OBGP+: An improved path-vector protocol for multi-domain optical networks

M. Yannuzzi *, X. Masip-Bruin, S. Sánchez-López, E. Marín-Tordera

Department of Computer Architecture, Technical University of Catalonia (UPC), 08800 Vilanova i la Geltrú, Barcelona, Spain

ARTICLE INFO

ABSTRACT

One of the essential components for the dynamic provisioning of lightpaths across multiple domains is the *Routing and Wavelength Assignment* (RWA) strategy adopted. The consolidation that path-vector protocols have had in practice, has motivated the optical extension to BGP (OBGP). We claim, however, that a routing model mostly centered on the exchange of reachability information–like the one we have today with BGP or the one offered by OBGP–will not be sufficient for multi-domain optical networks. Routing domains must be able to exchange both reachability as well as aggregated *Path-State Information* (PSI). Understanding that this is a missing piece in the routing models provided by BGP and OBGP is easy nowadays, but contributing with solutions capable of highly improving the performance of a path-vector without impacting on key aspects of the protocol–fundamentally, its scalability, its convergence properties, and the number of routing messages exchanged between domains–is a challenging task.

In this paper we propose *OBGP+*, which is a very simple extension of a path-vector protocol supporting the computation and advertisement of PSI between optical domains. The PSI that we propose to use is highly condensed in the form of a single integer value. In order to avoid the typical increase in the number of routing messages associated with the update of PSI, we propose to piggy-back the updates in non-dummy Keepalive messages exchanged between OBGP+ neighbors. Extensive simulations reveal that, despite its simplicity: (i) OBGP+ is able to drastically reduce the blocking experienced with a path-vector protocol like OBGP; (ii) OBGP+ needs much less number of routing messages than OBGP to achieve such performance; and (iii) the convergence and restoration features of OBGP+ are also better than those of OBGP, which is particularly important for connections that lack a protection path.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Future optical networks will require new protocols in order to route and support the on-demand provisioning of lightpaths between different domains. In light of this, some researchers have started to analyze the possibility of adopting an *Optical Border Gateway Protocol* (OBGP) as the future inter-domain routing protocol for optical networks [1–4]. The aim of these proposals is to extend BGP so that it can convey and signal optical information between OBGP neighbors. The strength of this approach is that future optical networks will benefit from the well-known advantages of the BGP-based routing model, such as scalability, clear administrative limits of routing domains, fully-distributed network administration based on filtering and routing policies, etc. The weakness, on the other hand, is that the routing model of future optical networks will inherit the well-known issues in BGP [5]. Indeed, a multi-domain routing model mostly centered on the exchange of reachability information, like the one offered by OBGP, is not going to be sufficient. This is confirmed by a number of recent research initiatives,

* Corresponding author. Tel.: +34 938967293; fax: +34 938967700.
E-mail address: yannuzzi@ac.upc.edu (M. Yannuzzi).

like [6] in Europe and [7] in the US. Accordingly, in this paper we consider that neighboring domains are able to exchange both *Network Reachability Information* (NRI), and enriched routing information consisting of aggregated *Path-State Information* (PSI).

The subject of this study is to show that even a minor modification of a path-vector protocol like OBGP, is enough to drastically improve its performance.[1] More precisely, we show that by integrating only plain and highly aggregated PSI in OBGP, it is possible to outperform the latter, and this can be accomplished without increasing the number or the frequency of routing updates exchanged between domains. We also compare the convergence properties of OBGP with those of our extended path-vector protocol, which we call *OBGP+*. Our results show that, upon a link failure, OBGP+ is able to restore paths faster than OBGP, and also that OBGP+ exchanges fewer routing messages than OBGP during a convergence. These properties are particularly important for connections that lack of a protection path. We hope that the results presented here will encourage other researchers to devise novel ways for improving path-vector routing protocols for multi-domain optical networks.

The rest of the paper is organized as follows. Section 2 reviews related work. Section 3 presents the routing model, including the NRI and the PSI exchanged between optical domains. Then, in Section 4, the Routing and Wavelength Assignment (RWA) algorithm proposed for OBGP+ is detailed. The comparison between OBGP and OBGP+ is shown in Section 5. Finally, Section 6 summarizes the main conclusions and describes our future work.

2. Related work

Most of the research and standardization efforts carried out so far in the area of routing in optical networks have focused on intra-domain aspects. The discussions concerning multi-domain issues are still in a very early stage, so although some topics have started to be analyzed by the three relevant standardization bodies, namely, the International Telecommunications Union (ITU) [8], the Internet Engineering Task Force (IETF) [9], and the Optical Internetworking Forum (OIF) [10], the situation is that the majority of the multi-domain issues are largely open at present.

A few years ago, the OIF proposed the Domain-to-Domain Routing Protocol (DDRP) [11]. DDRP is basically a hierarchical extension of OSPF-TE, supported by a modified version of Dijkstra's algorithm. However, DDRP has two main drawbacks. First, it represents a major change in the routing system, since it proposes to move towards a fully hierarchical routing model. Second, the modified Dijkstra algorithm still offers limited functionality. For instance, it returns a single path at a time, so complementary algorithms need to be adopted for path protection and diverse routing purposes.

With a more conservative approach, the standardization efforts being carried out at the IETF mention the need

to work on new protocols, or extensions to the existing ones, in order to enable the advertisement of inter-domain Traffic Engineering (TE) information. For instance, [12] describes a framework for establishing and controlling Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) Traffic Engineered (TE) Label Switched Paths (LSPs) in multi-domain networks. Even though this work mentions the possibility of adding TE extensions to BGP, there is neither an indication of how this would be implemented, nor how frequently TE data should be updated between domains.

Some authors went one step further, proposing the optical extension of BGP [1–4]. However, there is a wide consensus that neither DDRP nor OBGP will be able to provide the routing functionality and expected performance for multi-domain optical networks. Given that inter-domain routing in optical networks is becoming an active research area, a sound approach is to start analyzing the issue of how routing protocols can be endowed with the ability to compute and efficiently convey aggregated PSI between domains. Some recent proposals in this direction can be found in [13,14].

