Analysis of Performance Degradation in Sleep-Mode Enabled Core Optical Networks [Invited]

Albert Pagès, Massimo Tornatore, Jordi Perelló, Salvatore Spadaro, and Annalisa Morea

I. Introduction

Today's access networks represent around 70% of the overall power consumption of telecommunications networks, while core networks represent about 15% of it. However, as the traffic on the Internet increases, it is forecast that the power consumption in telecommunications networks will almost double by 2017, with access and core networks accounting for around 38% and 42% of such power consumption, respectively [1]. For these reasons, it is of capital importance to reduce the power consumption in core networks. Wavelength-division-multiplexing (WDM)-based optical networks have helped in reducing the power consumption of core networks by allowing large amounts of traffic to bypass routers through optical switching [2,3]. However, additional energy savings in core optical networks can be achieved by using more energy-efficient opto-electronic (OE) devices. In [4], the authors presented a novel architecture for both transponders (TSPs) and regenerators (REGs) capable of either putting some of their modules in a low-power consumption mode (*idle* mode) or turning them off (*off* mode) when they do not support traffic. OE device modules can then be turned back in a fully operative (*on*) mode whenever new connections are established. The dynamic management of the OE devices' power states is carried on according to the current traffic load and to the requirements in terms of connection setup time. Figure 1 depicts the inter-state transitions for such power adaptive devices.

With such power management capabilities, the authors of [4] demonstrated that up to 56% power savings can be attained with respect to traditional WDM networks, where devices are always powered on regardless of whether they are transmitting data or not. Other studies (e.g., [5,6]) confirm the benefits of putting some elements in core optical networks into a low-power consumption mode. In [7], different classes of traffic are considered, and, as a consequence, the setup time of the connections becomes a critical parameter, since typically high-priority traffic (e.g., real-time traffic) requires short setup times, while low-priority traffic does not. Considering that the boot-up time (i.e., the time required for the transition of a TSP or a REG from the *off* to *on* mode) can be quite long (tens of seconds [4]), available TSPs and REGs in the *off* mode cannot be allocated to support high-priority traffic.

Therefore, TSPs and REGs in the *off* mode should only be allocated to low-priority traffic, while TSPs and REGs in *idle* mode should be used to allocate high-priority traffic, since *idle* devices have a much lower activation time (time required for the transition from the *idle* to *on* mode) that can match the more stringent setup time requirements. Note that, by reserving some of the available devices per node for high-priority traffic and putting them into *idle* state, it is possible to improve the overall reactivity of the network, satisfying both high- and low-priority connection requirements, and, at the same time, reduce the power consumption of the network.

Manuscript received July 1, 2014; revised December 19, 2014; accepted January 15, 2015; published February 17, 2015 (Doc. ID 214854).

A. Pagès (e-mail: albertpages@tsc.upc.edu), J. Perelló, and S. Spadaro are with the Advanced Broadband Communications Center (CCABA), Universitat Politècnica de Catalunya (UPC), Barcelona, Spain.

M. Tornatore is with the Department of Electronics and Information (DEIB), Politecnico di Milano, Milan, Italy.

A. Morea is with Alcatel Lucent, Bell Labs, Nozay, France.

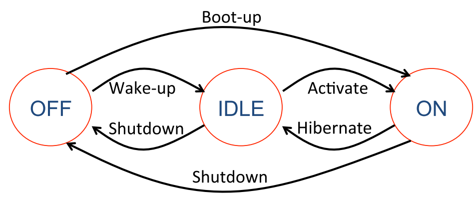

Fig. 1. Diagram of the transitions between the three power states of a power adaptive OE device.

Although there are works that analyzed the blocking probability in sleep-mode enabled optical networks with low- and high-priority traffic, such as in [8], there are no works, to the best of our knowledge, that evaluate the impact of the OE devices' wake-up time on the overall network blocking probability (*especially for high-priority traffic*) in core optical networks. Essentially, depending on the duration of the TSP/REG wake-up process, and the arrival rate of the connections, there may occur some situation in which devices in the *idle* state may not be available to be allocated to incoming high-priority connections.

Therefore, a proper dimensioning of the number of devices in *idle* mode per node to support high-priority traffic must be done. In [9], we evaluated the impact of TSPs' and REGs' wake-up times in a translucent optical network in terms of blocking probability of the connections as a function of the traffic dynamicity, i.e., the arrival rate and the duration of the connections. Under the same load conditions, more dynamic traffic implies that the arrival/departure rate of the connections is faster than in less dynamic traffic, with connections lasting less and arriving at a higher rate. Under such conditions, higher traffic dynamicities may suffer from higher blocking due to nonnegligible wake-up times, since more connections may arrive at a network node during a wake-up operation. Thus, we found that, although short wake-up times do not entail significant performance degradations even for highly dynamic traffic, as the duration of the wake-up time grows, less and less dynamic traffic can be supported without incurring severe performance degradations. For instance, for a wake-up time of 1 min, the blocking probability increased by a factor of 14 when compared against an ideal case (wake-up time = 0) for low traffic dynamicities, while it increased by a factor of almost 140 for higher traffic dynamicities.

In light of these results, it becomes clear that the proper dimensioning of the *idle* resources reserved for high priority is tightly related to the traffic dynamicity and the duration of the wake-up operation of the devices. To this end, in [9] we also analyzed the necessary extra *idle* resources per node needed to achieve sustainable blocking figures. We found that it is possible to compensate the effects of nonnegligible wake-up times with a small increase (3%–6%) in the number of reserved resources per node for high-priority traffic, even for high traffic dynamicities. The required increase in the number of *idle* resources translates to a marginal increase of the average daily power consumption (around 1.3%), since the additional devices enjoy the benefit of the low power consumption of the *idle* mode when not in use. Moreover, despite the slight increase in the energy consumption, average power savings around 65% were found compared to a traditional WDM optical network, where devices always stay powered on regardless of whether they are sending/receiving traffic.

