 100 peer-reviewed articles. His research interests are in the fields of optical networks with emphasis on packet-based switching technologies, quality of service (QoS) provisioning, energy-efficiency, and multi-layer traffic engineering.

Miroslaw Klinkowski received the M.Sc. degree from Warsaw University of Technology, Warsaw, Poland, in 1999, and the Ph.D. degree from Universitat Politècnica de Catalunya (UPC), Barcelona, Catalonia, Spain, in 2008. He is an Assistant Professor in the Department of Transmission and Optical Technology at the National Institute of Telecommunications, Warsaw, Poland, and is a Collaborating Researcher at UPC. His publications include several book chapters and more than 60 papers in relevant research journals and refereed international conferences. He has participated in many European projects dealing with topics in the area of optical networking. He is currently involved in the COST IC0804 action. His research interests include optical networking with emphasis on network modeling, design, and optimization.

Josep Solé-Pareta received the M.Sc. degree in telecom engineering in 1984, and the Ph.D. degree in computer science in 1991, both from the Technical University of Catalonia (UPC), Barcelona, Spain. In 1984 he joined the Computer Architecture Department of UPC. Currently he is Full Professor with this department. He did a Postdoc stage (summers of 1993 and 1994) at the Georgia Institute of Technology. His publications include several book chapters and more than 100 papers in relevant research journals, and refereed international conferences. His current research interests are in Nanonetworking Communications, Traffic Monitoring, Analysis and High Speed and Optical Networking and Energy Efficient Transport Networks.