In this work, we propose a very pragmatic approach. The proven scalability of path-vector protocols in the current Internet suggests that, if OBGP were deployed in the future, it would probably operate and scale similarly as BGP does today. This is remarkably important, since the operational aspects of new protocols, such as DDRP, can always be questioned in terms of scalability. From a practical viewpoint, OBGP offers a valid starting point, and it can be used as a reference model while assessing the performance achieved by alternative inter-domain RWA protocols. The aim of this paper is to introduce a set of minor modifications to a path-vector protocol that can be feasibly implemented in large scale multi-domain optical settings, as is the case of OBGP, and show that these modifications are sufficient to drastically improve the performance of OBGP.

3. The routing model in OBGP+

The routing information exchanged by means of OBGP+ fulfills the following requirements:

(i) PSI must be advertised between domains in addition to the usual reachability information.
(ii) The PSI received from downstream domains must be assembled and aggregated together with local PSI, and advertised to upstream domains.
(iii) This PSI flow must supply a standardized coupling between the different segments along a lightpath. This will support the computation of end-to-end optical paths in an efficient way.
(iv) The PSI exchanged must be completely independent of the intra-domain routing and signaling protocols. In this sense, enhancements or even a complete replacement of any of the protocols used inside a domain must not affect the routing information exchange model between domains.
(v) Special care must be taken while developing aggregated PSI schemes, and while deciding the frequency of the updates associated with the routing information sent across domain boundaries.

---

[1] The performance metric considered in this work is the blocking ratio of inter-domain lightpath requests.

These issues are the subject of the study in this section. A detailed description of the RWA algorithm provided by OBGP+ shall be introduced in Section 4.

OBGP+ is responsible for distributing inter-domain routing information, and deciding, within each domain, the best path to reach a destination. To this end, the routing advertisements in OBGP+ contain the usual NRI in addition to PSI.

The role of the PSI is to capture the "state" of resources along an inter-domain path. During the composition of the advertisements, OBGP+ nodes aggregate the PSI along a path, taking into account the state of both the intra- and the inter-domain segments of the path. The advertised PSI is rich enough so that upstream domains can drastically reduce the number of blocked lightpath requests, and at the same time is sufficiently aggregated so that administrative limits and business protection considerations of domains are respected.

The flow of routing advertisements between OBGP+ nodes from a destination domain or Autonomous System (AS) AS_D toward a source domain AS_S can be summarized as follows. A border node in AS_i assembles the PSI received from AS_{i+1} with its local PSI, and advertises AS_{i-1} the aggregate: $PSI^{(i-1)} = PSI^{(i)} \oplus PSI^{(i+1)}$, where the operator $\oplus$ denotes an appropriate PSI assembling function. The data conveyed in the PSI, as well as the strategy to update them, are detailed later in Section 3.2.

3.1. Network reachability information (NRI)

For simplicity, we assume that the optical nodes, namely, the *Optical Cross-Connects* (OXCs), do not perform wavelength conversion, so each lightpath computed by OBGP+ is subject to the wavelength continuity constraint. We proceed now to describe the NRI and the aggregated PSI conveyed by OBGP+.

Let L, F, and Ω denote the number of links, the number of fibers per-link, and the number of wavelengths per-fiber, respectively, at each destination OXC. For the sake of simplicity, we assume that all destination OXCs are identical, and that each network sinking traffic is connected to only one OXC.[2] Thus, $LF\Omega$ is an upper bound of the number of available wavelengths to reach any destination within a domain. Each AS may select–according to its local TE and routing policies–the particular subset of wavelengths that can be used by an upstream domain to reach the local networks. Consequently, the reachability information contained in the NRI messages conveyed by OBGP+ consists of:

(i) The set of destination networks $\{d\}$ and their associated AS-paths.
(ii) The Next-Hop (NH) to reach those destinations, i.e., the address of the ingress OXC in the neighboring domain from which the advertisement was sent. It is worth noticing that the NH concept is basically the same as in the case of BGP.

---

[2] The routing information exchange model described here can be easily generalized if these assumptions are not met.

(iii) A set of pairs $(\lambda_1, M_{\lambda_1}), \ldots, (\lambda_N, M_{\lambda_N})$ available for each destination d, where $\lambda_i, i \in \{1, \ldots, N\}$ denotes a particular wavelength, and M_{λ_i} denotes the maximum multiplicity of λ_i. Clearly, $N \leq \Omega$, and $M_{\lambda_i} \leq LF \forall i$.

In sum, the NRI distributed between OBGP+ nodes is composed of:

$$\Phi_{NRI}(d) = \left\{ \text{AS-path, NH}, (\lambda_i, M_{\lambda_i}) \right\}_d. \tag{1}$$

For each destination network, a transit AS may filter and advertise a subset of Φ_{NRI} to its upstream domains, or simply retransmit the NRI messages received. When a new destination network becomes available, or an already known one becomes unavailable, the NRI messages are triggered immediately by OBGP+. In any other case, the NRI should only change over large timescales compared to the PSI, according to the local optimizations and TE actions performed by the different routing domains.

Henceforth, and especially during the performance evaluation in Section 5, we assume that both OBGP and OBGP+ handle exactly the same NRI and treat it exactly in the same way.

3.2. Aggregated path-state information (PSI)

The PSI is composed of aggregated wavelength availability information. OBGP+ advertises PSI messages by aggregating and assembling the following three pieces of information:

(i) Intra-domain PSI.
(ii) PSI related to the inter-domain links toward its downstream domains.
(iii) The already aggregated PSI contained in the inter-domain advertisements received from downstream domains.

