Noname manuscript No.
(will be inserted by the editor)

Capacity Enhanced Receivers for Low Latency Burst Optical Slot Switching Rings

Bogdan Uscumlic · Annie Gravey · Philippe Gravey · Yvan Pointurier · Michel Morvan

Received: date / Accepted: date

Abstract We propose a new receiver architecture for coherent detection in slotted optical packet switching rings with elastic (rate adaptive) optical transponders. Such rings are a candidate solution for future datacenter and metropolitan networks. The new receiver can detect more than a single packet per time slot and consequently has higher flexibility (translating into higher supported capacity, or, equivalently, lower end-to-end latency, or a combination or both), at the cost of a moderate increase in the transponder complexity and energy consumption (less than 10%). We apply network planning and traffic engineering simulation tools (which we validate on small examples using theoretical models) to quantify the increase in network capacity and latency reduction that can be achieved thanks to the use of the new receivers. Finally, we identify the stability problem of the insertion process in the rings with the coherent receivers, and propose a polynomial network planning algorithm, for the case of fast-tunable transmitters. We evaluate the cost of enforcing the stability, in terms of the additional transponders needed, for the mentioned case.

Keywords Optical packet switching · metropolitan rings · coherent detection · computer simulation · network planning · stability · performance evaluation

B. Uscumlic · Y. Pointurier
Alcatel-Lucent Bell Labs, Nozay, France
Tel.: +33-1-60404067
Fax: +33-1-60406175
E-mail: *firstname.lastname*@alcatel-lucent.com

A. Gravey · P. Gravey · M. Morvan
Institut-Mines Télécom, Télécom Bretagne, Brest, France
Tel.: +33-2-29001111
Fax: +33-2-29001000
E-mail: *firstname.lastname*@telecom-bretagne.eu

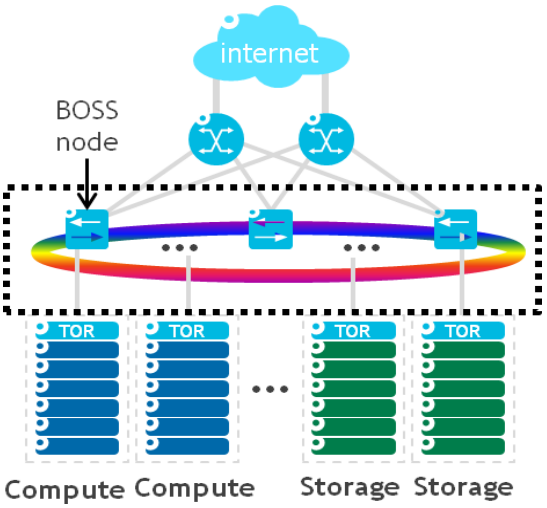

Fig. 1 An example of the application of BOSS rings for intra-data center connection.

1 Introduction

Optical packet switching leverages high bandwidth efficiency, network flexibility and low energy consumption, making this technology a viable candidate for future metropolitan and data center intra-connection networks. In this paper, we focus on a particular technology employing optical packet switching that is called Burst Optical Slot Switching (BOSS) rings where packet duration is fixed [1]. Such rings can be used, for instance, for the intra-connection of the switching equipment within a data-center, as illustrated in Fig. 1. BOSS rings interconnect e.g. the End of Row (EOR), or the Top of Rack (TOR) switches (the case shown in Fig. 1).

These rings were widely studied, from the point of view of the optical transmission technology, e.g. see [2,

3], and from the point of view of achievable network performance [4–6]. Many works have shown the advantages of BOSS rings over the optical circuit switching technologies. In particular, in [7] it is shown that optical traffic grooming allows to minimize the number of wavelengths used in the ring and to benefit from the capacity increase thanks to the statistical multiplexing.

In this paper, we propose an enhancement of the coherent receiver used in BOSS rings, in order to increase network capacity and decrease network latency. More precisely, we propose to duplicate the hardware parts that allow the reception of several optical slots simultaneously. Hence, the proposed receiver has a greater (optical) receiving capacity but the same (electrical) client-side capacity compared with a standard receiver. The new hardware in the proposed receiver increases the transponder power consumption only minimally (e.g. less than 10% for simulatenous stripping of two slots, as shown below). Compared with the standard receiver, the proposed receiver significantly improves the network capacity and the insertion latency of its nodes, thanks to the receiver's capability to extract several slots from the network at the same time.

This paper quantifies the gain in network capacity and latency that are made possible by this transponder. Although the proposed receiver introduces the additional latency at slot extraction, the end-to-end latency savings due to the increased flexibility of the new receivers largely offset the extraction queueing delay, as will be seen further.

For the first time, we identify the stability problem in the optical slot switching rings with coherent detection technology. We propose a network planning solution for allocating the transponders to a ring with known traffic matrix that employs the fast-tunable transmitters and we evaluate the cost increase of the BOSS network when the stability conditions are applied.

The remainder of the paper is organized as follows. The next section is devoted to a short description of the existing technologies for data centers and metropolitan area networks, which are the network segments of interest for the application of the technology that is considered in the paper. In the Section 3, we describe the architecture of a BOSS ring, focusing first on standard transponders and then on our proposed transponder. We validate the simulator tool in Section 4. The quality of service performances are evaluated in Section 5. In the following Section, we define the stability problem, propose the corresponding network planning algorithm, and evaluate the cost of the stabiliy in BOSS rings. Finally, the concluding remarks are given in the last section.

Note that some of the results presented in this paper have been reported in a communication given at ONDM 2015 [15].

2 State-of-the art technologies for data center intra-connection networks and metropolitan area networks

Today, further growth of data center networks is imposing to look for very scalable, cheap and energy efficient interconnection network inside data centers. Indeed, although widely accepted for their high bisection bandwidth and limited latency, the Ethernet based data centers suffer from limited scalability of high-radix electronic switches [16] and have very high power consumption and cost [17].

Regarding metropolitan area networks, the traditional solution is based on SONET/SDH architectures, whose successors are the OTN networks, employed over WDM multiplexed wavelength channels. In the following we provide an overview of the existing technologies for data center intra-connection and metropolitan area networks, which are the network segments to which the technology considered in this paper can be applied.

2.1 Data center intra-connection networks

Standard solution for data centers are the solutions based on Ethernet, electronic switching networks. Their limitations are the scalability, the power consumption and the installation cost (OPEX and CAPEX costs), since the entire traffic is processed electronically at each intermediate hop of the network.

The second group of solutions reposes on optical packet/burst switching (OPS/OBS) technologies. These solutions (e.g. [17,18,8,19]) are usually more efficient than standard solutions in terms of energy consumption, because they avoid using the large electronic switches and reduce the amount of traffic that is subject to optical-electronic-optical (OEO) conversion. Thanks to its high speed interfaces and the optical transparency of transit traffic, the optical packet switching is a viable solution for reducing the CAPEX cost of the future network inside of data centers, to improve its scalability and reduce its OPEX cost. In the same time, although the new optical packet switching network will probably have different intra-connection topology than the Ethernet based data center, the novel all-optical data center shall be able to provide the same or improved service flexibility, QoS management and virtualization benefits w.r.t. the currently available solutions, which

is a research challenge of many research teams working on this topic.

Third group of solutions are the hybrid solutions, which either combine electronic packet switching with optical circuit witching [20], or combine optical packet switching and optical circuit switching [21,22]. The optical circuits are often introduced to provide a quick, optical "bypass" between different groups of end hosts, in order to offload the large traffic flows from the main switching elements, and consequently to improve the latency.

2.2 Metropolitan area networks

2.2.1 Optical circuit switching solutions

Legacy solution is based on SONET/SDH rings [23], where the client traffic (ATM, Ethernet, IP packets,) is multiplexed into a set of hierarchical data rates. This solution is synchronized, and enables monitoring and fault detection, and fast protection within 50 ms.

The SONET/SDH solutions have as the successors the Optical Transport Network (OTN) [24] based solutions, which enable SONET/SDH alike OAM support for the DWDM networks.

Both previous solutions assume that the underlying optical switching technology is optical circuit switching, which has benefit of offering optically transparent bypass of transit traffic. However, once the wavelength channel/lightpath is reserved for the communication, it cannot be shared by several network nodes simultaneously. Temporal and spatial sharing of wavelengths is possible only with optical packet/burst switching.

2.2.2 Electronic packet switching solutions

Among the electronic switching solutions, there are Resilient Packet Ring [25], MPLS-TP [26], Ethernet based PBB-TE [27]. These solutions are based on electronic packet switching, and have sophisticated algorithms for supporting various aspects, such as addressing, traffic engineering, fairness, fast protection/restoration (within 50 ms), etc. The optical packet switching solution that we propose is compatible with these technologies. Indeed, previous technologies are usually transported via optical circuits, and thus, suffer from the same bandwidth inefficiency as SONET/SDH and OTN. Transporting these technologies over optical packet switching network would increase the network capacity.

