

# Toward a New Route Control Model for Multidomain Optical Networks

*Marcelo Yannuzzi, Xavi Masip-Bruin, Guillem Fabregó, and Sergio Sánchez-López,*

*Technical University of Catalonia*

*Alex Sprintson, Texas A&M University*

*Ariel Orda, Technion — Israel Institute of Technology*

## ABSTRACT

The design of the control plane model for multidomain optical networks poses complex challenges and introduces many open problems. Some initiatives have proposed Optical BGP (OBGP), which is an extension of BGP supporting the advertisement and signaling of optical information between routing domains. We argue, however, that future optical networks offer the opportunity to avoid inheriting the limitations of BGP, especially in terms of routing and traffic engineering control. In this article we present a route control model replacing BGP/OBGP. Extensive simulations confirm that our route control model is able to drastically reduce the blocking experienced with OBGP, and this can be accomplished without increasing the number or frequency of routing updates exchanged between domains.

## INTRODUCTION

In the coming years, customers will no longer be limited to purchase monthly or yearly contracts for one-time capacity allocation. Instead, providers will be expected to offer end-to-end optical connections in real time and for shorter periods, such as hours or even minutes (e.g., for resilience reasons). Addressing these requirements in the framework of a multidomain optical network represents a challenging problem. As a first step in this direction, some researchers have started to analyze the possibility of adopting the Optical Border Gateway Protocol (OBGP) as the future interdomain routing protocol for optical networks [1–3]. The aim of these proposals is to extend BGP [4] so that it can convey and signal optical information between OBGP neighbors. The strength of this approach is that optical networks will benefit from the advantages of the BGP-based route control model (e.g., proven scalability). The weakness, on the other hand, is that the routing model will inherit the well-known issues in BGP, such as:

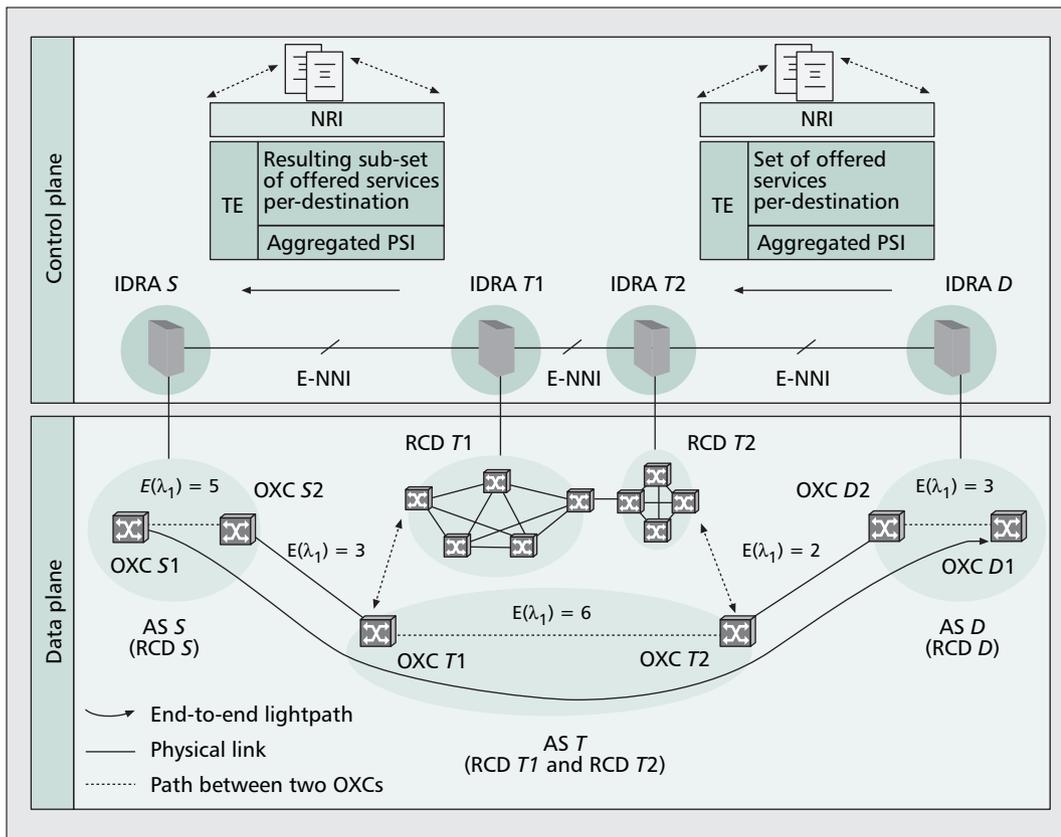
- The inability to convey useful traffic engineering (TE) information and exploit it in practice
- Lack of multipath routing capabilities
- Slow convergence and chattiness, impeding fast detection and response to network impairments

In brief, we argue that a multidomain routing model mostly centered on the exchange of network reachability information (NRI) — like the one we currently have with BGP or the one offered by OBGP — will not be sufficient. It is widely accepted now that, in addition to NRI, neighboring domains should be able to exchange aggregated path state information (PSI), and this requirement will also be present in next-generation optical networks [5]. We argue in favor of a change, with emphasis on borrowing the best of BGP while avoiding its key limitations. In particular, it is necessary to investigate how to design a distributed control plane with the ability to compute, convey, and efficiently exploit aggregated PSI in a multidomain setting.