The aggregation process operates as follows. Let r and q be a pair of OXCs inside an AS, $P(r, q)$ be a candidate path between r and q, and l be a link within the path $P(r, q)$. OBGP+ nodes compute the *Effective Number of Available Wavelengths* (ENAW) of type λ_i between the OXCs r and q as follows:

$$W_{r,q}(\lambda_i) = \max_{P(r,q)} \left\{ \min_{l \in P(r,q)} W_l(\lambda_i) \right\}. \tag{2}$$

The rationale in (2) can be easily interpreted by means of the Fig. 1. For instance, in AS1 the ENAW of type λ_1 between the nodes *OXC* 15 and *OXC* 12 is $W_{15,12}(\lambda_1) = 3$. This is because, from the two possible paths between these nodes, the path that goes through *OXC* 13 has a minimum $W_{13,12}(\lambda_1) = 1$, whereas the one that goes through *OXC* 11 has a minimum $W_{11,12}(\lambda_1) = 3$. Then, the maximum between both of them is 3. The ENAW given in (2) is especially important between two border OXCs in a transit domain, since it conservatively captures the practical availability of wavelength λ_i within the domain. In addition, (2) offers highly aggregated network state information, so this is the intra-domain portion of the wavelength availability component of a PSI aggregate.
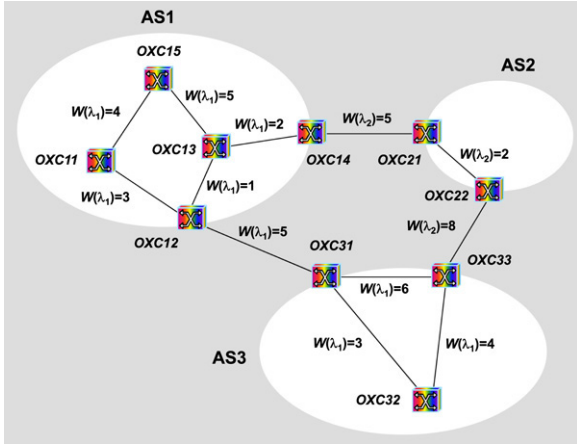
Fig. 1. Path-State Information (PSI) exchanged between optical domains.

For the inter-domain portion, each OBGP+ node is aware of which wavelengths are actually being used on its inter-domain links, and it also knows which wavelengths are effectively available downstream through the PSI advertisements received from neighboring OBGP+ nodes. Let $W_{l_b,r_b}(\lambda_i)$ denote the number of available wavelengths of type λ_i in the inter-domain link between the local border node l_b, and a neighbor border node r_b. For instance, in Fig. 1 the OBGP+ nodes in AS1 are aware that $W_{12,31}(\lambda_1) = 5$. Similarly, let $W_{r_b,d}^{adv}(\lambda_i)$ denote the ENAW of type λ_i between the neighbor border node r_b and the destination node d, advertised by r_b. Using these two inter-domain components and (2), an OBGP+ node advertises upstream that the ENAW between a local border node l_b and a distant destination node d is:

$$W_{l_b,d}^{adv}(\lambda_i) = \min\left\{W_{l_b,l_b'}(\lambda_i), W_{l_b',r_b}(\lambda_i), W_{r_b,d}^{adv}(\lambda_i)\right\}_d. \quad (3)$$

For instance, in Fig. 1 the border node $OXC14$ advertises to its neighbor $OXC21$ in AS2 that the ENAW of type λ_1 to reach $OXC32$ is:

$$W_{14,32}^{adv}(\lambda_1) = \min_{\lambda_1}\left\{W_{14,12}, W_{12,31}, W_{31,32}^{adv}\right\}$$
$$= \min\{2, 5, 4\} = 2. \quad (4)$$

In summary, the PSI received by an OBGP+ node for destination d is the set of ENAWs:

$$\Phi_{PSI}(d) = \left\{W_{r_b,d}^{adv}(\lambda_i)\right\}_d. \quad (5)$$

To advertise the PSI associated with the destinations contained in the NRI messages, we take advantage of the *Keepalive* messages exchanged between neighboring OBGP+ nodes. Similarly to the case of BGP, OBGP+ nodes exchange Keepalive messages to confirm that neighboring nodes are still operative. In BGP, Keepalive messages are of fixed length, consisting only of the 19-byte BGP header. In our OBGP+ model, we extend the BGP Keepalive concept with the purpose of conveying PSI, when relevant PSI

needs to be updated. In other words, the update of PSI is supported by the exchange of non-dummy Keepalive messages between routing domains. A major advantage of this strategy is that the update of PSI can be made without incrementing the number of routing messages exchanged between domains.

In Section 5 we shall show that when the RWA algorithm supporting OBGP+ exploits the highly aggregated PSI in (5), it is possible to achieve drastic reductions in the number of blocked inter-domain lightpath requests compared to that obtained with OBGP.

4. Routing and wavelength assignment strategy

Similarly to BGP, OBGP is essentially a shortest AS-path routing algorithm that exchanges NRI, but it does not handle PSI. Our OBGP+, however, handles the highly aggregated PSI supplied by the ENAW introduced in Section 3.2. Accordingly, each OBGP+ node computes and advertises the ENAW along the candidate paths, as described in (2) and (3). Algorithm 1 shows a simplified version of the OBGP+ decision process. This algorithm is the result of a set of enhancements that we introduced to OBGP [1].

From Algorithm 1, it is clear that OBGP+ is essentially a "shortest AS-path highest ENAW" RWA algorithm, given that it usually prefers the shortest AS-path (step 2 of the algorithm), but if more than one candidate lightpath exists, then it chooses the one with the highest ENAW (step 3).