2.2.3 Optical packet/burst switching solutions

Apart from Burst Optical Slot Switching (BOSS) from Bell Labs, there is a number of other optical packet/burst switching technologies for metropolitan rings. For detailed overview see, e.g. [5]. Here we describe some of them.

DBORN [28] is a bidirectional optical packet switching ring with slotted or un-slotted operation, where all the traffic passes through a special node, called hub. In difference to BOSS, in DBORN, the packet removal is done at the hub node, i.e. destination stripping is not allowed. DBORN is characterized with out-of-band control channel, and fixed receivers. The packets travel on the entire ring also in OPS rings DAVID [29] and in a variant of RINGO [30] technology. DAVID exists in passive and active architecture, and support or not the separation of upstream and downstream wavelengths. RINGO has 1 tunable laser per node, and fixed array of receivers. Another concurrent optical packet switching solution to BOSS is HORNET [31]: these rings are bidirectional and slotted and spatial reuse is leveraged. HORNET has tunable lasers, and the number of transmitters and receivers per node is fixed to $\lceil \frac{W}{n} \rceil$, where n is the ring size, and W is the number of wavelengths in the network. OPST [32] is an optical packet switching network proposed by Intune Networks. In this network, each destination has a dedicated wavelength, and the network is asynchronous, which makes this network different to BOSS. In OPST, the sources employ fast tunable lasers, and receivers are fixed.

Regarding the optical burst switching solutions, the important examples are TWIN [33] and OBTN [34]. TWIN is Time-Domain Wavelength Interleaved Network proposed by Bell Labs that is characterized with a passive network core, and active edge nodes which perform the scheduling optical bursts. The network is time-slotted, as in BOSS, but a single wavelength is allocated to each destination. TWIN and BOSS have been compared in terms of dimensioning cost in [35].

In OBTN, which stands for "Optical Burst Transport Network", each source is allocated a separate wavelength for the emission, and the network is synchronized. Sources employ fixed-tuned lasers, while destinations use burst mode receivers that can receive on desired wavelengths.

3 Burst Optical Slot Switching (BOSS) Ring

The physical topology of an BOSS network is a ring consisting of two parallel fibers, used to connect the network nodes. The two fibers are usually used in counter-

rotating directions, in a time slotted manner. One of the fibers can be used to carry backup traffic for fast failure recovery. In this paper, for the sake of simplicity, we consider unidirectional BOSS rings. Optical packets are used for the transport of encapsulated client data and are sent on different wavelength channels of wavelength division multiplexing (WDM) signal comb. A dedicated, possibly low-datarate (e.g., 10 Gb/s) wavelength control channel carries the headers (including slot source and destination information) of all synchronous slots on a separate wavelength, which is extracted, processed and re-inserted at every node. Each node consists of an electronic aggregation layer that encapsulates client frames into fixed-duration optical slots, one or several transponders (TRX) that effectively receive and transmit slots from/on the fiber medium, the aforementioned TRX for the control channel, and a "slot blocker" that can erase the received optical slots in order to make room for new inserted slots (see [2] for a possible implementation of the slot blocker). The transmitting side of a transponder is noted with TX and the receiving side with RX.

3.1 BOSS rings with direct detection receivers

The BOSS rings with direct detection receivers are considered in the previous works (e.g., see the work in [12]). In such studies, the network configuration where the nodes are equipped with fast-wavelength tunable transmitters and fixed wavelength receivers is analyzed. A BOSS node could either receive on a single wavelength or on a set of Z fixed wavelengths, in the case of the so-called "WDM receivers" (see e.g. [7]). In the latter case, the extraction latency appears due to the extraction queueing process of the received optical packets, since the "WDM receivers" can only temporarily receive on all Z wavelengths, while their effective capacity is limited to the single wavelength capacity. In the same time, the insertion latency in BOSS rings exists due to the queueing process of the optical packets waiting for the insertion to the ring. In [7] it is shown that the gain in terms of insertion latency is much more important than the penalty due to the extraction latency, when the direct detection "WDM receivers" are employed.

3.2 BOSS rings with coherent receivers

In this paper, we focus on the BOSS rings with coherent detection receivers. As discussed below, we consider either fixed or fast-tunable transmitters and different receiver flavors, depending on their complexity. In all cases, the receivers are able to extract a packet on any of the available wavelengths in the ring, which is a substiantial difference w.r.t. the BOSS ring configuration with direct detection receivers. Furthermore, the BOSS nodes are equipped with the TRXs being able to change the symbol rate and consequently their transmission rate. The supported datarates of the transponders are $\geq$100 Gbit/s.

3.2.1 Standard transponder

In standard configuration, the receiving side is able to extract a single packet per slot on any wavelength channel. The transmitting side can operate in one of two modes: either it can insert a single packet per slot on a fixed wavelength ("fixed TX"), or it can adjust the transmitting wavelength at every time slot ("fast-tunable TX"). Fast wavelength tunability at the TX side (demonstrated for instance in [3]) allows to maximize the network capacity, by equally sharing the inserted traffic over all available wavelengths in the network, i.e., to perform the so-called "load-balancing". Load-balancing is known to efficiently maximize the BOSS ring capacity, by minimizing the load on the wavelengths used in the network [7].

A standard coherent receiver is illustrated in Fig. 2(a) and consists of: a local oscillator, the coherent mixer, the photodiodes, analog-to-digital converters (ADC), followed by DSP processor. The entering signal is an optical WDM multiplex, while the received signal consist of demultiplexed electrical client frames. This RX architecture is called "standard RX". Thus, the standard RX is able to extract a single optical packet per time slot, from any of the available network wavelengths. A standard RX is fast wavelength tunable (using a fast wavelength-tunable laser as a local oscillator), whether the TX is fixed or fast wavelength tunable. Hence, in either case, fast wavelength selection can be performed at the RX side.

3.2.2 Proposed transponder

We propose a receiver architecture that is able to extract several optical packets simultaneously. This can be achieved by multiplying the number of optical front-ends, i.e., the hardware components that allow: light detection, coherent mixing with a local oscillator, and analog to digital conversion (ADC). In the remainder of the paper, we use N to note the number of optical front-ends of the new RX (the "receiver size"), ie. the number of optical packets that can be synchronously detected. The novel receiver for $N = 2$ is illustrated in Fig. 2(b). Obviously, when $N = 1$, the new receiver is equivalent to the standard one.

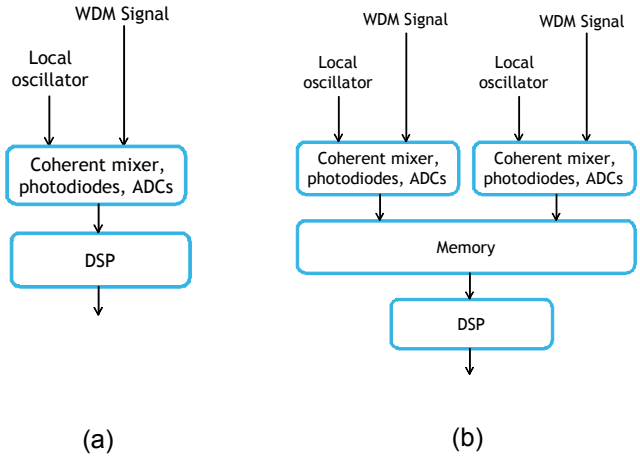

Fig. 2 Receiver architectures: (a) standard receiver; (b) proposed receiver for N=2 optical front-ends.

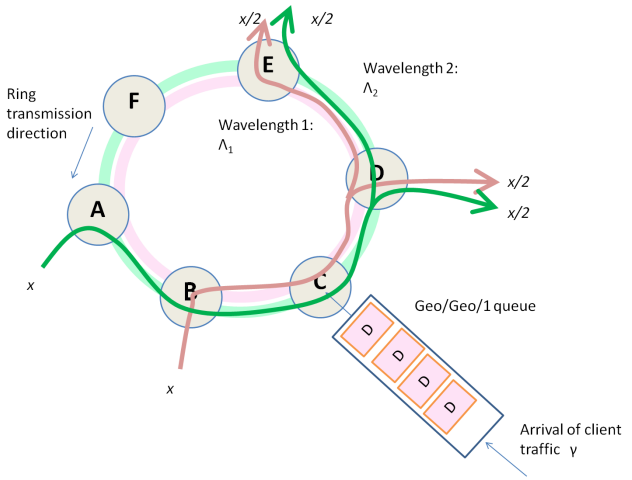

Fig. 3 Validation scenario: node C inserts traffic γ to node D in a ring network with 2 wavelengths, each loaded with traffic x.