In this context we present a new model for multidomain optical networks that blends interdomain routing and TE control. We describe its architecture, as well as the details of the NRI and PSI exchanged between domains. Our model tries to extend neither BGP nor OBGP, but rather proposes an alternative approach in place of them. We also introduce a routing and wavelength assignment (RWA) strategy that exploits the NRI and PSI to compute interdomain lightpaths in a highly effective way.

## A NEW ROUTE CONTROL MODEL

Our multidomain route control model is supported by a fully distributed and decoupled control plane. The decoupling consists of clearly splitting the control and data planes, where independent circuits are used to physically connect the nodes within the control plane. Other highly scalable and successful networks have relied on this kind of separation in the past, such as Sig-



■ **Figure 1.** Architecture of the IDRA-based routing and TE control model.

Our IDRAs will act as the glue between the intradomain and interdomain routing schemes of RCDs. They can compute primary and backup lightpaths subject to performance and/or reliability constraints, using the TE information received through the reference points of their RCD.

naling System #7 used in telephony. In the future, having dedicated fibers and nodes to distribute routing and signaling information between routing control domains (RCDs) is in fact desired. This approach leverages the evolution toward more advanced and reliable routing and TE control models, where the key is to release the traffic forwarders from the burden of exchanging control information and performing complex computations based on it.

### THE ROUTE CONTROL ARCHITECTURE

The nodes within our route control model are called interdomain routing agents (IDRAs) [6]. Each RCD may allocate one or more IDRAs depending on its scale; for example, Fig. 1 shows a source domain *S*, a destination domain *D*, and a transit provider *T* that is split into two RCDs, *T*<sub>1</sub> and *T*<sub>2</sub>.

The role of the IDRAs is twofold. On one hand, they are the ones that distribute the routing and signaling information between RCDs. On the other hand, they are in charge of the computation and establishment of interdomain lightpaths in a distributed way — similar to the path computation element (PCE) model [7]. A major difference, however, between the IDRA-based and PCE-based route control models is that BGP/OBGP is not present in the case of IDRAs.

Our IDRAs will act as the glue between the intradomain and interdomain routing schemes of RCDs. They can compute primary and backup lightpaths subject to performance and/or reliability constraints, using the TE information received

through the *reference points* of their RCD. A reference point represents one of the three standardized interfaces used to connect optical networks: the user-network interface (UNI), internal network-network interface (I-NNI), and external network-network interface (E-NNI). In particular, the E-NNI supports the communication and signaling operations either between different RCDs within the same autonomous system (AS) or between different ASs. The visibility of the internal resources within each RCD is controlled by the policies ruling the exchange of information through the E-NNIs.

### ROUTING AND TE INFORMATION EXCHANGE MODEL

The establishment of an interdomain lightpath is performed in three phases: routing, signaling, and setup. During the routing phase, the IDRA in the source domain uses the information advertised by neighboring IDRAs to find a loose end-to-end lightpath between the local optical node requesting the path and the destination node. In the second phase the source IDRA signals the chosen path to the IDRAs of the RCDs traversed by this path. The setup of the lightpath takes place in the third phase, where the IDRA at each of the traversed RCDs is responsible for establishing the portion of the path that runs through its RCD.

The advertisements distributed by the IDRAs contain the usual NRI, in addition to TE information consisting of PSI and the set of offered *services* by the RCDs along a path (Fig. 1). Dur-

In order to preserve the confidentiality across RCD boundaries, as well as to provide loop-free paths, the routes advertised by the IDRAs consist of a set of loose hops, wherein each intradomain subpath is abstracted as a single hop.

ing the composition of the advertisements, the IDRAs aggregate the PSI along a path, taking into account the state of both the intradomain and interdomain segments of the path. The role of the services is to endow the route control model with the capability for RCDs to exchange more complex TE data structures. For instance, an RCD could advertise that it offers wavelength conversion for a group of destinations, while it can offer multihop traffic grooming for another group. In our model an IDRA can determine the best path to reach a destination based on a number of specific requirements in terms of services, based on the PSI along the candidate routes, or based on a combination of them.

### NETWORK REACHABILITY INFORMATION

For simplicity of exposition, we assume that the optical nodes, that is, the optical cross-connects (OXC), do not perform wavelength conversion, so each lightpath computed by the IDRAs is subject to the *wavelength continuity constraint*. Each domain may select — according to its TE and routing policies — the subset of wavelengths that can be used to reach the local networks. In this framework, the information contained in the NRI messages sent by an IDRA consists of:

- The set of destinations.
- The next hop (NH) to reach those destinations (i.e., the address of the ingress OXC in the RCD from which the advertisement was sent)
- A set of pairs  $(\lambda_i, M_{\lambda_i})$  for each destination, where  $\lambda_i$  denotes a particular wavelength,  $i$  denotes the wavelength's identifier, and  $M_{\lambda_i}$  denotes the maximum multiplicity advertised for  $\lambda_i$ , since a destination OXC might be connected through multiple fibers