Algorithm 1   OBGP+($\{P(s, d), \lambda_i, M_{\lambda_i}, z_i\}$)

Input: $\{P(s, d)\}$ − set of paths between nodes s and d
  λ_i − a particular wavelength on path $P(s, d)$
  M_{λ_i} − Multiplicity of wavelength λ_i on path $P(s, d)$
  z_i − ENAW of type λ_i along the path $P(s, d)$

Output: $(P^{best}, \lambda^{best})$ −The best lightpath between s and d

1: Choose the (path, wavelength) pair with the highest local preference (LOCAL_PREF)   /* As in BGP */.
2: If the LOCAL_PREFs are equal, choose the shortest AS-path and assign the wavelength with the highest ENAW among the ones available on that path. If more than one wavelength has the same (highest) ENAW along the shortest AS-path, choose the wavelength with the lowest identifier i.
3: If the AS-path lengths are equal choose the (path, wavelength) pair associated with the highest ENAW.
4: If the ENAWs are equal prefer external paths over internal paths.
5: If the paths are still equal prefer the one with the highest ENAW to the next-hop OXC (i.e., to the OXC r_b in the neighboring domain).
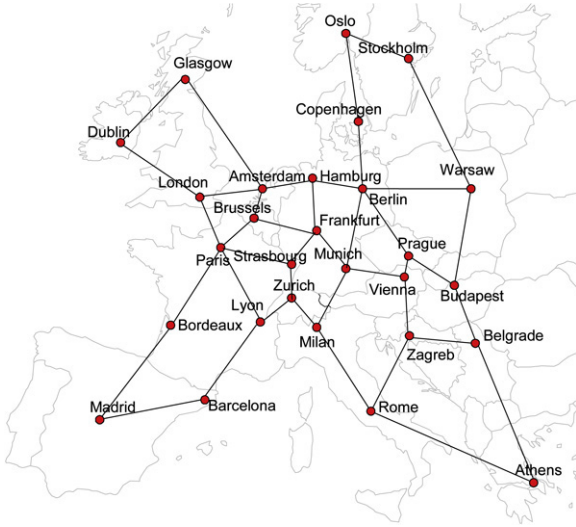6: If more than one path is still available run OBGP tie-breaking rules   /* As in BGP */.

Fig. 2. PAN European network topology.

5. Performance evaluation

The aim of this section is to contrast the performance of OBGP+ against OBGP. Our interest here is to compare four different performance metrics, namely, (i) the *Blocking Ratio* (BR) of inter-domain lightpath requests, (ii) the overall number of routing messages exchanged to achieve this blocking, (iii) the convergence and overall time needed to restore a set of connections after a link failure, and (iv) the number of routing messages exchanged during a convergence.

To this end, we have conducted extensive simulations using OPNET Modeler [15]. The simulation results presented here can be reproduced using the modules that we developed in OPNET, which are available online from [16]. The inter-domain scenario chosen for the trials was the PAN European network topology illustrated in Fig. 2 [17]. This multi-domain network is composed of 28 domains and 41 inter-domain links.

For the network topology inside each domain in the PAN, we have randomly chosen a minimum number of OXCs equal to the number of inter-domain links of that domain, up to a maximum of 6 OXCs inside each domain. This approach guarantees that each inter-domain link of a domain in the PAN is supported by a different border OXC. The OXCs inside each domain were connected by a full-mesh. In this setting, we have randomly placed 18 sources and 10 destinations, covering in this way the entire PAN European network with a source or a destination OXC inside each of its 28 domains. The overall set up process was repeated 100 times, and the results shown here are the averages over those 100 settings. It is worth emphasizing that these results are the outcome of more than 1800 h of event-driven simulations, so our criterion to choose a maximum of 6 OXCs per-domain reflects the trade-off between the size of the network, and the time needed to run the tests.

We have used 5 fibers per-link, and 12 wavelengths per-fiber throughout the entire PAN European network. In order to assess the impact of the frequency of update in the PSI, we have used different *Keepalive Update Intervals* K_T during the trials. K_T corresponds to the time interval between the delivery of non-dummy Keepalive messages conveying PSI. At present, most implementations of BGP use a default Keepalive value of 60 s, and three consecutive Keepalive messages need to be lost so that a BGP router proceeds to shut down a BGP session. In our simulations, we have tested three different scaled and normalized values: $K_T = 1$, $K_T = 3$, and $K_T = 5$ units through the simulation run-time. Clearly, the higher the values of K_T, the more time is needed by OBGP+ nodes to detect and react when a neighbor becomes inoperative. Therefore, a major advantage of conveying PSI piggy-backed in Keepalive messages, is that low values of K_T are desired, both to increase the responsiveness between OBGP+ neighbors as well as to support updating PSI more frequently.

5.1. Assessing the Blocking Ratio (BR)

As shown in Fig. 3, the trials were performed for different traffic loads, varying from 100 Erlangs up to 300 Erlangs, where traffic was modeled according to a Poisson distribution with exponentially distributed arrival and departure rates. Fig. 3 shows the BR and the standard deviation of inter-domain lightpath requests obtained with OPNET, for the different traffic loads, and the different Keepalive update intervals K_T. Clearly, OBGP+ outperforms OBGP, and it becomes evident that even minor PSI, like the one proposed in this paper, is enough to drastically reduce the blocking obtained $\forall K_T$. Whereas OBGP experiences blocking for all traffic loads tested, OBGP+ starts to show some negligible blocking only after reaching 200 Erlangs.

Fig. 3 also shows that the performance of both OBGP and OBGP+ degrade when K_T increases. The reason for this in the case of OBGP is the following. In OBGP [1], a source node is not aware of the subset of wavelengths $\mathcal{W}(P)$ that are no longer available along the different segments of an AS-path P. A source OBGP node will receive a reachability message indicating the withdrawal of path P, only after all the candidate wavelengths in that path have been consumed. A simple way to considerably reduce the BR experienced by OBGP is to update the subset $\mathcal{W}(P)$ through the Keepalive messages exchanged between OBGP neighbors. This approach not only offers more granular and updated NRI at the source OBGP node, but also, allows us to compare the performance of OBGP and OBGP+ under fairer conditions. Our implementation of OBGP in OPNET follows this approach. As expected, Fig. 3 confirms that the blocking experienced by OBGP increases with K_T. The corollary is that in a regular implementation of OBGP (i.e., where the subset $\mathcal{W}(P)$ is not conveyed and updated through the Keepalive messages) the BR yield by OBGP becomes independent of K_T, but it is much higher than that shown for $K_T = 5$ in Fig. 3.