Although the proposed DSP is able to receive data in the optical domain at rate $N \cdot C$ (where C is the capacity of a standard TRX), the effective capacity as seen from the electrical client layer is only C, because the proposed TRX has a single DSP unit with same capacity as that of a standard TRX.

Observe that, compared with a standard RX, the proposed RX only requires duplication of the coherent mixer, photodiodes and ADCs and the addition of a fast memory. According to [11], the power consumption of a 100 Gb/s coherent receiver is around 351 W, including 6.6 W for the (fixed wavelength) local oscillator, 1.6 W for the 4 photodiodes (and the associated transimpedence amplifiers), and 8 W for all 4 ADCs, amounting overall for 16.2 W (i.e., 5.5% of 351 W). Including for the fast driving circuit to make the local oscillators fast wavelength-tunable and for the additional memory, we expect the loss of energy efficiency of a TRX implementing RX according to our proposal to be limited to only 10%, w.r.t. a TRX implementing the standard RX.

Because of the additional memory, the new RX will introduce an extraction delay; however, we will show further that this additional delay is more than offset by a decrease in the insertion delay. In the following sections, we quantify both the insertion and extraction delays in the network, when new RX are used instead of standard RX.

4 Validating the ns2 based simulator

The computer simulations are performed by using an enhanced version the ns2 [9] simulator, to which we added the new functionalities. To validate this simulation tool we consider the simulation scenario in Fig. 3.

The BOSS ring in Fig. 3 is unidirectional, and contains 6 nodes, and 2 wavelengths Λ_1 and Λ_2. In this and following examples, the distance between the neighbor nodes is 100 time slots each of duration of 10 μs.

We consider four network configurations, depending on the TRX configurations per BOSS node:

- a single fixed transmitter and a single standard receiver are installed per BOSS node;
- a single fixed transmitter and a single new receiver are installed;
- a single fast tunable transmitter and a single standard receiver are installed;
- a single fast tunable transmitter and a single new receiver are installed.

The former two configurations allow each network node to emit a single optical packet on a predefined wavelength, while in the latter two configuration, a node can emit a single optical packet on any of the available wavelengths. In this paper, we consider that traffic is already aggregated into time slots. Matters pertaining to aggregation of client traffic into time slots are dealt with in [6], for instance.

Consider the following scenario, depicted in Fig. 3: node A sends to node D a traffic flow of amplitude $x/2$ (notation: $A \rightarrow D = x/2$). The other flows, with the same notation, are: $A \rightarrow E = x/2$, $B \rightarrow D = x/2$, $B \rightarrow E = x/2$ and $C \rightarrow D = \gamma$. Traffic x is normalized to the TX (or channel) capacity, and can be interpreted as the probability of optical packet arrival per time slot. This leads to Bernoulli assumption of traffic arrivals, that is validated in [7].

We observe the insertion process at node C, that is modeled with the Geo/Geo/1 queue as in [7]. We are

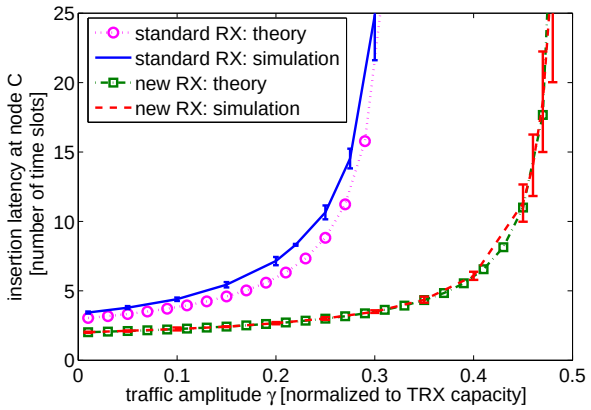

Fig. 4 Results for validation scenario with Fixed transmitters.

interested by the maximum amount of traffic γ that this node can insert such that losses and latency are bounded. Such value of γ can be approximated with the service probability of queue at node C. The queueing latency at node C is also reported in this section. Note that the entire traffic that node C inserts to the ring is destined to node D.

Two different ring configurations are considered: with fixed-wavelength (Section 4.1) or with fast wavelength-tunable transmitters (Section 4.2), when network is equipped either with standard TRX and with our proposed TRX.

4.1 Fixed transmitters

We suppose that each node has a single fixed transmitter (TX). It is also supposed that nodes B and C insert on Λ_1, and node A on Λ_2.

Two subcases are possible:

4.1.1 Each node has a single "standard" receiver

Standard receiver can receive a single packet per time slot on any wavelength.

The service probability μ_{S1}, that the incoming slot will be free for service at node C can be calculated as follows:

$$\mu_{S1} = 1 - P_{NE}, \tag{1}$$

where P_{NE} is the probability that there is no emission possibility for node C in the incoming time slot. Obviously, P_{NE} can be calculated as follows:

$$P_{NE} = P_{\Lambda_1 B} + P_{\Lambda_1 F \cap \Lambda_2 B, D}, \tag{2}$$

where $P_{\Lambda_1 B}$ is the probability that wavelength 1 is "busy", and $P_{\Lambda_1 F \cap \Lambda_2 B, D}$ is the probability that wavelength 1 is "free", while in the same time, wavelength 2 is "busy" with a packet towards node D. Indeed, note that if wavelength 2 already contains a packet towards node D, no slot can be sent in parallel on wavelength 1, which is the wavelength used by node C, since node D cannot receive more than a single packet per time slot, and, according to our scenario, all packets sent by C are destinated to D.

Next, we have:

$$P_{\Lambda_1 B} = x. \tag{3}$$

$$P_{\Lambda_1 F \cap \Lambda_2 B, D} = P_{\Lambda_1 F | \Lambda_2 B, D} \cdot P_{\Lambda_2 B, D} \tag{4}$$

Here, $P_{\Lambda_1 F | \Lambda_2 B, D}$ is the conditional probability that wavelength 1 is "free", given that the wavelength 2 is busy with a packet towards D. $P_{\Lambda_2 B, D}$ is the probability that Λ_2 is busy with a packet toward D.

We know that, according to the introduced notation:

$$P_{\Lambda_2 B, D} = x/2. \tag{5}$$

Let us now calculate $P_{\Lambda_1 F | \Lambda_2 B, D}$. If Λ_2 is busy with a packet towards D (which is inserted by node A), the probability that Λ_1 is free depends on the insertion process at node B, which inserts on Λ_1. Note that node B cannot insert a packet towards node D, if Λ_2 is busy with a packet towards D. In this case, we approximate the probability that node B will not insert a packet on Λ_1, with the ratio of probability that no packet arrives at this node over the probability that no packet has been sent towards node D.

Thus:

$$P_{\Lambda_1 F | \Lambda_2 B, D} = \frac{1 - x}{1 - x/2}, \tag{6}$$

and finally:

$$\mu_{S1} = 1 - (x + \frac{1 - x}{1 - x/2} \cdot x/2) \tag{7}$$

For $x = 0.5$, we obtain $\mu_{S1} \approx 0.335$. The theoretical curve, for latency δ (in number of time slots), in this case is obtained with the formula for Geo/Geo/1 queue with "arrival first" policy:

$$\delta = \frac{1 - \gamma}{\mu_{S1} - \gamma}. \tag{8}$$

Both theoretical result and the ns2 simulation curve are plotted in Fig. 4 for "standard RX". We see a pretty good accordance of the simulation and analytical results.

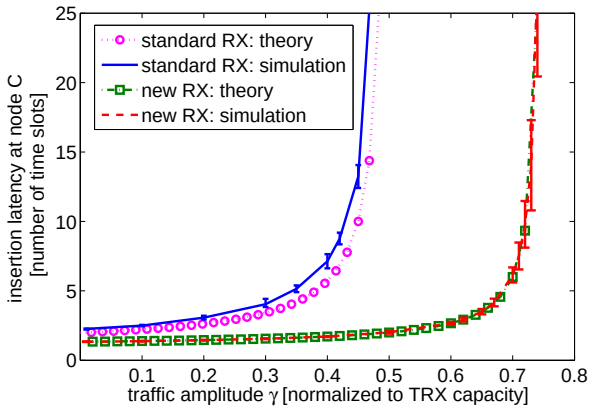

Fig. 5 Results for validation scenario with Fast tunable transmitters.

4.1.2 Each node has a single "new" receiver

Each node is now equipped with a new RX (of size 2), being able to receive two packets per time slot on any wavelength.