When a new destination becomes available, or an already known one becomes unavailable, the NRI messages are triggered immediately by an IDRA. Conversely to BGP, the NRI exchanged among the IDRAs does not include the AS-path to reach a destination. In our model, rather than comparing candidate routes according to the length of the AS-path, the IDRAs use the TE information contained in the routing advertisements to compare the routes. However, the performance of the RWA strategy used by the IDRAs will substantially depend on the length of the lightpaths chosen, so the path length is embedded in the PSI messages exchanged between the IDRAs. Another important difference between BGP and the IDRA-based model is that instead of advertising only the “best” route for any given destination, the IDRAs can advertise multiple routes per destination, even with the same NH address.

### PATH STATE INFORMATION

In order to preserve confidentiality across RCD boundaries, as well as to provide loop-free paths, the routes advertised by the IDRAs consist of a set of loose hops, wherein each intradomain subpath is abstracted as a single hop. For example, in Fig. 1 IDRA  $T1$  advertises IDRA  $S$  the path  $P(T1, D1)$ , including the list of loose hops:  $OXC T1 \rightarrow OXC T2 \rightarrow OXC D2 \rightarrow OXC D1$ .

Each path advertised by the IDRAs has associated PSI, which is composed of *aggregated*

*wavelength availability* and *aggregated load information*. In this article we present a very simple approach in which both aggregates are integer values. The IDRAs advertise PSI messages by aggregating and assembling the following three pieces of information:

- Intradomain PSI
- PSI related to the interdomain links toward its downstream domains
- The already aggregated PSI contained in the interdomain advertisements received from downstream domains

We proceed to describe how this process is performed.

### Aggregated Wavelength Availability Information

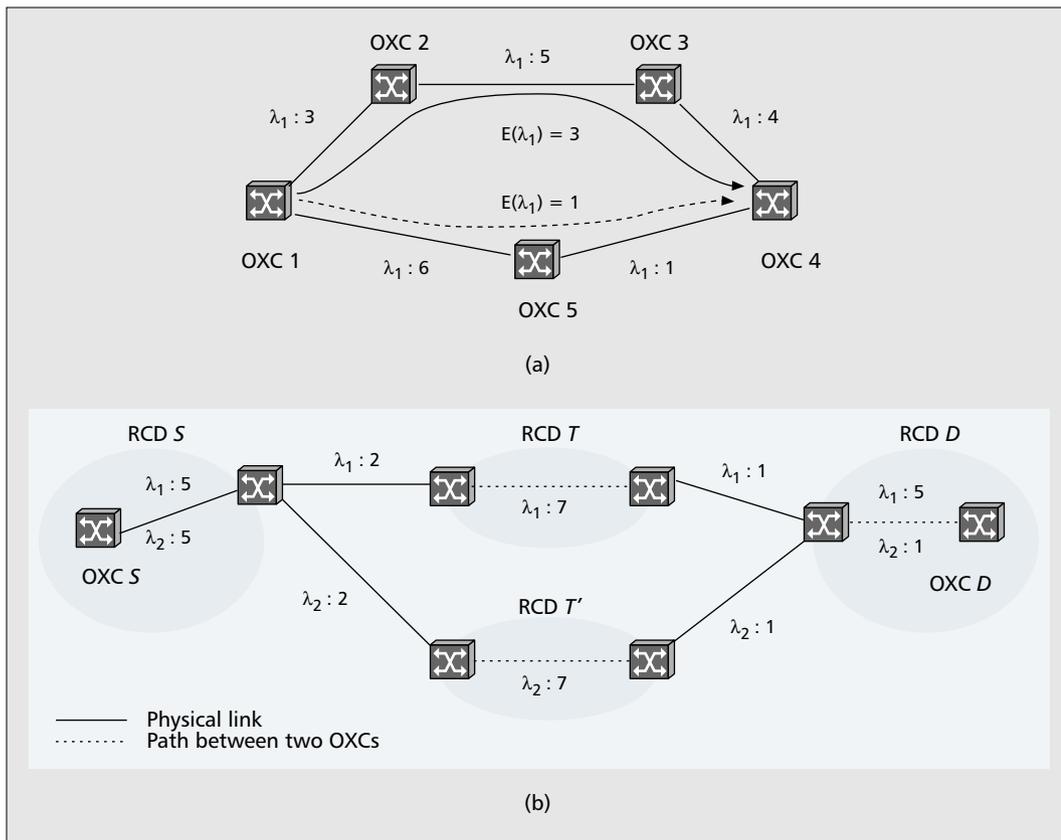
— For any pair of OXCs inside an RCD, the local IDRA computes the effective number of available wavelengths (ENAW) of type  $\lambda_i$  between them, as shown in Fig. 2a. In this example there are two candidate paths between  $OXC 1$  and  $OXC 4$ . Figure 2a also shows the number of available wavelengths of type  $\lambda_1$  on each link. For the path that runs through  $OXC 2$ , the number of wavelengths  $\lambda_1$  that can be effectively used between  $OXC 1$  and  $OXC 4$  is 3. This is because at most three lightpaths can be established between  $OXC 1$  and  $OXC 4$  through  $OXC 2$  using  $\lambda_1$ . Similarly, for the path that runs through  $OXC 5$ , the number of wavelengths  $\lambda_1$  that can be used between  $OXC 1$  and  $OXC 4$  is 1. The ENAW  $\lambda_1$  between  $OXC 1$  and  $OXC 4$  is computed by an IDRA as the maximum between these two (i.e.,  $E_{1,4}(\lambda_1) = 3$ ). In summary, the computation of the ENAW is a simple process, where an IDRA first keeps the minimum number of available wavelengths on the links of a candidate path, and then computes the maximum among all candidates.