However, an open issue is how this wake-up operation affects different network architectures of core optical networks, namely 1) opaque optical networks, where connections undergo an optical–electrical–optical (O-E-O) conversion in every intermediate node along the end-to-end path; 2) translucent optical networks, where connections may undergo a regeneration process on some of the nodes along the end-to-end path; and 3) transparent optical networks, where connections always remain in the optical domain along the entire end-to-end path. Due to the different numbers of devices that may undergo a wake-up operation in these three types of optical networks, significant differences in terms of blocking associated with the wake-up time can arise.

To this end, in this work we evaluate the degradation of the overall network performance in terms of blocking probability due to the wake-up time of the devices in a core optical network in all three architectures. Moreover, evaluation of the power savings with respect to traditional WDM scenarios is also performed. Additionally, we introduce a novel routing algorithm called wake-up time aware routing (WTAR) that aims to compensate the blocking due to the presence of nonnegligible wake-up times.

The rest of the paper is structured as follows: Section II describes the employed resource reservation strategy and the proposed WTAR algorithm. Section III depicts the evaluation scenario and the results obtained during the performance evaluation. Finally, Section IV draws the main conclusions.

II. TSPS AND REGS ALLOCATION STRATEGY

In this paper, we consider the dynamic management of a core network, where connections arrive and disconnect at random to/from the network. Connections in such a situation are understood as the aggregation at the source node of the core network of multiple client connections coming from the same access/metro network that are going to the same destination access/metro network. Future core optical networks are expected to be more dynamically operated than nowadays [10], mainly due to the continuous traffic growth of the Internet and new emerging applications, such as video on demand and multimedia social networks, which require a more flexible resource allocation in order to satisfy their needs. In such a scenario, it becomes highly important to keep the associated operational expenditures (OPEX) of the network at low levels, so as to cope with these new requirements in an efficient way. A very important parameter in this regard is the power consumption of the network. Thus, it is possible to save costs by adjusting the number of active devices according to the current traffic.

Additionally, future networks are predicted to be more service aware in order to cope with the heterogeneity of the applications that will run on top of them. Technologies such as deep packet inspection [11] will allow network operators to classify traffic in multiple categories, for instance, high-priority traffic coming from financial applications or telemedicine, which require a high quality of service (QoS), and low-priority traffic with less stringent QoS. By treating the traffic according to their priority, it is predicted that network operators will see an increase in their revenues

For these reasons, in this paper we are targeting the dynamic management of the power states of OE devices in a core optical network, considering that multiple classes of traffic coexist. To this end, we consider two classes of connections, namely, high-priority and low-priority connections, characterized by two different requirements in terms of setup time. Specifically, we assume that the high-priority connections have more stringent setup time requirements than low-priority connections. To manage the establishment of the high-priority connections according to their setup time target, a possible energy-efficient strategy consists of reserving part of the node OE resources (TSPs and REGs) for high-priority traffic and setting them in *idle* mode for a prompt establishment of such connections. Assuming that every node in the network is equipped with exactly N TSPs and N' REGs, a fraction of such resources (m and m', respectively) are set in the *idle* state and will be exploited by high-priority traffic. Note that both values N and N' as well as m and m' are the same for all nodes and fixed *a priori*. Then, the rest of the unused resources (those that are neither transmitting/receiving data nor reserved for high-priority traffic) are left in *off* mode and used for future low-priority incoming connections or to be switched to *idle* when m or m' decreases.

In order to maintain the pool of resources reserved for high-priority traffic, every time an *idle* device (either a TSP or REG) is allocated for a new high-priority connection (and thus set to *on* mode), if and only if spare *off* devices exist, a wake-up operation will be triggered and an *off* device will start changing to *idle* mode. Hence, high-priority connections are blocked if no *idle* devices are available. As for low-priority connections, if no *off* devices are available when they arrive, they are also blocked, since the devices in the *idle* state are reserved only for high-priority connections. Additionally, when a high- or low-priority connection is torn down, if the number of *idle* devices is lower than m (TSPs) or m' (REGs), the released devices are hibernated to *idle* mode. Otherwise, the released devices are completely shut down (*off* mode).

Regarding the wake-up operation (transition from *off* to *idle*), its duration can dramatically affect the blocking levels on the network, specifically for high-priority traffic. In this regard, high-priority connections can be blocked mainly for two reasons: 1) lack of free resources (TSPs, REGs, or wavelengths) due to connections competing for them, or 2) lack of *idle* devices (TSPs or REGs) due to wake-up time delays.

As explained above, after an *idle* device is selected to support an incoming high-priority connection, a wake-up operation is also triggered to maintain the pool of reserved

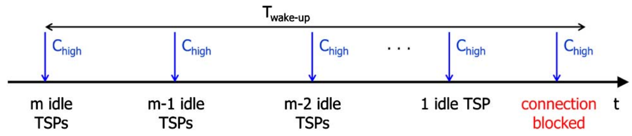

Fig. 2. Example of blocking of high-priority connections due to the influence of the wake-up time.

resources for high-priority traffic (a spare *off* device is set to *idle*). But, in the case that additional $m-1$ ($m'-1$) high-priority connections have to be established at that particular node during the wake-up operation, all *idle* devices will become exhausted. So, any further high-priority connection arriving before the ending of the wake-up operation will be blocked due to the lack of *idle* devices. Figure 2 depicts an example of this event.