The service probability is now:

$$\mu_{N1} = 1 - P_{NE} = 1 - x. \tag{9}$$

For $x = 0.5$ we get $\mu_{N1} = 0.5$, which means:

$$\mu_{N1} \geq \mu_{S1} \tag{10}$$

The capacity gain is more than 30% in this case.

The theoretical results for the insertion latency and the simulation results for fixed TX and new RX are plotted in Fig. 4 for the same x. There is a very high accordance of the simulation and analytical results.

4.2 Fast tunable transmitters

Now consider that each node has a single **fast-tunable** transmitter (TX). Each node at transmission equally splits the traffic over both wavelengths, i.e. the load-balancing is performed, which is the main difference w.r.t. to case with "fixed transmitters".

4.2.1 Each node has a single "standard" receiver

The service probability μ_{S2}, that the incoming slot will be free for service can be calculated as follows (taking into account that the BOSS nodes can insert packet only on both wavelengths) is :

$$\mu_{S2} = 1 - P_{NE}, \tag{11}$$

where

$$P_{NE} = P_{\Lambda_1 B \cap \Lambda_2 B} + P_{\Lambda_1 F \cap \Lambda_2 B, D} + P_{\Lambda_2 F \cap \Lambda_1 B, D}. \tag{12}$$

Here, $P_{\Lambda_1 B \cap \Lambda_2 B}$ is the probability that both wavelengths (Λ_1 and Λ_2) are "busy". $P_{\Lambda_1 F \cap \Lambda_2 B, D}$ is the probability that wavelength 1 is "free", while in the same time, wavelength 2 is "busy" with a packet towards node D. Similarly, $P_{\Lambda_2 F \cap \Lambda_1 B, D}$ is the probability that wavelength 2 is "free", while in the same time, wavelength 1 is "busy" with a packet towards node D.

$$P_{\Lambda_1 F \cap \Lambda_2 B, D} = \frac{1-x}{1-x/2} \cdot x/2, \tag{13}$$

like in the previous case. Furthermore, because of the symmetry of traffic matrix, we have $P_{\Lambda_2 F \cap \Lambda_1 B, D} = P_{\Lambda_1 F \cap \Lambda_2 B, D}$.

Next:

$$P_{\Lambda_1 B \cap \Lambda_2 B} = P_{\Lambda_1 B | \Lambda_2 B} \cdot P_{\Lambda_2 B}, \tag{14}$$

where $P_{\Lambda_1 B | \Lambda_2 B}$ is the probability that Λ_2 is busy knowing that Λ_1 is busy, and $P_{\Lambda_2 B}$ is the probability that Λ_2 is busy. Obviously,

$$P_{\Lambda_2 B} = x. \tag{15}$$

If Λ_2 is busy, it can be busy with equal probability either with a packet towards D or towards E. If it is busy with a packet towards D, the probability that node B will insert a packet on Λ_1 can be approximated with the ratio of the probability that, at this node, no packet arrives towards E over the probability that no packet arrives towards D. If Λ_2 is busy with a packet towards node E, because of the symmetry the same probability shall be accounted for. Thus,

$$P_{\Lambda_1 B | \Lambda_2 B} = 0.5 \frac{x/2}{1-x/2} + 0.5 \frac{x/2}{1-x/2} = \frac{x/2}{1-x/2}. \tag{16}$$

Finally, we have:

$$\mu_{S2} = 1 - \left(\frac{x/2}{1-x/2} \cdot x + 2 \cdot \frac{1-x}{1-x/2} \cdot x/2 \right) = 1 - x. \tag{17}$$

For $x = 0.5$, we get $\mu_{S2} = 0.5$. For this value of x, the analytical and simulation curves for the latency of Geo/Geo/1 queue at node C, are plotted in Fig. 5, for the case of fast tunable TX and standard RX. We see an excellent concordance of the simulation and analytical results.

Note that the results of the equation (17) do not depend on the number of wavelengths in the ring. This is due to a general property that will be formulated in Section 6.

4.2.2 Each node has a single "new" receiver

The service probability is now:

$$\mu_{N2} = 1 - P_{NE}. \tag{18}$$

Now, P_{NE} is equal to the probability that both wavelengths are busy, in the same time.

$$\mu_{N2} = 1 - x \cdot x. \tag{19}$$

For $x = 0.5$, we get $\mu_{N2} = 0.75$. Once again,

$$\mu_{N2} \geq \mu_{S2}, \tag{20}$$

the capacity of the network is increased by about 33% thanks to the use of the proposed receivers.

The analytical and simulation results have a very high accordance, as seen in Fig. 5 for fast tunable TX and new RX and $x = 0.5$, which validates our simulator.

5 Quality-of-Service Performance Evaluation for Coherent BOSS rings

After validating the simulator, in this section we perform a set of extensive simulations, in order to quantify the capacity and quality of service (QoS) gains, in the network using new RX instead of standard RX.

The simulation study is based on the use of two tools. First tool is the optimal network dimensioning solution for elastic BOSS rings, based on Mixed Integer Linear Programming (MILP), and published in [10]. The MILP formulation is implemented in CPLEX optimization software. The second tool is the customized ns2-based computer simulator.

The following diagrams are the output of both simulation tools. The simulation operations are performed in the following order:

1. random traffic generation by a dedicated algorithm;
2. network dimensioning with CPLEX, by using the MILP formulation;
3. for the obtained network configuration, we perform the ns2 event driven computer simulation to measure either the insertion or extraction latency of traffic flows, or both.

Here, we focus on a 6-node network with 40 wavelengths.

The traffic is generated randomly, by choosing the connections ("traffic flows") and their amplitudes uniformly, until the total network load is reached. Each "traffic flow" is defined by the unique (source, destination) pair. In our experiments, the overall network load is chosen from [100, 500] Gbit/s, and the TX capacity is supposed to be 100 Gbit/s. The traffic distribution is biased to ensure that the entire sent or received capacity is not greater than a single TRX. In another words, we assume that each network node has a single TX and a single RX. More precisely, the amplitudes of traffic flows are uniformly chosen in sets [0, σ] Gbit/s, where $\sigma \in [10,\ 90]$ Gbit/s are chosen to satisfy the TRX capacity limitation. Finally, we discard the traffic flows that are lower than 1 Gbit/s.

For such generated traffic, we run the network dimensioning tool from [10], that has the objective of minimizing the number of transponders in the BOSS ring. We adapt constraint (11) in [10] to limit the single wavelength occupancy to 90% (this is important in order to try to avoid the congestion in the network).

The previous network design yields a solution for the routing and wavelength assignment (RWA) problem. We then map this solution into our ns2 simulator to assess the network quality of service (QoS) performances. The diagrams presented in this section measure either the sum of insertion and extraction latency, averaged over all network nodes, or the extraction latency separately. All reported results are given with the confidence interval of 10%, achieved at the confidence level of 95%.

5.1 Impact of the new receiver on network end-to-end latency

The first set of results, shown in Tab. 1, shows the impact of the new receiver on the end-to-end latency. Note that the end-to-end latency of a traffic flow is the sum of the insertion latency and the extraction latency[1].

The insertion latency is due to the queuing process at each nodes that has some optical packets to insert. As previously, we suppose a single FIFO queue in front of each network node, that we describe with Geo/Geo/1 model.

The extraction latency exists due to the memory in the new RX. We suppose that each optical packet needs 1 time slot of processing before reaching the client layer, in which case the queueing process at the reception can be described with a nGeo/D/1 queue. (The same queueing model for the extraction process, but in the case of "WDM receivers" with direct detection, is used in [7].) Accordingly, we suppose that the extraction latency is exactly 1 slot in case of standard RX. Note that, as far as the finctionality is concernet, the standard RX can be considered as a special case of new RX, having a single optical front-end, i.e. obtained for N=1.

[1] Since the latency due to propagation is fixed by the topology, it is not accounted for in our results.

Table 1 The sum of **the insertion and the extraction latency** (in number of time slots), for different network configurations and different values of total traffic load

Traffic Load [Gbit/s]→ Configurations↓	100	200	300	400
fixed TX + standard RX	2.09	2.33	(packet losses)	
fixed TX + new RX	2.09	2.33	3.34	4.14
fast tunable TX + standard RX	2	2	(packet losses)	
fast tunable TX + new RX	2	2	2.65	3.8

Table 2 The extraction latency (in number of time slots), for different network configurations and different values of total traffic load

Traffic Load [Gbit/s]→ Configurations↓	100	200	300	400
fixed TX + standard RX	1	1	(packet losses)	
fixed TX + new RX	1	1.25	1.68	2.46
fast tunable TX + standard RX	1	1	(packet losses)	
fast tunable TX + new RX	1	1	1.62	2.65

The results reported in Tab. 1 show the sum of the insertion and extraction latencies for the following TX+RX combinations (at each ring node):

- fixed wavelength TX + standard RX;
- fast wavelength tunable TX + standard RX;
- fixed wavelength TX + new RX;
- fast wavelength tunable TX + new RX.