In order to quantify the reductions supplied by OBGP+ in terms of blocking, we define the following *Blocking Improvement Factor* (IF):

$$\text{BIF} \triangleq \left(\frac{\text{BR}^{(OBGP)}}{\text{BR}^{(OBGP+)}} \right)_{\text{Traffic (Erlangs)}}. \tag{6}$$
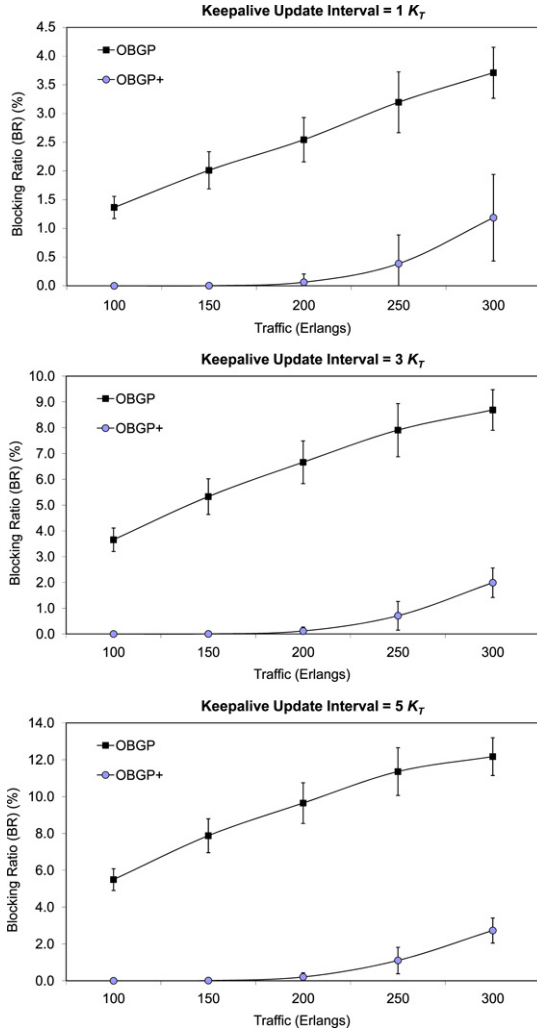
Fig. 3. Blocking ratio and standard deviation: Comparison between OBGP and OBGP+ for the different traffic loads and different Keepalive Update Intervals (K_T).

Table 1 summarizes the blocking improvement factor BIF for the different traffic loads and Keepalive update intervals assessed. The results show that:

- Up to 100 Erlangs, OBGP+ does not introduce blocking, so BIF = ∞.
- Up to 150 Erlangs, OBGP+ is able to reduce the BR by more than two orders of magnitude.
- For 200 Erlangs, OBGP+ is able to reduce the BR by more than one order of magnitude.
- For 250 Erlangs, OBGP+ is able to reduce the BR by approximately one order of magnitude.
- Even for the highest traffic load assessed, i.e., 300 Erlangs, OBGP+ achieves an improvement factor that roughly varies between 3 and 5.

It is worth mentioning that the purpose of Fig. 3 is to compare two different RWA strategies, so the values of the BRs per se should not be taken as representatives of those expected in operational networks. Studies like the one developed by the IST European project NOBEL [18], have

Table 1
Improvement Factor in the blocking requests for the different traffic loads and Keepalive update intervals assessed.

Blocking Improvement Factor (BIF) between OBGP and OBGP+			
Traffic (Erlangs)	Keepalive update interval		
	$K_T = 1$	$K_T = 3$	$K_T = 5$
100	∞	∞	∞
150	643.94	1239.50	567.94
200	40.42	57.74	46.50
250	8.27	11.18	10.28
300	3.13	4.37	4.46

taxonomized the expected performance in the dynamic provisioning of lightpaths for different applications. For instance, [18] recommends for real-time and streaming applications a blocking less than or equal to 0.1%. Fig. 3 shows that OBGP is unable to reach this bound for all the simulation conditions tested. To reach such bound, extra resources (e.g., more wavelengths and/or fibers) would be needed in the PAN-European network in the case of OBGP. Conversely, OBGP+ is able to reach the 0.1% bound for 150 Erlangs for the three values of K_T, and even for 200 Erlangs when $K_T = 1$. These findings show that even minor modifications to a path-vector protocol, can be capitalized in considerable CAPEX reductions.

5.2. Assessing the number of messages exchanged

Fig. 4 confirms that OBGP+ always needs less overall number of routing messages than OBGP. The reason for this is twofold. First, PSI updates are never triggered between OBGP+ neighbors. Instead, they are piggy-backed in the Keepalive messages used in both OBGP and OBGP+. Second, OBGP tends to exhaust the available wavelengths along the shortest AS-path before switching to an alternative path. This triggers network reachability messages and path exploration after paths become blocked. Conversely, OBGP+ explicitly considers the ENAW in the RWA algorithm when two or more paths exhibit the same AS-path length, so it is able to provide a much better traffic distribution than OBGP, with drastic reductions in the blocking, and hence less network reachability messages need to be exchanged.