Some studies suggest that, with today's technology, the duration of the wake-up time is on the order of tens of seconds (e.g., [4]). Depending on the inter-arrival rate of the connections, it may happen that the duration of the wake-up time is excessive, leading to unsustainable levels of blocking for network operators. This is especially critical in opaque optical networks, as they require the utilization of a TSP at every node on the end-to-end path. In this situation, a significant number of nodes may simultaneously exhaust their pool of reserved resources for high-priority connections. For this reason, this kind of opaque network is more sensitive to performance degradations due to nonnegligible wake-up times. On the other hand, this situation is less severe in transparent networks, since the source and the destination nodes are the only ones affected by the lack of *idle* resources; hence fewer nodes may simultaneously run out of *idle* devices. Translucent optical networks represent an intermediate case due to the use of some regenerators (that are also affected by the wake-up time) along the end-to-end path.

To reduce such blocking situations, we propose the novel WTAR algorithm, which considers as an input the available *idle* resources in the end-to-end paths in the network and tries to balance the connections to avoid the exhaustion of *idle* devices in the nodes. In the following, we provide details of the two versions of the algorithm, one for the opaque case and another for the translucent case. We did not consider the transparent case since in this case the wake-up time only affects the source and the destination of the connections; hence there is no chance to avoid the exhaustion of *idle* devices in the nodes by means of balancing the connections.

A. Opaque WTAR

As we mentioned before, opaque networks are more sensitive to nonnegligible wake-up times due to the fact that connections established in the network employ $2 \times n_{\text{hops}}$ TSPs, n_{hops} being the number of links belonging to the end-to-end path. As the duration of the wake-up operation increases, the chances of running out of *idle* TSPs in a node increases, leading to blocking situations. At this point it is worth mentioning that we are considering an opaque network in which nodes are equipped with W TSPs per incoming/outgoing link (TSPs are considered bidirectional).

A possible strategy to avoid this problem is to encourage the load balancing of the *idle* devices on the network, reducing the congestion of the nodes, hence avoiding the blocking of a high-priority connection due to the lack of free resources. Moreover, it is also interesting to balance the load of *off* devices since they will be employed to perform wake-up operations when needed. To this end, we can use the following path metrics:

$$\alpha \frac{l_p}{D_r} + (1-\alpha)\frac{m}{s_{\text{idle},p}}, \tag{1a}$$

$$\alpha \frac{l_p}{D_r} + (1-\alpha)\frac{N-m}{s_{\text{off},p}}, \tag{1b}$$

where l_p represents the length of the path and D_r the transmission reach of the signals (both in kilometers), N and m represent the total number of TSPs and the number of *idle* TSPs per node and bidirectional link (we remind the reader that the assumed node architecture has a dedicated set of TSPs for every incoming/outgoing link) reserved for high-priority connections, respectively, and $s_{\text{idle},p}$ and $s_{\text{off},p}$ represent the number of *idle* and *off* TSPs in the most congested node along the path, respectively. Note that the metrics are for the whole path, discriminating the candidate paths in terms of total length and most congested resources. Parameter α is a real value in [0, 1] and is used to set the emphasis of the routing policy. For instance, in the case of $\alpha = 0$, the routing algorithm chooses as candidate paths the ones having larger values of $s_{\text{idle},p}$ or $s_{\text{off},p}$, depending on the priority. That is, it chooses the paths that have the least congested nodes, avoiding blocking situations due to the lack of resources in the nodes. However, since the connections are balanced in the network, longer paths might be chosen, thus increasing the number of TSPs that will be powered on and, hence, the overall power consumption of the network.

On the other hand, if $\alpha = 1$, shorter paths in terms of physical distance will be preferred as candidate paths. In this situation, the energy consumption is kept low, since less resources are powered on. However, the availability of both *idle* and *off* TSPs will be not balanced, and the blocking associated with the wake-up time will be higher. Scenarios with α between 0 and 1 represent a routing policy offering a trade-off between power consumption and blocking probability.

Thus, the proposed WTAR algorithm can be described as follows:

1) When a connection arrives, calculate the first K shortest paths between source and destination (where K is an input parameter) utilizing the metric (1a) for a high-priority connection or (1b) for a low-priority connection. In both cases, if two or more of the calculated paths have the same weight, the path with the lowest number of hops is prioritized.
2) Select the first path as the candidate path and try to establish the connection employing a first-fit (FF) criteria for the wavelength selection, considering that, thanks to the O-E-O capability of the nodes, the wavelength continuity constraint may be relaxed and different wavelengths along the path may be used. Remember that high-priority connections can only be served employing *idle* TSPs while low-priority connections can only be served employing *off* TSPs.
3) If the connection can be set up along the selected path, mark the connection as successful. Otherwise, try the following path in the set.
4) If all the paths in the set have been explored and none has free resources, the connection is blocked.

B. Translucent WTAR

The WTAR algorithm for the translucent case presents slight differences from the one presented for the opaque case. In a translucent optical network, the connection always stays in the optical domain until regeneration is needed, which is performed by a REG at one of the regeneration pools of the network. In this case the REGs become the most critical resources, as they may be utilized in multiple points along the end-to-end path. As for add/drop TSPs, since they are only employed in the source and the destination nodes of the connection, there is no chance to balance their use by means of an appropriate routing algorithm. For these reasons, we develop a metric to help balancing the use of the REGs in the network similarly to how TSPs were balanced in the opaque WTAR case. Specifically,

$$\alpha \frac{l_p}{D_r} + (1-\alpha)\frac{m'}{r_{\text{idle},p}}, \tag{2a}$$

$$\alpha \frac{l_p}{D_r} + (1-\alpha)\frac{N'-m'}{r_{\text{off},p}}, \tag{2b}$$

where N' and m' represent the total number of REGs and the number of *idle* REGs reserved per node for high-priority connections, respectively, and $r_{\text{idle},p}$ and $r_{\text{off},p}$ represent the number of *idle* and *off* REGs in the most congested node along the path, respectively. The rest of the variables are as in the opaque case. With such definitions, the actual WTAR algorithm proposed for translucent networks is as follows:

1) Same as step (1) in the opaque scenario, utilizing Eqs. (2a) and (2b) as the metrics.
2) Select the first path as the candidate path and try to establish the connection employing a FF criteria for the wavelength selection, considering that the wavelength continuity constraint has to be guaranteed between regeneration spans. Please consider that wavelength continuity also has to be guaranteed between an add/drop TSP and a REG (source and destination). When regeneration is needed (the length of the path is longer than D_r), it will be performed in the first

possible node along the end-to-end path, since we assume that every node in the network is equipped with a pool of regenerators. Depending on the length of the path, regeneration may be needed multiple times. In such cases, it will be performed in the first possible node along the accumulated path (before the accumulated path is longer than D_r). Additionally, we consider that any REG is capable to perform wavelength conversion.

3) Same as (3) in the opaque case.
4) Same as (4) in the opaque case.

The behavior of the translucent WTAR algorithm is the same as in the opaque case: with $\alpha = 0$, the use of REGs in the network will be balanced, reducing the blocking associated with the effects of the wake-up time; with $\alpha = 1$, the shortest physical routes will be selected, keeping the energy consumption low; any case in between represents a compromise solution between power savings and blocking probability.

In the following section, we evaluate the proposed algorithms through extensive simulations.

III. Test Scenario and Results

A. Scenario Description

To analyze the impact of the wake-up time (T_w), we have executed various simulations employing two network topologies, namely, the 12-node Deutsche Telekom (DT) network [12] and the 30-node CORONET network [13]. Table I shows the main network statistics considering the shortest paths among all node pairs in the two networks to compute average values. For all the tests, we offered $4 \cdot 10^5$ connections to the network, which are assumed to be bidirectional and requesting one wavelength each. Connections arrive at the network following a Poisson distribution, with exponentially distributed inter-arrival times and holding times with average value IAT and HT, respectively. Thus, the network load can be defined as HT/IAT. Offered traffic is uniformly distributed among all source/destination pairs. The bit rate of the connections is 100 Gb/s, employing a 28 Gbaud PDM-QPSK modulation, for which the transparent reach without regeneration is 1200 km [14].

For the opaque and translucent scenario, the connections are routed employing the WTAR algorithm (the specific value of the α parameter will be annotated for each experiment). For the transparent scenario, connections are routed employing a K-shortest path algorithm with physical distance as the path metric. Only routes that are shorter than the transparent reach are considered, while the others are blocked. As for the wavelength assignment, a FF criteria is utilized, considering that the same wavelength must be provisioned in all links in the end-to-end path. For all three scenarios, the value of K (the number of candidate paths) is equal to 6.

TABLE I
Network Statistics

	DT	CORONET
Shortest path length (km)	48 (1 hop)	69 (1 hop)
Longest path length (km)	1095 (4 hops)	7078 (9 hops)
Average path length (km)	530.7424	3064.7195
Average number of hops	2.1363	4.8253

As a baseline scenario, we considered the DT topology network considering an opaque architecture with an ideal TSP T_w; that is, $T_w = 0$ s. With this scenario in mind, we dimensioned the network so as to have an overall blocking probability (considering high- and low-priority traffic classes) not higher than 1%, assuming an offered network load of 140 and mean HT = 1 h, with a share of high and low-priority traffic of 30% and 70%, respectively. As a result, we have found that nodes have to be equipped with 30 TSPs per incoming/outgoing link and the same value of wavelengths per link. The number of TSPs per node and incoming/outgoing link reserved for high-priority traffic in this basic case is $m = 2$. For the sake of simplicity, we will also employ the same number of resources in the CORONET network topology. In such a case, the overall blocking probability is no larger than 2%. To perform a fair comparison, although in a translucent and transparent optical network scenario TSPs are only employed in the source and destination nodes, we adopt the same number of network resources in terms of both TSPs and wavelengths as in the opaque case. This is done to avoid having the opaque scenario with much more node equipment (i.e., TSPs) than in the transparent or translucent case. Moreover, for the translucent scenario we consider that each node is equipped with $N' = 15$ REGs with $m' = 1$ reserved for high-priority traffic. As a result, for this base scenario, the overall blocking probability is not higher than 1%–5% of the total offered connections for both transparent and translucent cases.

On this base scenario, we analyzed the impact of the value of T_w on the network blocking probability (BP). Since the impact of T_w is tightly related to the traffic dynamicity (how many high-priority connections may arrive during a wake-up operation), we have investigated, for various degrees of dynamicity and values of T_w, how BP values of high-priority traffic evolve. To this end, we define $r = \mathrm{HT}_{\mathrm{base}}/\mathrm{HT}$ as the dynamicity of the traffic, denoting as $\mathrm{HT}_{\mathrm{base}}$ the mean HT in the base scenario (i.e., 1 h), modifying the mean IAT of the connections in all cases so as to maintain the same network load of 140. Hence, under the same offered load, more dynamic connections are being established.