From the table, we can see that at lower loads (up to 200 Gbit/s), the network has the same performance whether standard or new RX are used. However, for medium and high network loads (starting from the loads of 300 Gbit/sec), the sum of the insertion and the extraction latency in the network employing the new RX only slightly increases, while when standard RX are used, the network enters into the zone of instability. When network is the instability zone, some of its nodes have the insertion queues that are congested, and such nodes experience very high latencies and losses of optical packets. Since some packets are lost, the latency results for the instable loads are not shown in Tab. 1; instead, we mark the instability zone with "(packet losses)". These results clearly suggest that the network has much higher capacity when the new receivers are used. The use of new RX allows network to stay in the zone of stability, even at high loads. Consequently, the latency is reduced, and limited to only few time slots.

Note that the instability of the insertion queues is due to the packet nature of the traffic and can be prevented by the proper network dimensioning. One way to enforce the stability is to reduce the allowed occupancy of the wavelength channels. In the presented example the wavelength channels where allowed to be filled up to 90%. Obviously, this limitation is not enough to guarantee the stability of the network when standard RX are used. Note that here, we use the network dimensioning tool from [10], that does not enforce the network stability. The solution of the problem of stability in coherent BOSS rings in a case of fast-tunable transmitters, is the topic of the Section 6.

5.2 Impact of the new receiver on network extraction latency

Tab. 2 shows the extraction latency only, for the same scenario. The simulations confirm that when standard RX are used, latency is equal to exactly 1 time slot. (Once again, we report only the results for the loads where the packet losses where not observed.) Interestingly, the extraction latency is very small for new RX (equal to only few time slots), and is obviously not a limiting factor for the network to benefit from the use of novel receiver architecture.

5.3 Impact of the new receiver architecture on network capacity

Next, we were interested to quantify the effective capacity gain of the BOSS ring using new RX instead of standard RX. To evaluate this gain, we simulate once again the network with standard RX, and measure the insertion and extraction latency in the network. This time, we scale the inserted traffic by a "capacity scaling coefficient α," $(0 < \alpha < 1)$. The simulations are stopped when the value of α is found, at which a network with standard RX achieves the same latency figures (within only few percent of difference) as the network with new RX. For the reference latencies of the network with new RX, we use the values already reported in Tab. 1. The effective relative capacity gain G in the network (expressed in %), thanks to the use of new RX, corresponds then to $G = \frac{1-\alpha}{\alpha} \cdot 100$.

Fig. 6 presents the obtained results for the effective capacity gain G at different network loads. With the increase of load, the values of G increase, which is expected, since the gains in latency are higher, as shown in Tab. 1. The highest value for G is ≈42%, corresponding to α of 0.7. The gains in capacity are higher when TX leveraging fast-tunability are used, meaning that the

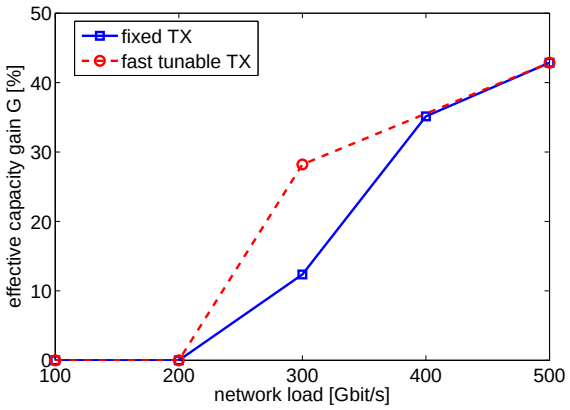

Fig. 6 The effective capacity gain in the network G, achieved thanks to the use of the new receivers, for different traffic loads.

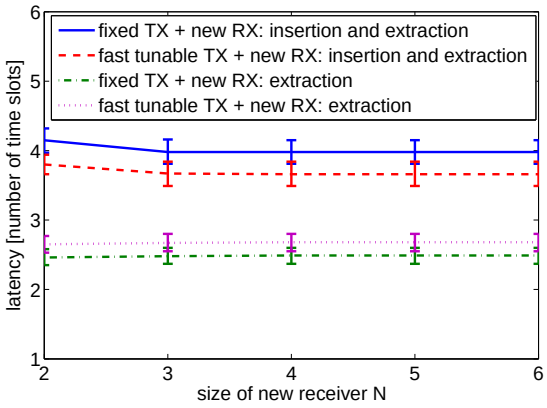

Fig. 7 Impact of the number of optical front-ends N (in the new receiver) on the network latency.

additional network flexibility (coming from this property of the transmitters) favors the savings that can be achieved by the new receiver architecture.

5.4 Impact of the size and number of the new receivers on the network latency

The number N of optical front-ends in the proposed receiver also impacts the network performance. This impact is evaluated in Fig. 7, for a fixed network load of 400 Gbit/s. Interestingly, N=2 already gives excellent results. Further increasing N only slightly decreases the insertion and extraction latency compared with N=2. Note that the results for the standard RX (obtained for $N = 1$) are not presented in Fig. 7, since for such receivers (at the same traffic load) the network is unstable, as shown in Tabs. 1 and 2.

Note that in this and previous diagrams, the fast-tunability decreases insertion latency but not extraction latency.

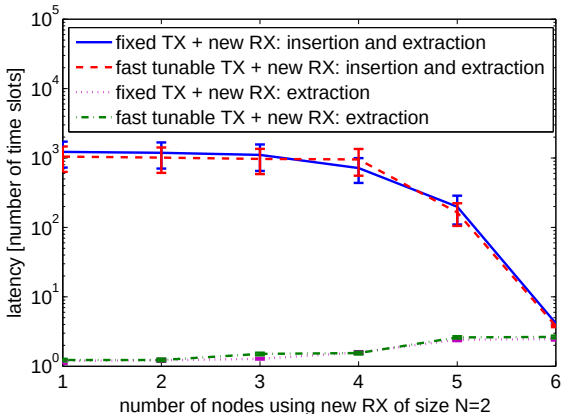

Fig. 8 Impact of the number of nodes using new receivers (N=2) on latency.

In the final set of simulations, we were interested to see what happens if only some of the nodes use the new RX. The latency versus the number of nodes that use new RX with N=2 front-ends, for the fixed total load of 400 Gbit/s is illustrated in Fig. 8. The nodes that use novel RX are randomly chosen, while the other nodes use standard RX.

The results presented in this section suggest that the network has more and more available capacity, for the increased number of RX employed, and for the greater values of N.

6 The cost of the stability in BOSS rings

So far, we have seen that the stability issue has the important negative consequencies on the QoS performances in the BOSS rings employing the standard RX. In this section, for the first time, we solve the problem of stability and evaluate its cost, in the optical packet switching rings employing the fast-tunable transmitters, standard RX, and coherent detection.[2]

As suggested by the results of Section 4, i.e. the equation (17), in the case of fast-tunable transmitters, when standard RX are used, the service probability depends only on the transit traffic amplitude (x in the eq. (17)), and not on the number of wavelengths in the ring. This is resumed in the following lemma, that is then used as the starting point for the analysis in this section.

[2] The network planning algorithms enforcing the stability in the other configurations of BOSS rings with coherent detection, including the configurations with the new receivers, are left for future study. Note that the ad-hoc stability condition applied on the design of a network with new receivers (that consists in limiting the wavelength occupancy to 90%), seems to be sufficient for random traffic profile studied in this paper, as shown in Section 5.

Lemma 1. *In the case of fast-tunable transmitters, and standard receivers, at a given network node, the time slot is available for the transmission (toward the observed receiver) with the probability equal to* $1 - t$*, where* t *is the probability of traffic being sent to the same receiver, by the upstream nodes.*

Thanks to the Lemma 1, in the case of BOSS rings with fast-tunable transmitters we are able to derive the network planning algorithm, that enforces the stability conditions when dimensioning such networks.

6.1 Network planning and the stability issues

The optical network dimensioning/planning usually consist of addressing the routing and wavelength assignment (RWA) problem, and allocating the needed number of wavelengths, transponders and wavelength converters to the network nodes, in order to support the given traffic matrix. The given traffic matrix is usually over-estimated, but irrespective of that, the network configuration that is the result of the network dimensioning, shall provide the satisfactory network performance.