Similarly as we did for the blocking, we define the following *Message Improvement Factor* (MIF), in order to quantify the reductions supplied by OBGP+ in terms of the overall number of messages exchanged:

$$\text{MIF} \triangleq \left[\left(\frac{\text{Messages}^{(OBGP)}}{\text{Messages}^{(OBGP+)}} \right)_{\text{Traffic (Erlangs)}} - 1 \right] \times 100\%. \quad (7)$$

Table 2 summarizes the message improvement factor MIF for the different traffic loads and Keepalive update intervals assessed. The results show that OBGP needs between 92% more messages than OBGP+ (for 200 Erlangs and $K_T = 1$), and 6% more for the most demanding condition assessed (i.e., for a traffic of 300 Erlangs and a Keepalive $K_T = 5$). The results also reveal that the largest improvements are obtained for $K_T = 1$ (i.e., when PSI information is frequently updated among OBGP+ neighbors) and medium traffic loads (between 150 and 250 Erlangs).
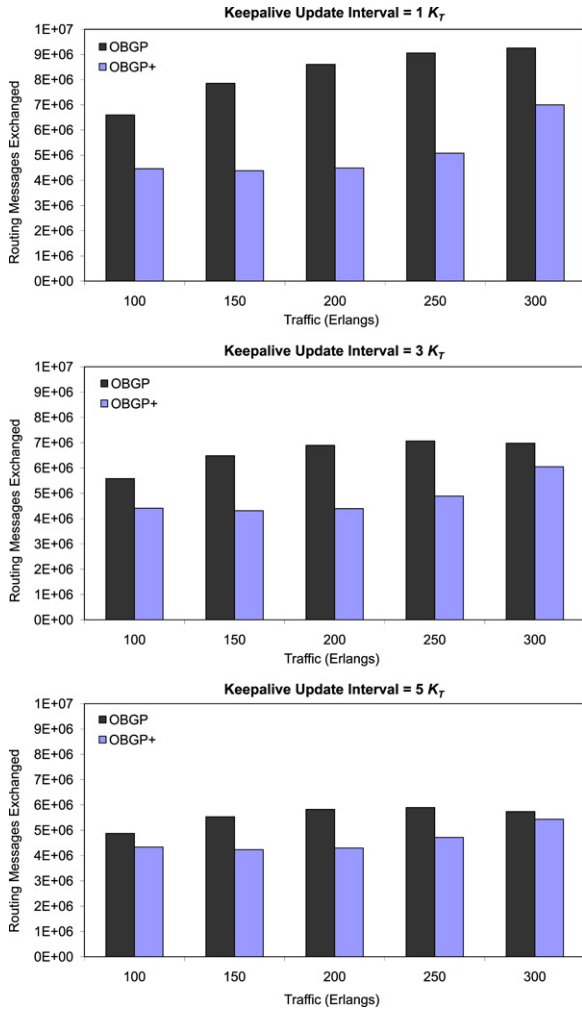
Fig. 4. Number of routing messages exchanged for the different traffic loads and different Keepalive Update Intervals (K_T).

Table 2
Improvement factor in the number of messages exchanged for the different traffic loads and keepalive update intervals assessed.

Improvement in the number of messages exchanged between OBGP and OBGP+			
Traffic (Erlangs)	Keepalive update interval		
	$K_T = 1$	$K_T = 3$	$K_T = 5$
100	48%	27%	12%
150	79%	50%	31%
200	92%	57%	36%
250	78%	45%	25%
300	32%	15%	6%

5.3. Assessing the convergence time of OBGP+

The goal of this section is to compare the convergence time of OBGP and OBGP+. To this end, we have randomly chosen an inter-domain link in the PAN European Network topology, and considered two different transitions between states of that link. The first one is what we called ON–OFF, which represents the transition between the normal operation of the link, and a failing state. The second
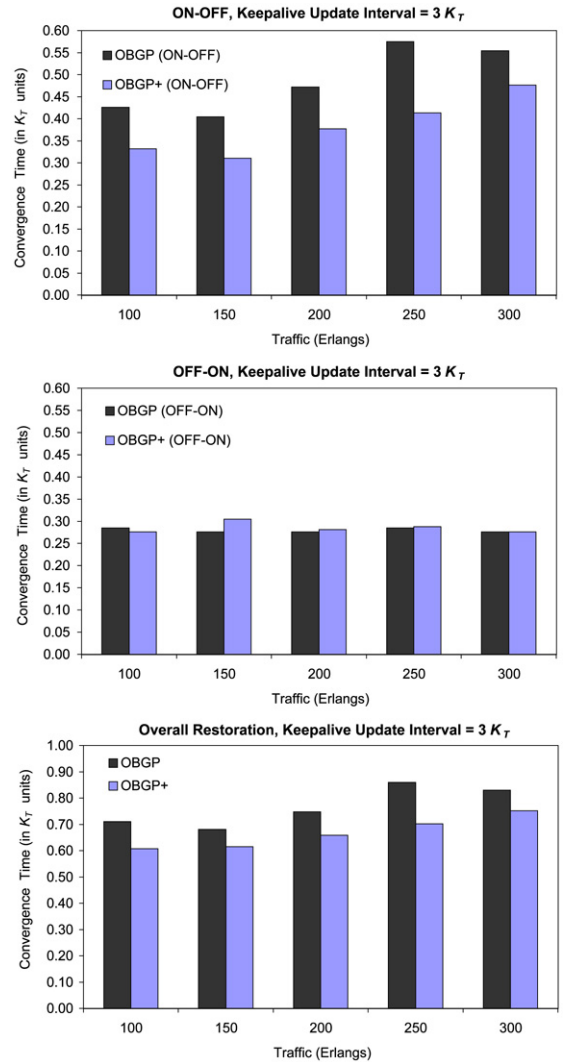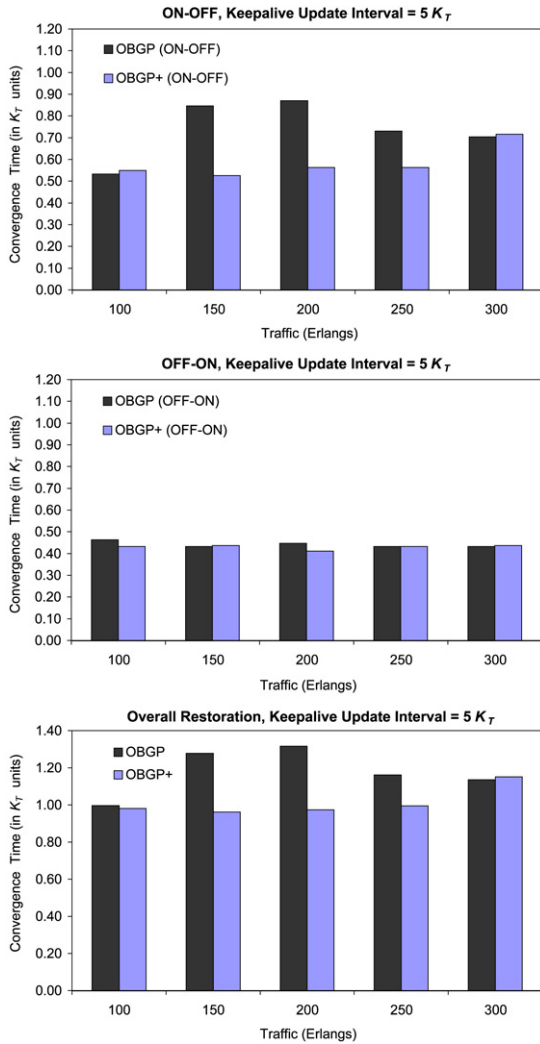
Fig. 5. (Top) Convergence time ON–OFF; (Center) Convergence time OFF–ON; (Bottom) Overall time needed to restore the initial state of the network (Keepalive Update Interval of 3 K_T).