B. DT Network Results

We start by evaluating the impact of T_w for all three cases (opaque, translucent, and transparent) considering that the routing algorithm is completely blind to T_w; that is, no WTAR is being considered ($\alpha = 1$ in the opaque and translucent case). Figure 3 shows the BP as a function of r for high-priority traffic, considering multiple values of T_w, for the DT network topology. The base scenario ($T_w = 0$ s) is also depicted as a reference. It can be appreciated that the opaque scenario is the most sensitive to the value of T_w,

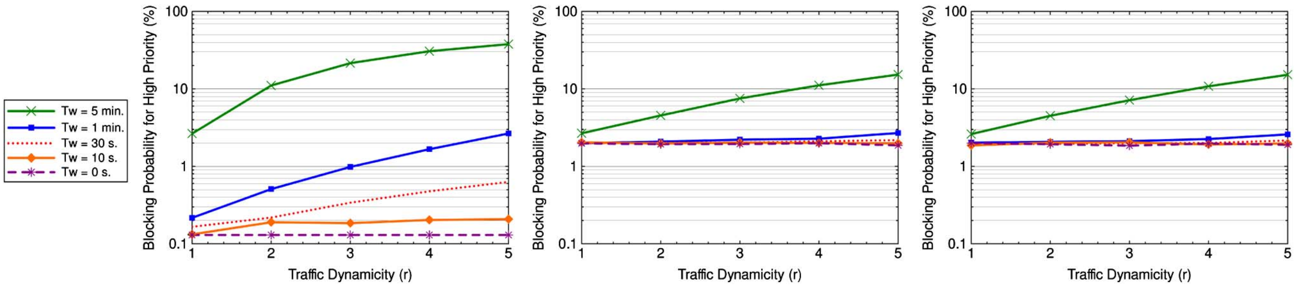

Fig. 3. Blocking as a function of r for the opaque (left), translucent (middle), and transparent (right) cases for the DT network scenario without WTAR.

since, as explained before, all the end-to-end routes of the connections entail the use of devices affected by T_w. More specifically, we can see that as T_w increases, less traffic dynamicity can be supported without incurring severe performance degradations. For instance, for $T_w = 1$ min and $r \geq 3$, the BP becomes higher than 1%, while for $T_w = 5$ min, the BP becomes higher than 1% even for the base dynamicity ($r = 1$). The translucent and transparent cases behave very similarly because, in the DT network, regeneration occurs very sparsely. Additionally, we can see that up to $T_w = 1$ min all the curves feature almost flat behavior. Although not shown for the sake of space, the BP of low-priority traffic is around 1%–2% in the opaque case for all traffic dynamicities and $T_w \leq 1$ min, while it experiences a slight reduction in the case of $T_w = 5$ min. This is due to the fact that higher values of T_w result in more high-priority connections being blocked; hence, more resources can be exploited by low-priority traffic.

Now let us investigate the main causes of blocking of high-priority connections. In the opaque scenario, the sole cause of blocking is the lack of *idle* TSPs, while in the translucent and opaque scenarios the vast majority of blocking (around 99%) is due to the lack of free contiguous wavelengths. For this reason, the BP in the translucent and transparent case for $T_w \leq 1$ min is higher than in the opaque case, since it is necessary to provision the same wavelength in all links along the path. Nevertheless, for $T_w = 5$ min, the lack of *idle* resources starts to be noticeable. Note, however, that the experimented degradations are still much less pronounced than in the opaque case, due to the fact that less resources are affected by T_w (only the source and destination nodes are affected).

In light of this, we have investigated for various values of r how many TSPs per node should be set to *idle* in order to reduce the BP of high-priority traffic below 1% considering $T_w = 1$ min. Figure 4 depicts the obtained results (the black thick solid line marks the desirable 1% BP value) for the opaque, translucent, and transparent scenarios.

At least one additional *idle* TSP per node should be reserved to meet the desired BP for $r = 3$, 4, and 5 in the opaque scenario. Note, however, that when more TSPs are set to *idle*, the BP of low-priority connections starts to increase, since fewer resources can be exploited by them. Hence, to maintain a low-priority BP constant, a number of *off* TSPs equal to the additional TSPs set to *idle* should be equipped in the nodes, thus increasing the overall network cost. In contrast, we can see that even increasing the number of *idle* TSPs does not decrease the BP in the translucent and transparent cases, since, as commented before, in such scenarios, the main cause of blocking is the lack of free wavelengths.

Now, we will evaluate how the blocking levels change when the routing algorithm takes into consideration the presence of T_w. Since in the DT network the opaque scenario is the only one that is highly affected by T_w, we will focus on the opaque case considering $\alpha = 0$. To this end, Fig. 5 depicts the improvement ratio with respect to the case in which no WTAR is performed ($\alpha = 1$), that is, the ratio between the BPs in the case without WTAR and the case with WTAR. We can see that substantial

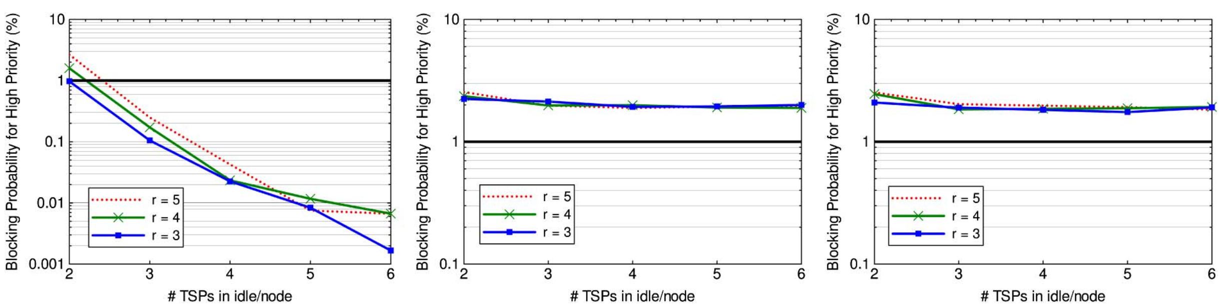

Fig. 4. Blocking as a function of the number of idle TSPs for the opaque (left), translucent (middle), and transparent (right) cases for the DT network scenario without WTAR.