In case of optical circuit switching (OCS) network dimensioning, the additional guarantees are not needed to provide the satisfactory network performance: as long as routing and wavelength assignment problem is resolved, the network can have the stable operation for the given traffic matrix. It is because the transported circuits are created at the electronic domain: once created, the optical slots are inserted/served without waiting, with fixed service time.

In difference to the network planning of OCS network, the network planning of optical packet switching (OPS) networks is much less studied and understood. The major difference in dimensioning of OCS and OPS networks is in the stability problem that appears only in optical packet switching networks. The problem of stability relates to the insertion process, and is a direct consequence of the possibility given to an optical packet switching transmitter to choose between several insertion wavelengths or destinations (resources), when inserting the optical packet to the ring. In such system, the insertion, i.e. the optical packet service probability depends on the availability of each resource. The total insertion capacity cannot be calculated simply as a sum of all resource availabilities; its value is below the mentioned sum. Note that the time multiplexing nature of the transmitter enables high bandwidth efficiency in optical packet switching rings, but the stability problem is the price to be paid, for the capacity gains.

In the context of intensity modulated transmission, the stability conditions for OPS rings are studied in [12]. In the current paper, for the first time we consider the problem of the stability, and the related network cost increase, in OPS rings with coherent detection. We explain the particularities of the stability issue, w.r.t. the coherent detection scheme, and point out how to resolve the head-of-line blocking problem and preserve the work conserving operation. (The system is work conserving, if whenever there is a free slot for the emission, and the queued packets that are waiting for the insertion, a packet is actually sent.) The stability conditions are formulated, and the optimal dimensioning algorithm containing the stability conditions are given.

6.2 Destination queueing for avoiding the head-of-line contention in the BOSS rings

Note that, in a BOSS ring, the sharing of wavelengths by different destinations is enabled. This is a natural consequence of the use of the coherent detection, as each node can adjust its receiver to the desired receiving wavelength. The sharing of wavelengths enables high network flexibility, and has a very positive impact on the network quality-of-service (QoS) parameters.

However, simple FIFO queuing cannot be used for the insertion of traffic in an BOSS node, as this would provoke the head-of-line (HOL) blocking. Indeed, each destination can receive only a single optical packet per slot, per each receiver. If the BOSS node has a single receiver, the total number of received optical packets per slot is limited to 1.

If the optical packets would be queued in the same FIFO waiting line, the HOL packet could block the subsequent packets, if there are no transmission opportunities for the HOL packet, but such opportunity exist for any other, subsequent packet.

Note that in difference to the reasons for the HOL blocking in BOSS rings with direct detection receivers[12], here the destination provoked HOL blocking is the reason for the non work-conserving operation of the network.

Thus, to keep the work-conserving operation of the insertion process at BOSS node, and avoid the HOL blocking problem, an BOSS node in a network with coherent detection, needs to create virtual output queues, according to the destination of each optical packet (see Fig. 9).

In the following, we present a small motivation example explaining more closely how the stability issue increases the total network cost in an BOSS ring.

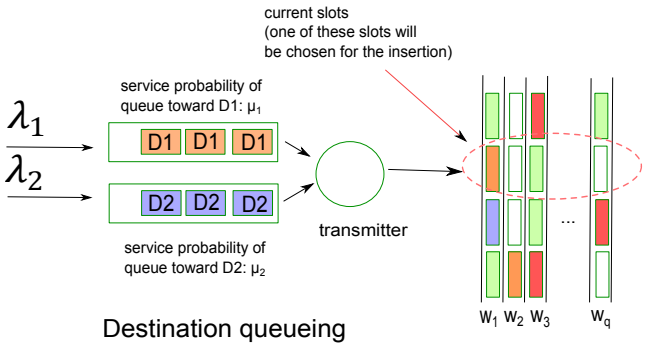

Fig. 9 Destination based queueing process at an arbitrary BOSS node.

6.3 Stability conditions

Let us observe a simple BOSS node, inserting the packets toward two downstream destinations D1 and D2 (Fig. 9). The destination queuing of the optical packets is applied, from the reasons mentioned above.

The system presented in Fig. 9 is a system of a single server (transmitter) and two parallel queues. There are q (e.g. $q = 40$) wavelengths in the system, on which the insertion is possible.

The scheduling decision performed by the server consists of two steps:

1. From which queue to select the optical packet for the insertion?
2. To which wavelength to insert the selected optical packet?

If we have a single transponder at BOSS node, the server capacity is equal to the insertion of 1 optical packet per time slot (on any available wavelength). Let us also suppose in this example, that each downstream destination has a single receiver. Thus, D1 or D2 can extract (receive) at most 1 optical packet per time slot (on any wavelength).

The queues are described with the following parameters:

- Let λ_1 be the traffic arrival rate to the queue D1. It is defined as the probability of optical packet arrival (creation) toward destination D1, at the current (observed) time slot (Fig. 9).
- Let λ_2 be the traffic arrival rate to the queue "D2". It is defined as the probability of optical packet arrival toward destination D2, at the current time slot.
- Let μ_1 be the service rate for the queue "D1". It is defined as the probability that among the "current slots" (Fig. 9) there is a "free" time slot, and that among the "busy" slots, no optical packet is already scheduled toward destination D1. Note that this probability depends on the transit traffic already sent to D1.
- Let μ_2 be the service rate for the queue "D2". It is defined as the probability that among the "current slots" there is a "free" time slot, and that among the busy" slots, no optical packet is already scheduled toward destination D2. Note that this probability depends on the transit traffic already sent to D2.

From the queueing theory it is known that the necessary conditions for the stability of the queues D1 and D2 are: $\lambda_1 < \mu_1$ and $\lambda_2 < \mu_2$, respectively. In another words, the necessary condition for the stability of the observed queue is that its service rate is strictly greater than its customer arrival rate ($\mu > \lambda$).

However, these conditions, although necessary, are not sufficient for the stability of the above queueing system. It is since in the above system, the queues share the same transmitter, and their insertion process cannot be observed independently.

From the studies of wireless networks, done by Tassiulas & Ephremides [13], we know that the necessary and sufficient stability conditions for the insertion at the observed BOSS node with 2 destinations are:

$$\lambda_1 < \mu_1,\ \lambda_2 < \mu_2,\ \lambda_1 + \lambda_2 < \mu_1 + \mu_2 - \mu_1\mu_2. \qquad (21)$$

Note that Tassiulas & Ephremides have shown [13] the previous result for a system with two queues, when the Longest Queue First queue selection is assured. Here, the "queue selection" is equivalent to the "destination selection", which corresponds to the step 1 of the previously defined "scheduling" decision in a BOSS node. The step 2 of the decision is the selection of wavelength, which can be done e.g. randomly, by following the uniform distribution.

If the traffic arrival and service probabilities do not satisfy the eq. (21), the insertion queues become unstable, and the optical packets are lost from the insertion buffers.

Let us consider the following numerical example: We suppose that the packet toward destination D1 can be inserted with a 50% chance, i.e. $\mu_1 = 0.5$. In another words, with probability 0.5, there is an insertion opportunity for a packet queued at head-of-line position of queue D1. Similarly, let us suppose $\mu_2 = 0.5$, for the service probability of queue D2. According to the conditions (21), we get:

$$\lambda_1 < 0.5,\ \lambda_2 < 0.5,\ \lambda_1 + \lambda_2 < 0.75. \qquad (22)$$

The last condition in (22) is particularly interesting: it states that the sum of all the insertion traffic shall not be greater than 75% of the transponder capacity.

If the BOSS node tries to insert the values of traffic that are higher than what is allowed by (22), the

queues on Fig. 9 will became unstable and the optical packets will be lost (since these queues are implemented in electronics and are thus limited).

In another words, we have seen that the stability issue decreases the network capacity. Consequently, the stability issue increases the cost of the BOSS network. For instance, if we suppose that the BOSS node in Fig. 9 needs to insert the capacity $\lambda_1 = 0.4$ and $\lambda_2 = 0.4$ (in total: $\lambda_1 + \lambda_2 = 0.8 > 0.75$), the system will be unstable, according to (22).

The only solution for the system to be able to accept this traffic is to increase the number of receivers (transponders) at one of the destination nodes D1 and D2 (or at both of them). By increasing the number of receivers at D1 (D2), the service probabilities μ_1, μ_2 will be greater, and the stability zone of the system will be expanded. However, the cost of the resulting network will be increased. In this section we evaluate the cost increase due to the stability conditions, in terms of the number of additional transponders needed to enable a stable network design. To do this, we propose a network dimensioning algorithm that takes into account the stability conditions of the network.