The ENAW is especially important between two border OXCs in a transit domain, since it “conservatively” captures the practical availability of wavelength  $\lambda_i$  within the domain. In addition, it offers highly aggregated state information (it is an integer number), so this is the intradomain portion of the wavelength availability component of a PSI aggregate.

For the interdomain portion, each IDRA is aware of which wavelengths are being used on its interdomain links, and it also knows which wavelengths are available downstream through the PSI advertisements received from neighboring IDRAs. Let  $E_{l_b, r_b}(\lambda_i)$  denote the number of available wavelengths of type  $\lambda_i$  in the interdomain link between the local border  $OXC l_b$ , and a remote border  $OXC r_b$ . Similarly, let  $E_{r_b, d}^{adv}(\lambda_i)$  denote the ENAW of type  $\lambda_i$  between the remote border  $OXC r_b$  and the destination  $OXC d$ , advertised by the downstream IDRA in  $r_b$ 's domain.

Using these two interdomain components and the intradomain PSI, an IDRA advertises upstream that the ENAW  $\lambda_i$  between a local border  $OXC l_b'$  and a distant destination  $d$  is  $E_{l_b', d}^{adv}(\lambda_i) = \min \{E_{l_b', l_b}(\lambda_i), E_{l_b, r_b}(\lambda_i), E_{r_b, d}^{adv}(\lambda_i)\}$ . For instance, in Fig. 1 IDRA  $T1$  advertises to IDRA  $S$  that the ENAW of type  $\lambda_1$  to reach  $OXC D1$  from  $OXC T1$  is  $E_{T1, D1}^{adv}(\lambda_1) = \min_{\lambda_1} \{E_{T1, T2}, E_{T2, D2}, E_{D2, D1}^{adv}\} = \min_{\lambda_1} \{6, 2, 3\} = 2$ .

**Aggregated Load Information** — In our



■ **Figure 2.** a) Computation of the ENAW; b) advantage of the cost computation.

In OBG, the optical information is encoded and advertised using multiprotocol BGP extensions and extended communities. The OBG routing process on the other hand, is basically the same used in BGP, so OBG would generally choose the lightpath that traverses the least number of ASs.

model a cost is associated with each candidate (path, wavelength) pair. For simplicity, we shall only outline here the motivations for its use, and succinctly describe its computation and advertisement; for a detailed formulation we refer the reader to [8].

The goal is that the cost reflects the current load in the availability of wavelengths in an interdomain path, allowing an IDRA to compare routes more accurately than directly using the ENAWs between the source and destination. To this end, the cost computed and advertised between the IDRAs increases when the ENAW along the different segments of an interdomain path decreases. Likewise, the cost increases when the length of an interdomain path increases, where the length is computed as the number of hops  $H$  from the source OXC to the destination OXC, considering each intradomain subpath as a single hop. The source IDRA will choose the (path, wavelength) pair with the minimum cost.

It is worth highlighting that different paths offering the same ENAW will frequently have different costs (loads). For instance, in Fig. 2b OXC S can reach OXC D through both RCD T and RCD T'. The ENAW of type  $\lambda_1$  through RCD T is  $E_{S,D}(\lambda_1) = 1$ , and this is also the case for  $\lambda_2$  through RCD T' (i.e.,  $E_{S,D}(\lambda_2) = 1$ ). By computing the costs, the IDRA in domain S can discriminate between these two paths and select the one through RCD T, given that  $\lambda_1$  is less loaded than  $\lambda_2$  (notice that  $H = 5$  for both paths). It is worth highlighting that in the cases where the difference in the availability of wavelengths between two candidate paths is not as

obvious as in Fig. 2b, our evaluations confirm that the proposed approach offers a reasonable trade-off between the length  $H$  and the wavelength load on the lightpaths chosen.

In summary, the PSI advertised by the IDRAs consists of a set of candidate (loose) paths, together with their ENAWs and costs.

## RWA STRATEGY

In OBG the optical information is encoded and advertised using multiprotocol BGP extensions and extended communities. The OBG routing process, on the other hand, is basically the same used in BGP, so OBG would generally choose the lightpath that traverses the least number of ASs. Similar to BGP, OBG exchanges NRI, but it does not handle PSI. Understanding that this is a missing piece in the routing models provided by BGP/OBG is easy nowadays, but contributing solutions capable of highly improving the performance of these routing protocols without increasing the number and frequency of the routing messages exchanged between domains is a challenging task.