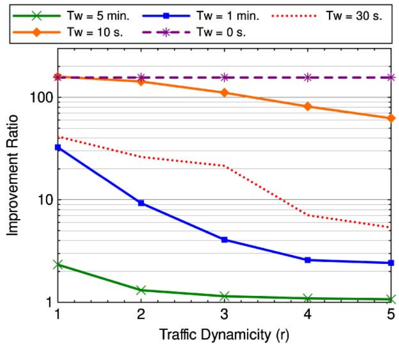

Fig. 5. Improvement ratio as a function of r for high-priority traffic for the DT network scenario with WTAR ($\alpha = 0$).

improvements in the BP can be achieved with the proposed WTAR algorithm, with improvement ratios up to 100 for low values of T_w, which translates into reductions of up to two orders of magnitude in the BP of high-priority connections. More modest improvements (around one order of magnitude) are achieved for higher values of T_w. Additionally, the blocking levels remain lower than 1% for all traffic dynamicities for $T_w \leq 1$ min by using WTAR. This means that we can reduce the overall blocking levels without provisioning extra *idle* TSPs or *off* TSPs. Although not shown, the blocking levels of low-priority connections are also reduced (up to one order of magnitude) for all traffic dynamicities and values of T_w. This happens because, as explained before, the WTAR algorithm also balances the use of *off* devices in the network ensuring that there will be enough of them to perform wake-up operations.

To complete the studies in the DT network, let us now evaluate the blocking levels of high-priority traffic as a function of α. To this end, we focus on the opaque case, considering $r = 4$. Figure 6 depicts the obtained results. We can see that the choice of α has a significant impact on the blocking probability. Generally, lower values of α result in lower blocking figures, since the load balancing term of the WTAR algorithm has more weight. This is especially evident for low values of T_w, with variations of some orders of magnitude in the blocking. As for higher values of T_w, the impact of α is lower due to the fact that nodes exhaust more rapidly their pool of *idle* resources; hence

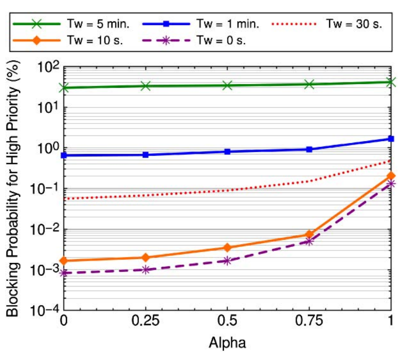

Fig. 6. Blocking as a function of α for the opaque case for the DT network scenario.

the routing algorithm has fewer chances to balance their utilization.

C. CORONET Network Results

At this point, let us evaluate the BP of the WTAR algorithm in a larger network topology, namely, the 30-node CORONET topology. We will focus on the opaque and the translucent cases, since the longer physical diameter of this network does not allow the establishment of transparent connections for all node pairs due to transmission reach reasons. As in the previous subsection, we start by evaluating in Fig. 7 the blocking of high-priority traffic as a function of r for various values of T_w considering that no WTAR is employed ($\alpha = 1$).

Results confirm that the opaque case is very sensitive to the value of T_w, allowing for less dynamic traffic as T_w increases. As for the translucent case, we can see that in such a larger network, since more regeneration is needed, the effects of T_w start to be noticeable already for $T_w \geq 10$ s. In this scenario, we observed that the main cause of blocking is the lack of *idle* REGs (around 60%), whereas the second cause is the lack of free wavelengths for low values of T_w or the lack of *idle* TSPs for high values of T_w. Moreover, as in the DT scenarios, as the duration of T_w increases low-priority traffic experiences a slight improvement in its blocking figures, since more resources can be exploited by it.

We will now evaluate the amount of extra *idle* resources (TSPs for the opaque case and REGs for the translucent case) needed to achieve a BP under 1%. We consider

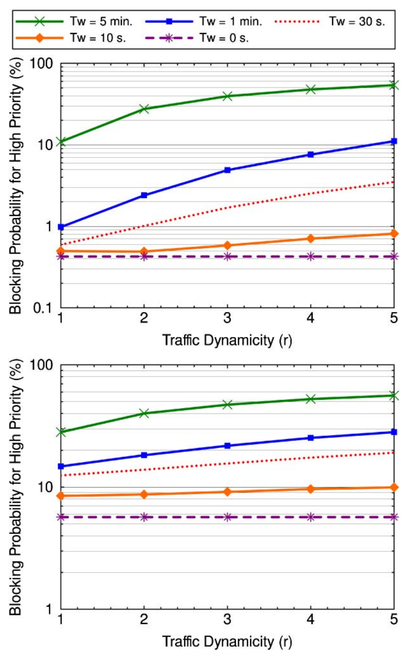

Fig. 7. Blocking as a function of r for the opaque (top) and translucent (bottom) cases for the CORONET network scenario without WTAR.

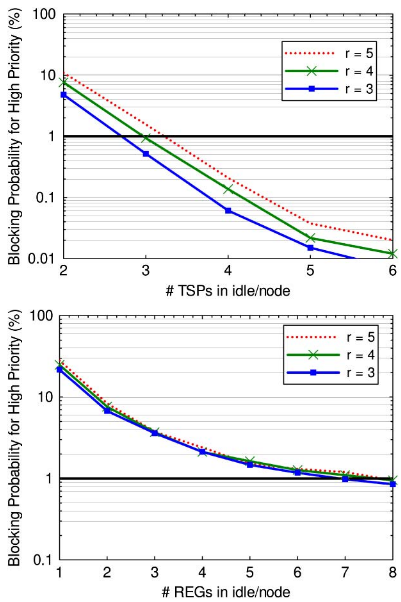

Fig. 8. Blocking as a function of the number of idle resources for the opaque (top) and translucent (bottom) cases for the CORONET network scenario without WTAR.