The formulas found by Ephremides and Tassiulas can be generalized for the system of n queues (corresponding to n destination queues in our network):

$$\sum_{i \in Q} \lambda_i < 1 - \prod_{i \in Q} (1 - \mu_i), \quad \forall Q \in \{1, 2, ..., n\}, \tag{23}$$

where λ_i and μ_i are the traffic arrival and traffic service probability, toward the destination i. Note that in order to guarantee the stable network operation, the previous stability conditions need to be satisfied for each BOSS node separately, and for each of its transmitters. Furthermore, n can differ from one BOSS node (transmitter) to the another.

Enforcing the stability conditions (23), and using the Longest Queue First scheduling policy for the queue selection (step 1 of the scheduling decision in a BOSS ring) has been shown to be throughput optimal in [13]. The work on the stability conditions was further extended by Stolyar [14], to show that the stable operation can be achieved for the MaxWeight scheduling policies (including the common policies, such as Oldest Packet First, and also the Longest Queue First). Note that the stability conditions proposed by Stolyar are different from those proposed by Tassiulas & Ephremides. In [12] we formulate a theorem, showing that the two stability conditions are equivalent. This means that the stability conditions (23) hold for any MaxWeight scheduling policy.

Note that when the stability conditions (23) are applied on the network configurations from Tab. 1 (or Tab. 2, in Section 5), the instability of the insertion process in the BOSS ring is confirmed at least at a single BOSS node, when the traffic load is greater or equal to 300 Gbit/sec.

6.4 The algorithm for the stable design of the coherent BOSS rings employing the fast-tunable transmitters

Here, we propose the network dimensioning algorithm, that enforces the stability conditions (23). We consider a unidirectional BOSS ring with set of nodes K, where the network nodes are equipped with the transponders, composed of fast-tunable transmitters and standard receivers. Thanks to the Lemma 1, the service probabilities are easily calculated in this case, while the traffic arrival probabilities are given by the traffic matrix T. T is composed of the traffic demands of the amplitude $t_{i,j}$ (normalized to the single channel capacity), between any two nodes i and j in the ring ($i \neq j$). Like in the traffic matrix considered in the Section 5, we consider a special case where $(\forall i \in K)(\sum_{j,j \neq i} t_{i,j} \leq 1, \sum_{j,j \neq i} t_{j,i} \leq 1)$.

The goal of the optimization algorithm is to verify whether for a given traffic matrix and set of nodes K, the BOSS ring is stable, and if it is not, to allocate the minimum number of additional transponders to the network nodes, so that the resulting network design is stable. More precisely, the algorithm allocates the additional receivers, that is equivalent to the allocation of the additional transponders, since each transponder is composed of a transmitter and a receiver part.

The pseudo-code of the network planning algorithm is given in Algo. 1. The following notation is introduced: $\pi(s, d)$ is the path in the ring, between the source node s and the destination node d; $TRXNumberAtNode(d)$ is the number of transponders at node d, and is the output of the algorithm.

In the case of a ring node instability, the Algo. 1 allocates randomly an additional transponder to a destination. It is supposed that each connection going toward some destination is equally shared between all the receivers of that destination. In the same time, adding the new receivers results in the increase of the number of the variables on which the stability conditions are calculated.

Note that the problem of finding a stable network configuration of a BOSS ring with direct detection receivers is addressed in [12]. Although the form of the stability conditions applied in the current algorithm and in the solution in [12] is the same, in the latter case the queues were organized per wavelength (and a wavelength could be possibly shared between different

nodes), while here we have per destination queues (and a destination may receive on any wavelength). Consequently, the two network planning problems are not equivalent.

```
//initialization:
for each ring node p ∈ {1,2,...,|K|} do
    TRXNumberAtNode(p)=1;
end
//main code:
while 1 do
    RingStable=true;
    for each ring node p ∈ {1,2,...,|K|} do
        InstabilityAtNode(p)=false;
    end
    for each ring node p ∈ {1,2,...,|K|} do
        for each receiver rx(i) of destination node i
        (i ∈ {1,2,...,|K|}, i ≠ p) do
            set the traffic arrival probabilities λ^p_{rx(i)}
            to t_{p,i} / TRXNumberAtNode(i);
            set the service arrival probabilities μ^p_{rx(i)}
            to 1 − (Σ_{s,i,p∈π(s,i)} t_{s,i}) / TRXNumberAtNode(i);
            apply the stability conditions (23) on the
            system described by λ^p_{rx(i)}-s and μ^p_{rx(i)}-s;
            if ring is unstable at node p then
                RingStable=false;
                InstabilityAtNode(p)=true;
            end
        end
    end
    if RingStable==false then
        while 1 do
            RingUpdated=false;
            for each ring node p ∈ {1,2,...,|K|} do
                if InstabilityAtNode(p)==true then
                    randomly choose destination d so
                    that t_{p,d} ≠ 0;
                    TRXNumberAtNode(d)+=1;
                    RingUpdated=true;
                    break;
                end
            end
            if RingUpdated==true then
                break;
            end
        end
    end
    if RingStable==true then
        break;
    end
end
```

Algorithm 1: Network Planning Algorithm for the Stable BOSS rings.

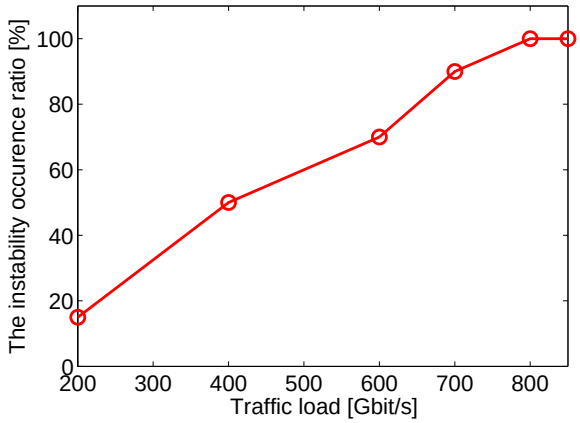

Fig. 10 The instability occurence ratio.

6.5 Evaluating the cost of the stability in BOSS rings

In this section, we report the results of network dimensioning, obtained by using the previously introduced polynomial algorithm.

The traffic matrix is generated according to the same procedure as in Section 5. The Algo. 1 is implemented in the Matlab software, and all the results in this section are the averages over 100 randomly generated traffic matrices. The TRX capacity of 100 Gbit/s is supposed

For a 10-node ring, the Algo. 1 is run to calculate the occurency ratio of the instability among 20 randomly generated traffic matrices. The results are summarized in Fig. 10. As it can be seen from this figure, the instability occurence ratio increases very quickly, and reaches 100%, starting from 80% traffic load. (Note that 80% of the maximum traffic load for a ring with 10 nodes and TRX capacity of 100 Gbit/s corresponds to 800 Gbit/s in Fig. 10)

For different ring sizes, and different levels of maximum traffic load, the results on the relative average increases in the TRX number in the network design, calculated by Algo. 1 are presented in Fig. 11. For high network load of 70%, the cost penalty due to the stability enforcing is as high as 50% of the total cost. (The cost increase values from Fig. 11 are obtained by using the formula $\frac{st-unst}{unst} \cdot 100[\%]$, where st and $unst$ respresent the numbers of TRX in the network with and without stability conditions, respectively.)

From the figure, we can see that increasing the traffic load increases the cost penalty gain. Even at low traffic load (of 20%), the minimum cost penalty is above 20%. Furthermore, the cost penalty is smaller for the higher number of ring nodes. Indeed, the number of TRX is greater in the larger networks, which results in the smaller relative cost penalty.

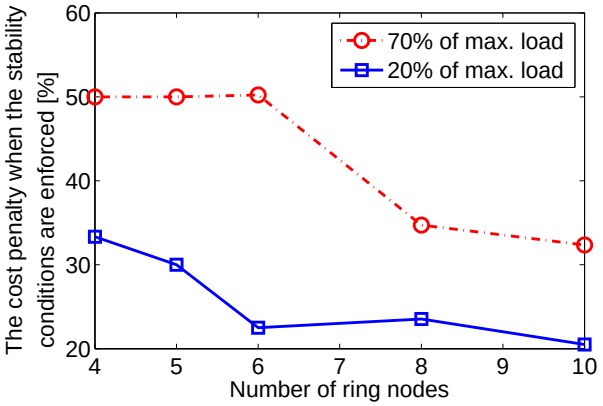

Fig. 11 The cost penalty when the stability conditions are enforced.