In this article we propose a simple method to cope with this problem. As in the case of BGP, the IDRAs exchange *Keepalive* messages to confirm that the processing modules (i.e., the electrical and software modules) of neighboring IDRAs are still operational. It is worth recalling that neighboring IDRAs are physically linked to each other, so failures at the optical layer can be detected and repaired by much faster means than the exchange of *Keepalive* messages. In

BGP, Keepalive messages are of fix length, consisting only of the 19-byte BGP header. In our model we extend the Keepalive concept with the purpose of piggybacking PSI only when relevant PSI needs to be updated.

A simplified version of the lightpath selection process is shown in Fig. 3 (for simplicity we only show the selection of a single route). Figure 3 shows that the IDRA choose minimum-cost paths (step 1), and if more than one path shows the same (minimum) cost, the IDRA break the tie first by the highest ENAW along the candidate paths, then by the shortest number of hops  $H$ , and after that by following essentially the same steps as BGP.

## PERFORMANCE EVALUATION

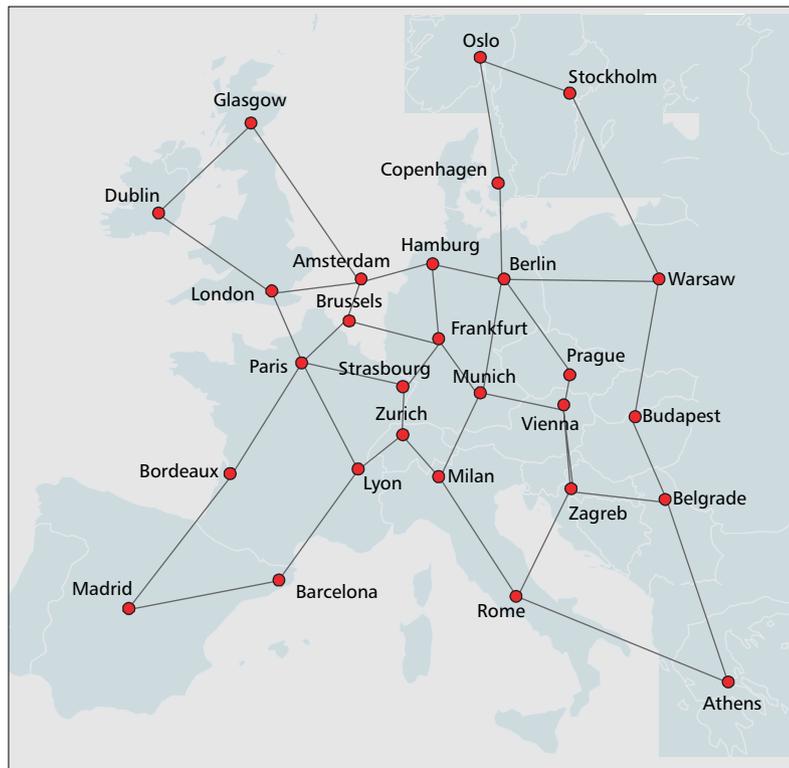
The aim of this section is to compare the performance of the IDRA-based route control model against OBG. Our interest is to examine the

**Input:** NRI associated with each destination  $d$   
PSI between OXCs  $s$  and  $d$

**Output:** The best (path, wavelength) pair between  $s$  and  $d$

- 1: Choose the (path, wavelength) pair with the minimum cost
- 2: If the costs are equal choose the path with the highest ENAW
- 3: If the ENAWs are equal choose the path with the shortest number of hops  $H$ , and assign the wavelength  $\lambda_i$  with the lowest identifier  $i$
- 4: If the hops  $H$  are equal prefer the path with the highest ENAW to the remote border OXC
- 5: If more than one path is still available run BGP tie-breaking rules [4]

■ **Figure 3.** IDRA RWA decision process.



■ **Figure 4.** Pan-European reference network topology.

blocking ratio (BR) of interdomain lightpath requests and the number of routing messages exchanged to achieve such blocking. To this end, we have conducted extensive simulations using OPNET.

The scenario chosen for the trials is illustrated in Fig. 4. This network was introduced in [9] as a reference topology suited for a pan-European fiber optic network. It is composed of 28 domains and 41 interdomain links, and the nodes were chosen in [9] to include some of the main European Internet exchange points. During the last years, this sample topology has become increasingly used as a reference simulation scenario. For the topologies inside each domain, we have randomly generated 10 different scenarios. For each of the 10 random scenarios, we have in turn chosen 10 different configurations by randomly placing 18 sources and 10 destinations, covering the pan-European network with a source or a destination inside each domain. This yields a total of 100 different settings for our tests. The results shown here are the averages over those 100 settings for both the BR and the number of routing messages exchanged throughout the simulation runtime.