$T_w = 1$ min and three dynamicity values, namely 3, 4, and 5. Figure 8 depicts the obtained results. As in the DT network scenario, the increase of *idle* TSPs per node in the opaque case allows for reducing the blocking levels until the desired threshold is achieved, needing one extra *idle* TSP for $r = 3$ and 4, and two for $r = 5$. In the translucent case, even though increasing the number of *idle* REGs per node reduces the blocking, the reductions do not allow us to achieve the targeted performance without substantial increases in the reserved resources. For example, it would be necessary to provision at least six to seven extra *idle* REGs per node to achieve a BP below 1%.

Let us now evaluate the BP for the WTAR algorithm with $\alpha = 0$. As in the DT case, we will depict the improvement ratio with respect to the case without WTAR. Figure 9 depicts the obtained results. We can see that more moderate reductions are achieved when compared to the DT network. This is due to the fact that in the CORONET network, since connections go through a larger number of devices, the overall blocking figures are higher due to multiple nodes simultaneously running out of *idle* devices. In such a situation, the routing algorithm has fewer chances to influence the utilization of the OE devices by means of load balancing. Nevertheless, reductions in the blocking up to around 70% can be achieved in some cases. As in the DT scenario, low-priority connections also experience some reduction on their blocking figures, since the routing algorithm also balances the use of *off* devices. Next, we evaluated the evolution of the blocking probability as a function of α for both opaque and translucent network scenarios, fixing $r = 4$. Figure 10 depicts the obtained results. It can be seen that the impact of α in the CORONET network is less acute when compared to the DT network, due to the fact that connections are experiencing more blocking.

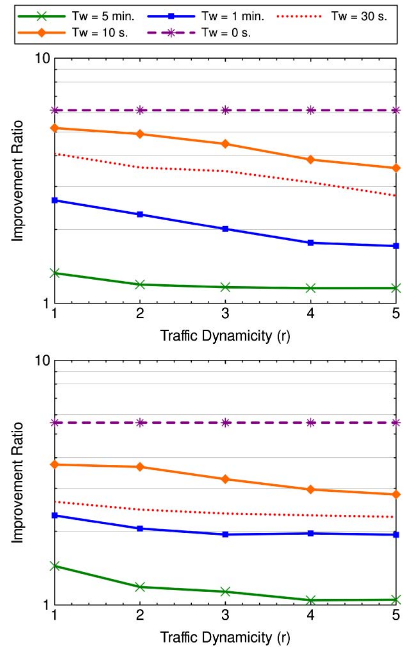

Fig. 9. Improvement ratio as a function of r for high-priority traffic for the opaque (top) and translucent (bottom) cases for the CORONET network scenario with WTAR ($\alpha = 0$).

Nevertheless, it can be seen that lower values of α generally entail lower blocking figures, even for high values of T_w.

To conclude our study, we have evaluated the average daily power consumption under the fluctuating daily traffic profile presented in [15] considering three cases: 1) all

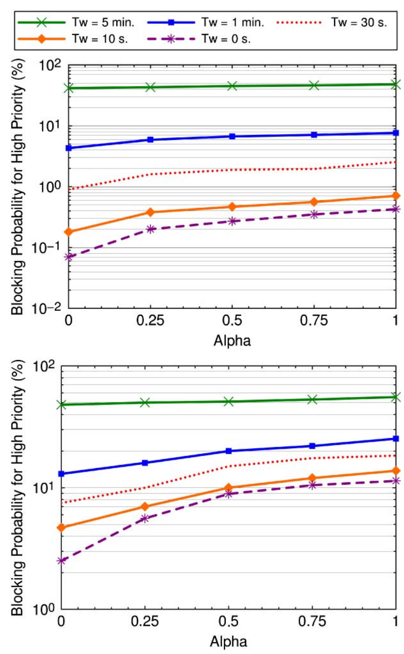

Fig. 10. Blocking as a function of α for the opaque (top) and translucent (bottom) cases for the CORONET network scenario.

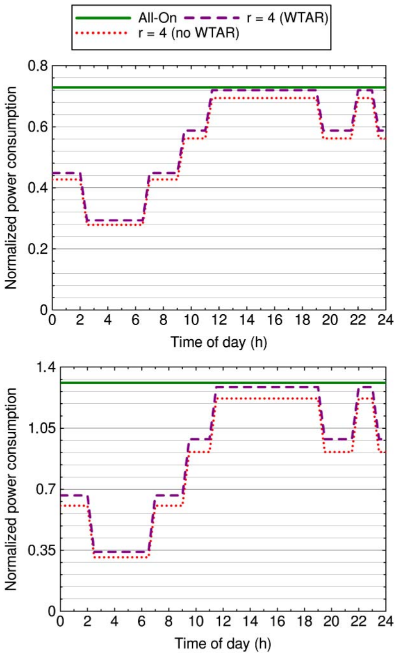

Fig. 11. Normalized daily power consumption for the opaque (top) and translucent (bottom) cases in the CORONET network.

devices stay always powered on, as in traditional WDM networks; 2) a sleep-mode strategy is adopted, but the routing algorithm is blind to the effects of T_w ($\alpha = 1$); and 3) a sleep-mode strategy is adopted and the presented WTAR algorithm is used ($\alpha = 0$). We use the CORONET network topology, considering both opaque and translucent architectures. Additionally, we set $T_w = 1$ min and $r = 4$. For the sake of fairness, in the all-on scenario, the same dimensioning used to get the other results has been assumed. Moreover, for the cases in which sleep-enabled devices are employed, nodes are equipped with extra *idle* and *off* devices per node whenever necessary according to the results presented previously. We assume the same power figures for the devices as in [4].