one is what we called OFF–ON, and represents the opposite transition, i.e., between the failing state and the normal operation of the link.

In order to assess the convergence time of both protocols, we have considered the time that elapses between the advertisement (withdrawal) of a new (known) destination, and the time when the last message originated by the OFF–ON (ON–OFF) event is processed.

Figs. 5 and 6 show the convergence time of OBGP and OBGP+ normalized in K_T units, for 3 K_T and 5 K_T, respectively. We have omitted the results for 1 K_T, since they do not provide additional information from that contained in Figs. 5 and 6. The results shown at the top of these figures represent the ON–OFF transition, while the ones at the center show the OFF–ON transition. Our results confirm that the ON–OFF transition is the most demanding in time, since it is in this transition when path exploration occurs. The results shown at the bottom of Figs. 5 and 6 correspond to the accumulated time for an ON → OFF →
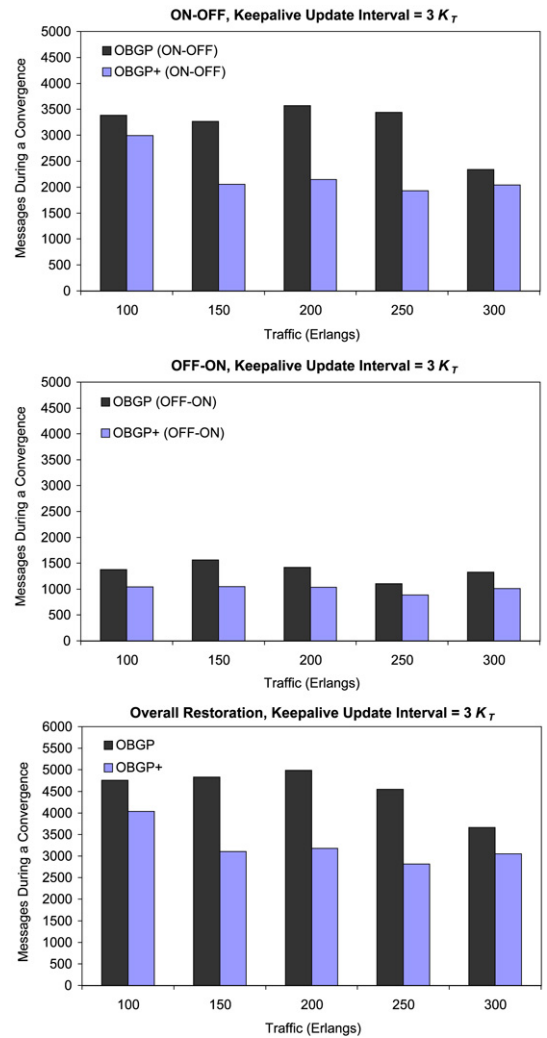
Fig. 6. (Top) Convergence time ON–OFF; (Center) Convergence time OFF–ON; (Bottom) Overall time needed to restore the initial state of the network (Keepalive Update Interval of 5 K_T).

Fig. 7. (Top) Number of routing messages exchanged during a convergence ON–OFF; (Center) Number of routing messages exchanged during a convergence OFF–ON; (Bottom) Overall number of routing messages exchanged during the restoration of the initial state of the network (Keepalive Update Interval of 3 K_T).

ON transition. The particular interest for these values is that they represent lower bounds for the time needed by the protocols to advertise and restore the initial state of the network, i.e., without considering the offset introduced by the transient state OFF.

The results obtained show that OBGP+ needs to explore less paths, hence converges faster than OBGP for the ON–OFF transition. For the OFF–ON case, the protocols offer almost the same performance, but the time saved by OBGP+ in the ON–OFF transition translates in lower bounds while restoring the initial state of the network in a cycle ON → OFF → ON.

5.4. Assessing the number of messages exchanged during a convergence

Following the same approach described in Section 5.3, this section aims at providing insight on the number of routing messages exchanged by both OBGP and OBGP+

during a convergence state. The results are shown in Figs. 7 and 8. Once again, we observe that the majority of the messages are exchanged during the ON–OFF transition, i.e., when path exploration takes place. We also confirm that OBGP+ needs to process fewer routing messages, which is consistent with the previous finding that OBGP+ shows shorter convergence times than OBGP.