We 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 ( $K_T$ ). BGP nodes usually use a default Keepalive value of 60 s, and three consecutive Keepalive messages need to be lost so that a node proceeds to shut down a BGP session. We have tested three scaled and normalized values:  $K_T = 1$ ,  $K_T = 3$ , and  $K_T = 5$  units through the simulation runtime. Clearly, the higher the values of  $K_T$ , the more time is needed by the IDRA to detect and react when the electrical or software modules of a neighboring IDRA become inoperative. Therefore, a major advantage of conveying PSI piggybacked on Keepalive messages is that low values of  $K_T$  are desired both to increase the responsiveness between neighboring IDRA as well as to support updating PSI more frequently.

The results shown here were obtained using cross traffic between all the sources and destinations in the pan-European network. Traffic was modeled according to a Poisson distribution with exponentially distributed arrival and departure rates, and, as shown in Fig. 5, the trials were performed for different traffic loads, varying from 100 up to 300 Erlangs.

In order to contrast the performance of our IDRA-based route control model against OBG under comparable conditions, we made the following decisions:

**Decision 1:** For the simulation results shown here, the multipath routing capabilities of the IDRA were switched off.

**Decision 2:** In OBG a node is not aware of the subset of wavelengths  $W(P)$  that are no longer available along the different segments of a path  $P$ . A node only receives a reachability message indicating the withdrawal of path  $P$  when *all* the candidate wavelengths in  $P$  have been consumed. A straightforward way to significantly reduce the BR experienced by OBG is to update

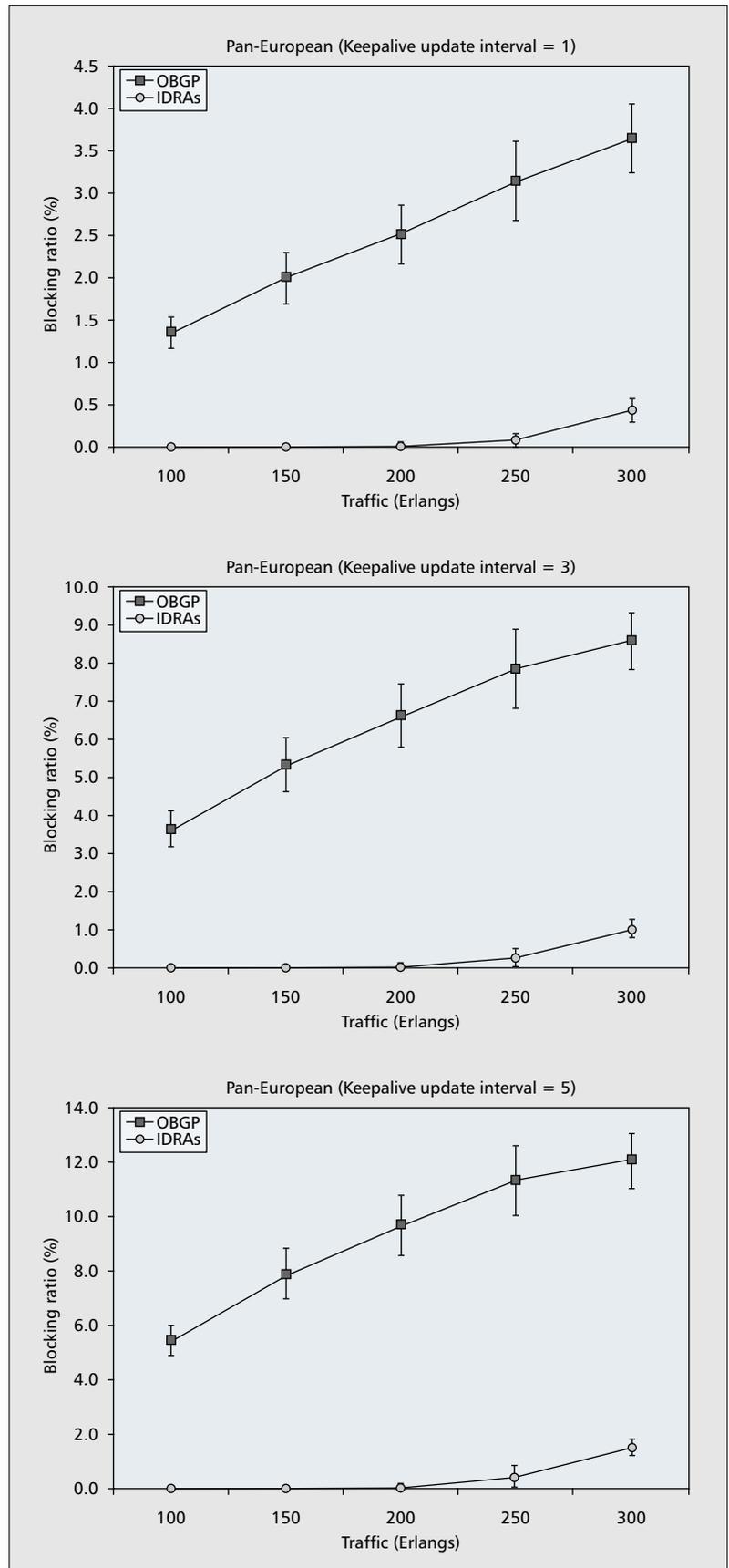
the subset  $W(P)$  through the Keepalive messages exchanged between OBGP neighbors. This provides more granular and updated NRI at the source OBGP node. Our implementation of OBGP follows this approach. We shall show that without this strategy, the BR experienced by OBGP would be much higher than that shown in Fig. 5.