In order to present a fair comparison between the three cases and decouple the power consumption from the blocking figures, we define the normalized power consumption as the ratio of the average total power consumption in the network to the total number of accepted connections. Utilizing this metric, Fig. 11 shows the evolution of the power consumption along the day for the three cases. It can be appreciated that huge savings (around 70% during low traffic periods and around 30%–40% on average) are attained when the proposed power management strategy is employed. Focusing on the differences between employing or not the proposed WTAR algorithm, we can see that the WTAR algorithm results in higher average power figures. This is basically due to the fact that with the WTAR algorithm more connections are accepted, so more power is consumed. Additionally, as the WTAR algorithm balances the use of the network elements (TSPs and REGs), slightly longer paths are employed, increasing the number of devices that need to be powered on. Nevertheless, this increase in power when compared against the case without WTAR is very small and gets compensated with the substantial reductions in the BP that can be achieved (up to two orders of magnitude) and the fact that, under certain conditions, it is not required to equip network nodes with extra devices, allowing for a lower network CAPEX.

IV. Conclusion

In this work, we evaluated the impact of the wake-up times of transponders and regenerators in a multiservice core optical network employing sleep-mode enabled devices. We showed that various ranges of degradations can happen depending on the type of optical network (transparent, translucent, and opaque), the duration of the wake-up operation of the optical devices, the dynamicity of the traffic, and the geographical size of the network. Particularly, we showed that opaque networks are the most sensitive to degradations due to long wake-up time operations, due to the fact that the whole end-to-end path composing a connection is affected.

To compensate such effects, we proposed a novel routing algorithm called WTAR that aims to balance the use of *idle* resources on the network so as to reduce the blocking associated with the lack of them. We showed that significant reductions on blocking can be achieved with the proper WTAR algorithm (up to two orders of magnitude in the opaque case and up to 70% in the translucent case), with the added benefit that such reduction can be achieved with very small increments, or no increments at all, in the equipment amount per network node, maintaining the overall low network cost.

Additionally, we showed that substantial reductions can be achieved (up to 70%) in the average daily network power consumption when sleep-mode enabled transponders and regenerators are employed when compared to traditional WDM network scenarios. The additional power consumption that entails the use of the WTAR algorithm gets compensated by the significant reductions in connection blocking and network CAPEX when compared with a routing algorithm that is completely blind to the effects of the wake-up time.

Acknowledgments

The work in this paper has been funded by the Spanish Science Ministry through the project ELASTIC (TEC2011-27310) and the GreenTouch consortium.

References

[1] C. Lange, D. Kosiankowski, R. Weidmann, and A. Gladisch, “Energy consumption of telecommunication networks,” in *Proc. ECOC*, Sept. 2009.

[2] L. Wosinska, A. Jirattigalachote, P. Monti, A. Tzanakaki, and K. Katrinis, “Energy efficient approach for survivable WDM optical networks,” in *Proc. ICTON*, July 2010.

[3] Y. Zhang, P. Chowdhury, M. Tornatore, and B. Mukherjee, "Energy efficiency in telecom optical networks," *IEEE Commun. Surv. Tutorials*, vol. 12, no. 4, pp. 441–458, July 2010.

[4] A. Morea, S. Spadaro, O. Rival, J. Perelló, F. Agraz, and D. Verchere, "Power management of optoelectronic interfaces for dynamic optical networks," in *Proc. ECOC*, Sept. 2011, paper We.8.K.3.

[5] A. Muhammad, P. Monti, I. Cerutti, L. Wosinska, P. Castoldi, and A. Tzanakaki, "Energy-efficient WDM network planning with dedicated protection resources in sleep mode," in *Proc. GLOBECOM*, Dec. 2010.

[6] I. Cerutti, N. Sambo, and P. Castoldi, "Sleeping link selection for energy-efficient GMPLS networks," *J. Lightwave Technol.*, vol. 29, no. 15, pp. 2292–2298, June 2011.

[7] A. Morea, J. Perelló, F. Agraz, and S. Spadaro, "Demonstration of GMPLS-controlled device power management for next generation green optical networks," in *Proc. OFC/NFOEC*, Mar. 2012, paper OM3G.6.

[8] F. Musumeci, M. Tornatore, M. Riunno, and A. Pattavina, "Blocking analysis for green WDM networks with transponder power management," in *Proc. ICTON*, June 2013, paper We.A4.5.

[9] A. Pagès, M. Tornatore, J. Perelló, S. Spadaro, and A. Morea, "Impact of transponders and regenerators wake-up time on sleep-mode enabled translucent optical networks," in *Proc. OFC/NFOEC*, Mar. 2014, paper W4B.4.

[10] N. Sambo, P. Castoldi, F. Cugini, G. Bottari, and P. Iovanna, "Toward high-rate and flexible optical networks," *IEEE Commun. Mag.*, vol. 50, no. 5, pp. 66–72, May 2012.

[11] M.-Y. Liao, M.-Y. Luo, C.-S. Yang, C.-H. Chen, P.-C. Wu, and Y.-C. Chen, "Design and evaluation of deep packet inspection system: A case study," *IET Netw.*, vol. 1, no. 1, pp. 2–9, Mar. 2012.

[12] J. Perelló, A. Morea, S. Spadaro, A. Pagès, S. Ricciardi, M. Gunkel, and G. Junyent, "Power consumption reduction through elastic data rate adaptation in survivable multi-layer optical networks," *Photonic Netw. Commun.*, vol. 28, no. 3, pp. 276–286, Dec. 2014.

[13] Monarch Network Architects, http://monarchna.com/topology .html.

[14] A. Klekamp, R. Dischler, and F. Buchali, "Transmission reach of optical-OFDM superchannels with 10–600 Gb/s for transparent bit-rate adaptive networks," in *Proc. ECOC*, Sept. 2011, paper Tu.3.K.2.

[15] A. Morea, O. Rival, N. Brochier, and E. Le Rouzic, "Data rate adaptation for night-time energy savings in core networks," *J. Lightwave Technol.*, vol. 31, no. 5, pp. 779–785, Mar. 2013.