7 Conclusion

In this paper, we proposed a novel coherent receiver architecture with N optical front-ends but with same DSP and client-side capacity, such that N optical slots can be synchronously received. The duplicated components increase the TRX power consumption by less than 10%. Using detailed simulation study, where we used both network dimensioning and performance evaluation simulation tools, we showed that the use of the novel receivers significantly reduces the end-to-end latency of the network flows, by expanding the stability region of the insertion queueing process at network nodes. Furthermore, we showed that a network that uses our proposed TRX can carry up to 40% more traffic than a network using standard TRX. We showed that using more than 2 optical front-ends in the receiver does not further improve network latency.

For the first time, we assessed the network planning problem in the coherent optical packet switching rings when the stability conditions are enforced. The results show that the stability issue increases in 100% of cases, when the network is sufficiently saturated. Enforcing the stability conditions results in the increase of the number of needed transponders in rings for up to 50%. Note that, thanks to their high bandwidth efficiency, the optical packet switching networks remain much cheaper than optical circuit switching networks, in many applications, despite the previous cost increase.

Acknowledgements This work was supported by the CELTIC-Plus SASER SAVENET project.

References

1. D. Chiaroni et al.: Packet OADMs for the next generation of ring networks. Bell Labs Technical Journal, vol. 14, no 4, pp. 263-285 (2010).
2. G. de Valicourt, M.A. Mestre, L. Bramerie, J.-C. Simon, E. Borgne, L. Vivien, E. Cassan, D. Marris-Morini, J.-M. Fédéli, P. Jennevé, H. Mardoyan, Y. Pointurier, A. Le Liepvre, G.H. Duan, A. Shen, S. Bigo: Monolithic Integrated Silicon-based Slot-Blocker for Packet-Switched Networks. Proc. ECOC, Cannes, France, Sep. 2014, paper We.3.5.5.
3. J.E. Simsarian et al.: Fast-tuning 224-Gb/s intradyne receiver for optical packet networks. OFC/NFOEC (2010), paper PDPB5.
4. L. Sadeghioon, A. Gravey, B. Uscumlic, P. Gravey, M. Morvan: Full Featured and Lightweight Control for Optical Packet Metro Networks. Journal of Optical Communications and Networking, February 2015, vol. 7, n 2, pp. A235-A248
5. B. Uscumlic, I. Cerutti, A. Gravey, P. Gravey, D. Barth, M. Mprvan, P. Castoldi: Optimal dimensioning of the WDM unidirectional ECOFRAME optical packet ring. Photonic network communications, July 2011.
6. N. Benzaoui, Y. Pointurier, T. Bonald, J.-C. Antona: Impact of the electronic architecture of optical slot switching nodes on latency in ring networks. IEEE/OSA Journal of Optical Communications and Networking, vol. 6, no. 8, Aug. 2014, pp. 718–729.
7. B. Uscumlic, A. Gravey, P. Gravey, I. Cerutti: Traffic grooming in WDM optical packet rings. 21st International Teletraffic Congress (ITC'21), Paris, France, Sep. 2009
8. M.A. Mestre et al.: Optical slot switching-based datacenters with elastic burst-mode coherent transponders. Proc. ECOC, Cannes, France, Sep. 2014, paper Th.2.2.3.
9. http://www.isi.edu/nsnam/ns/
10. B. Uscumlic, Y. Pointurier, A. Morea, S. Bigo: On the Cost of Protection in Optical Slot Switching Rings with Elastic Transponders. Proc OFC, Los Angeles, California, USA, Mar. 2015, paper Th2A.46.
11. A. Morea et al.: Power Management of Optoelectronic Interfaces for Dynamic Optical Networks. Proc. ECOC, Geneva, Switzerland, Sep. 2011, paper We.8.K.3.
12. B. Uscumlic, A. Gravey, I. Cerutti, P. Gravey, M. Morvan: Stable Optimal Design of an Optical Packet Ring with Tunable Transmitters and Fixed Receivers. ONDM 2013 : the 17th International Conference on Optical Network Design and Modeling, 16-19 april 2013, Brest, France
13. L. Tassiulas, A. Ephremides: Dynamic server allocation to parallel queues with randomly varying connectivity. IEEE Transactions on Information Theory, vol. 39, pp. 466 478, March 1993.
14. A. L. Stolyar: MaxWeight Scheduling in a Generalized Switch: State Space Collapse and Workload Minimization in Heavy Traffic. Annals of Applied Probability, vol. 14, no. 1, pp. 153, 2004.
15. B. Uscumlic, Y. Pointurier, A. Gravey, P. Gravey, M. Morvan: Optical Receivers with Multiple Front-Ends for Low Latency Optical Slot Switching Rings. ONDM 2015 : the 19th International Conference on Optical Network Design and Modeling, 11-14 May 2015, Pisa, Italy
16. N. Binkert et al.: The role of optics in future high radix switch design. 38th Annual International Symposium on Computer Architecture (ISCA), pp.437,447, 4-8 June 2011
17. I. Cerutti et al.: Designing Energy-Efficient Data Center Networks Using Space-Time Optical Interconnection Architectures. IEEE Journal of Selected Topics in Quantum Electronics, vol.19, no.2, pp.3700209,3700209, March-April 2013.
18. Raluca-Maria Indre, Jelena Pesic, James Roberts: POPI: A Passive Optical Pod Interconnect for high performance data centers. Proc. ONDM 2014.

19. Y. Pointurier, B. Uscumlic, M.A. Mestre, P. Jennev, H. Mardoyan, A. Dupas, and S. Bigo: Green Optical Slot Switching Torus for Mega-Datacenters. ECOC 2015
20. N. Farrington, A. Forencich, P. Sun, S. Fainman, J. Ford, A. Vahdat, G. Porter, and G. C. Papen: A 10 us Hybrid Optical-Circuit/Electrical-Packet Network for Datacenters. Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2013, OSA Technical Digest (online) (Optical Society of America, 2013), paper OW3H.3.
21. Ken-ichi Kitayama et al.: Optical Packet and Path Switching Intra-Data Center Network: Enabling Technologies and Network Performance with Intelligent Flow Control. ECOC 2014, paper Tu.1.6.1
22. W. Miao et al.: SDN-Enabled OPS with QoS Guarantee for Reconfigurable Virtual Data Center Networks. Journal of Optical Communications and Networking, Vol. 7, No. 7, July 2015.
23. B. Mukherjee: Optical WDM Networks. Springer, 2006.
24. Fujitsu White Paper: The Key Benefits of OTN Networks. Fujitsu Network Communications Inc., 2010.
25. 802.17 Resilient packet ring (RPR). IEEE Computer Society, 2004.
26. MPLS Transport Profile (MPLS-TP), A Set of Enhancements to the Rich MPLS Toolkit. White paper, Juniper Networks, 2011.
27. ADDING SCALE, QoS AND OPERATIONAL SIMPLICITY TO ETHERNET: White paper, Provider Backbone Transport. Nortel Networks, 2007.
28. N. Bouabdallah, G. Pujolle, and H. Perros: Cost-effective single-hub wdm ring networks. IEEE International Conference on Communications, ICC '06, Vol. 5, pp. 24212426, June 2006.
29. L. Dittmann et al.: The European IST Project DAVID: A Viable Approach Toward Optical Packet Switching. IEEE Journal On Selected Areas In Communications, Vol. 21, No. 7, Sept. 2003.
30. A. Carena et al.: RingO: An Experimental Packet Network for Metro Applications. IEEE Journal On Selected Areas In Communications, VOL. 22, No. 8, OCTOBER 2004.
31. I.M. White, M.S. Rogge, K. Shrikhande, and L.G. Kazovsky: A summary of the HORNET Project: A Next-Generation Metropolitan Area Network. IEEE Journal On Selected Areas In Communications, VOL. 21, pp. 1478-1494 (2003)
32. J. Dunne, T. Farrell, and J. Shields: Optical packet switch and transport: A new metro platform to reduce costs and power by 50% to 75% while simultaneously increasing deterministic performance levels. Sixth International Conference on Broadband Communications, Networks, and Systems, BROADNETS 2009, pp. 15, September 2009.
33. I. Widjaja, I. Saniee, R. Giles, and D. Mitra: Light core and intelligent edge for a flexible, thin-layered, and cost-effective optical transport network. Communications Magazine, IEEE, vol. 41, no. 5, pp. S30S36, May 2003.
34. S. Cao, N. Deng, T. Ma, J. Qi, X. Shi, J. He, and J. Zhou: An optical burst ring network featuring sub-wavelength- and wavelength-granularity grooming. Photonics Global Conference (PGC) 2010, pp. 13, Dec 2010.
35. I. Popescu, B. Uscumlic, Y. Pointurier, P. Gravey, M. Morvan, and A. Gravey: A cost comparison of survivable subwavelength switching optical metro networks. Proceedings of the 26th International Teletraffic Congress, ITC 26, 2014.