Figure 5 shows the BR and standard deviation for the different traffic loads and different Keepalive update intervals assessed. Clearly, the IDRAs are able to drastically reduce the blocking obtained with OBGP. Whereas OBGP experiences blocking for all traffic loads tested, the IDRAs only start to show some negligible blocking after reaching 200 Erlangs. In order to quantify the gain in terms of blocking, we define the improvement factor (IF) as the ratio between the BR obtained with OBGP and the BR obtained with the IDRAs for the same traffic load. Table 1 summarizes the IF for 200, 250, and 300 Erlangs, as well as the number of routing messages exchanged for the different traffic loads and values of  $K_T$  tested. The results show that OBGP yields an overall BR between 7.93 and 363.97 times larger than that obtained with the IDRAs, depending on the traffic load and update interval  $K_T$ . Even in the case of the highest traffic load simulated, 300 Erlangs, the BR obtained with OBGP is approximately one order of magnitude larger than that experienced by the IDRAs.

The purpose of Fig. 5 is to reflect the contrast between two different route control strategies, so the values of the BRs per se should not be taken as representative of those expected in operational networks. More specifically, recent studies like the one developed in the framework of IST project NOBEL [10] have benchmarked the blocking performance for different applications. For example, [10] recommends that for real-time and streaming applications, the blocking should be less than or equal to 0.1 percent. Our results show that OBGP is unable to reach this bound for all the simulation conditions tested. To reach such a bound, extra resources (e.g., more wavelengths and/or fibers) would need to be added to the pan-European network in the case of OBGP. On the other hand, the IDRAs are able to reach the 0.1 percent bound for 200 Erlangs for the three values of  $K_T$ , and even for 250 Erlangs when  $K_T = 1$ . In particular, for 200 Erlangs and  $K_T = 1$ , the IF in the BR offered by the IDRAs is 363.97, and Table 1 shows that this can be achieved with approximately a third of the routing messages needed in the case of OBGP.

Figure 5 also confirms what we anticipated while describing our implementation decision 2. It is clear that the performance of OBGP degrades as  $K_T$  increases. Thus, when the subset  $W(P)$  is not conveyed and updated through the Keepalive messages, the BR yield by OBGP becomes rather independent of  $K_T$ , but the results obtained are worse than those shown for  $K_T = 5$  in Fig. 5.

A fundamental aspect is that for the conditions described here, our route control model always needs a smaller overall number of routing messages than OBGP (Table 1). The reason for



■ Figure 5. Comparison between OBGP and the IDRAs for different traffic loads, and Keepalive Update Intervals ( $K_T$ ).

	Keepalive update interval ( $K_T = 1$ )			Keepalive update interval ( $K_T = 3$ )			Keepalive update interval ( $K_T = 5$ )		
	200 Erlangs	250 Erlangs	300 Erlangs	200 Erlangs	250 Erlangs	300 Erlangs	200 Erlangs	250 Erlangs	300 Erlangs
<b>IF</b>	363.97	39.48	8.40	315.69	29.63	8.41	158.00	24.20	7.93
<b>Traffic (Erlangs)</b>	Routing messages OBGP	Routing messages IDRAs	Routing messages OBGP	Routing messages IDRAs	Routing messages OBGP	Routing messages IDRAs	Routing messages OBGP	Routing messages IDRAs	Routing messages IDRAs
100	6,564,525	2,819,949	5,539,285	2,771,408	4,842,449	2,764,530			
150	7,907,963	3,013,904	6,544,983	2,961,622	5,574,075	2,876,943			
200	8,607,917	3,141,911	6,905,969	3,041,394	5,822,980	2,946,896			
250	8,992,258	3,288,572	7,033,482	3,149,322	5,864,259	3,027,520			
300	9,198,274	3,661,793	7,071,856	3,393,776	5,928,454	3,179,430			

**Table 1.** Improvement factor in the blocking requests for 200, 250, and 300 Erlangs, and overall number of routing messages exchanged.

this is threefold. First, PSI updates are never triggered between neighboring IDRAs, but are rather piggybacked on the Keepalive messages used by both OBGP and the IDRAs. Second, whereas in OBGP the interdomain routing messages can potentially reach all the nodes in the network, in our route control model these messages only need to reach the IDRAs, which represent a small fraction of the total number of OBGP nodes in the network. Third, 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, the IDRAs explicitly consider the ENAW and cost in the RWA algorithm, so they are able to provide much better traffic distribution than OBGP. This produces drastic reductions in the BR; hence, fewer network reachability messages need be exchanged.

It is worth emphasizing that for the multidomain network tested here, the global overhead in the size of the Keepalive messages exchanged between the IDRAs is negligible.

## CONCLUSIONS

In this article we have argued in favor of the opportunity to make a change offered by future optical networking, and as a first step in this direction we have proposed and tested a route control model that is an alternative to BGP/OBGP. One of the major advantages of our model is its simplicity. It is built on well-known techniques in networking; it reuses some of the strengths of BGP/OBGP, and effortlessly integrates highly aggregated PSI in the form of two integer values.