6. Conclusions and future work

In this paper, we have shown that by endowing a path-vector protocol like OBGP with the capability to compute, aggregate, and convey only minor path-state information, it is possible to drastically reduce its blocking ratio. We have shown that these significant improvements can be achieved without needing to exchange more routing messages than with OBGP. In fact, OBGP+ reduces the number of routing messages exchanged between optical
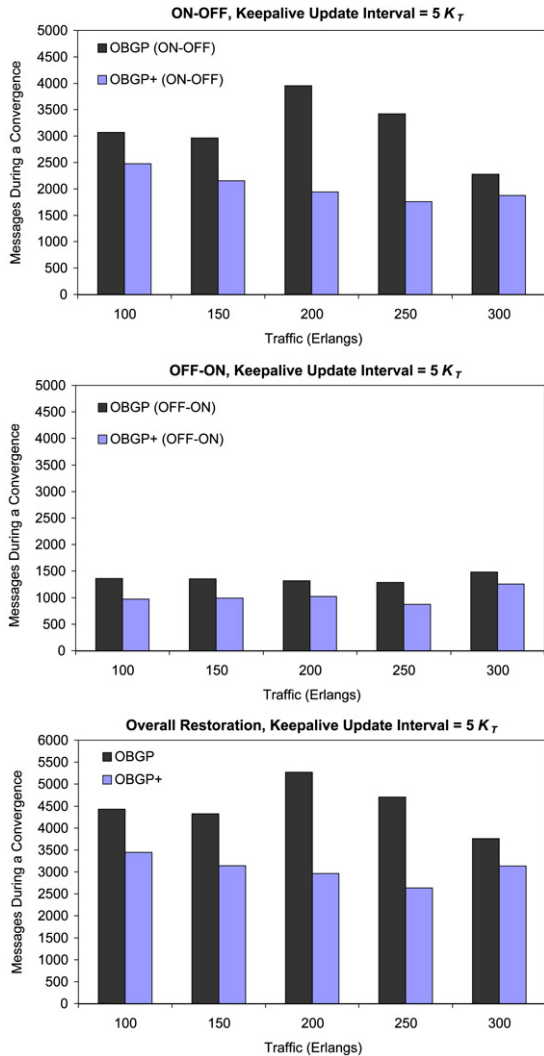
Fig. 8. (Top) Number of routing messages exchanged during a convergence ON–OFF; (Center) Number of routing messages exchanged during a convergence OFF–ON; (Bottom) Overall number of routing messages exchanged during the restoration of the initial state of the network (Keepalive Update Interval of 5 K_T).

domains, given that by decrementing the blocking, it is possible to reduce the exchange of network reachability messages and path exploration when blocking starts to occur. This is possible due to the strategy of piggy-backing PSI updates in the Keepalive messages exchanged between OBGP+ neighbors. We have also analyzed the convergence and restoration properties of OBGP+. Our results reveal that OBGP+ not only converges faster than OBGP, but also exchanges less number of routing messages than the latter during a convergence.

Although these are promising findings, more research is needed in this direction. Our main results and conclusions still apply to a rather small multi-domain optical network (the PAN European topology shown in Fig. 2), so further studies are needed to analyze the performance of the proposals made here in a large-scale environment composed by thousands of ASs. We plan to explore this in the near future.

Overall, multi-domain RWA approaches like the one proposed here, offer a promising line of work to address the trade-off between obtaining a low blocking ratio, and keeping the path-state information as limited as possible.

Acknowledgments

References

[1] M. Blanchet, F. Parent, B. St-Arnaud, Optical BGP (OBGP): InterAS Lightpath Provisioning, IETF draft, ietf-draft-parent-obgp-01, March 2001.
[2] M.J. Francisco, L. Pezoulas, C. Huang, I. Lambadaris, End-to-end signaling and routing for optical IP networks, in: Proceedings of IEEE ICC, NY, USA, April 2002.
[3] L. Wang, et al. A Novel OBGP-based mechanism for Lightpath Establishment in WDM mesh networks, in: Proceedings of ECOC 2003, Rimini, Italy, September 2003.
[4] L. Wang, H. Zhang, L. Zheng, Reducing the OBGP protection switching time in WDM mesh networks, in: Proceedings of OFC, Anaheim, CA, USA, March 2006.
[5] M. Yannuzzi, X. Masip-Bruin, O. Bonaventure, Open issues in interdomain routing: A survey, IEEE Network 19 (6) (2005) 49–56.
[6] Future Internet, The future networked society: A white paper from the EIFFEL Think-Tank, available from: http://future-internet.eu/.
[7] Future Internet Design: http://www.nets-find.net/.
[8] International Telecommunications Union (ITU): http://www.itu.int/.
[9] Internet Engineering Task Force (IETF): http://www.ietf.org/.
[10] Optical Internetworking Forum (OIF): http://www.oiforum.com/.
[11] G. Bernstein, et al. Domain to domain routing using GMPLS, OSPF extension V1.1(Draft), OIF2002.23.06, July 2002.
[12] A. Farrel, J.P. Vasseur, A. Ayyangar, A framework for inter-domain multiprotocol label switching traffic engineering, IETF RFC 4726, November 2006.
[13] G. Liu, C. Ji, V. Chan, On the scalability of network management information for inter-domain light-path assessment, IEEE/ACM ToN 13 (1) (2005).
[14] M. Yannuzzi, X. Masip-Bruin, G. Fabregó, S. Sánchez-López, A. Sprintson, A. Orda, Toward a new route control model for multi-domain optical networks, IEEE Communications Magazine 46 (6) (2008) 104–111.
[15] Opnet Technologies, Inc.: http://www.opnet.com/.
[16] Opnet Modules developed: http://www.craax.ctvg.upc.es/opnet/opnet-craax.html.
[17] R. Hulsermann, et al., A set of typical transport network scenarios for network modelling, ITG Fachbericht (182) (2004).
[18] NOBEL: Next generation optical networks for broadband European leadership, IST Integrated Project (FP6-506760): http://www.ist-nobel.org/.