Despite its simplicity, the proposed route control model is able to drastically reduce the blocking obtained with OBGP, and such improvements can be achieved without needing to exchange more routing messages than with OBGP. In fact, our strategy reduces the number of routing messages exchanged between domains, since by decrementing the blocking, it

is possible to reduce the exchange of network reachability messages and path explorations when blocking occurs.

These are promising findings, but more research is needed in this direction. First, our results and conclusions apply to a rather small multidomain optical scenario (the pan-European reference network shown in Fig. 4), so further studies are needed to analyze the performance of the proposals made here in a large-scale environment. Second, our results were obtained under the wavelength continuity constraint, so it is necessary to analyze the potential impact of wavelength conversion, particularly when the conversion is performed at domain boundaries. Third, standardization efforts will be needed to appropriately incorporate the exchange of NRI and PSI at the I-NNI and E-NNI interfaces.

## ACKNOWLEDGMENTS

This work was partially funded by both the Spanish Ministry of Science and Technology under contract TEC2005-08051-C03-01 and the Catalan Government under contract 2005 SGR00481 for UPC authors. The authors also acknowledge the support received from OPNET Technologies Inc.

## REFERENCES

- [1] M. Blanchet, F. Parent, and B. St-Arnaud, "Optical BGP (OBGP): InterAS Lightpath Provisioning," IETF draft, ietf-draft-parent-obgp-01, Mar. 2001.
- [2] M. J. Francisco *et al.*, "End-To-End Signaling and Routing for Optical IP Networks," *Proc. IEEE ICC*, New York, NY, Apr. 2002.
- [3] L. Wang, H. Zhang, and L. Zheng, "Reducing the OBG Protection Switching Time in WDM Mesh Networks," *Proc. OFC*, Anaheim, CA, Mar. 2006.
- [4] Y. Rekhter, T. Li, and S. Hares, "A Border Gateway Protocol 4 (BGP-4)," IETF RFC 4271, Jan. 2006.
- [5] G. Liu, C. Ji, and V. Chan, "On the Scalability of Network Management Information for Inter-Domain Light-Path Assessment," *IEEE/ACM Trans. Networking*, vol. 13, no. 1, Feb. 2005.
- [6] M. Yannuzzi *et al.*, "Interdomain RWA based on Stochastic Estimation Methods and Adaptive Filtering for Optical Networks," *Proc. GLOBECOM '06*, San Francisco, CA, Nov./Dec. 2006.

- [7] A. Farrel, J. P. Vasseur, and J. Ash, "A Path Computation Element (PCE)-Based Architecture," IETF RFC 4655, Aug. 2006.
- [8] M. Yannuzzi, *Strategies for Internet Route Control: Past, Present, and Future*, Doctoral thesis, Tech. Univ. of Catalonia, Spain, 2007.
- [9] S. De Maesschalck et al, "Pan-European Optical Transport Networks: An Availability-Based Comparison," *Photonic Network Commun.*, Springer, vol. 5, no. 3, May 2003, pp. 203–25.
- [10] NOBEL: Next Generation Optical Networks for Broadband European Leadership, IST Integrated Project (FP6-506760), <http://www.ist-nobel.org/>

## BIOGRAPHIES

MARCELO YANNUZZI ([yannuzzi@ac.upc.edu](mailto:yannuzzi@ac.upc.edu)) received a degree in electrical engineering from the University of the Republic (UdelaR), Uruguay, in 2001, and DEA (M.Sc.) and Ph.D. degrees in computer science from the Department of Computer Architecture, Technical University of Catalonia (UPC), Spain, in 2005 and 2007, respectively. He is with the Advanced Network Architectures Lab at UPC, where he is an assistant professor. He held previous positions with the Physics Department of the School of Engineering, UdelaR, from 1997 to 2003, and with the Electrical Engineering Department of the same university from 2003 until 2006. He worked in the industry for 10 years at the national telco in Uruguay (1993–2003).

XAVI MASIP-BRUI ([xmasip@ac.upc.edu](mailto:xmasip@ac.upc.edu)) received M.S. and Ph.D. degrees from UPC, both in telecommunications engineering, in 1997 and 2003, respectively. He is currently an associate professor of computer science at UPC. His current research interests lie in broadband communications, QoS management and provision, and traffic engineering. His publications include around 50 papers in national and international refereed journals and conferences. Since 2000 he has participated in many research projects: IST projects E-NEXT, NOBEL, and EuQoS; and Spanish research projects SABA, SABA2, SAM, and TRIPODE.

GUILLEM FABREGO-GINEBRA ([gfab9618@alu-etsetb.upc.edu](mailto:gfab9618@alu-etsetb.upc.edu)) received a degree in telecommunications engineering from UPC in 2008. His research interests are in the areas of

interdomain routing and network stability in optical networks.

SERGIO SANCHEZ-LOPEZ ([sergio@ac.upc.edu](mailto:sergio@ac.upc.edu)) got his Ph.D. in telecommunications engineering in 2003 from the Department of Computer Architecture, UPC. Since 1992 he has been an associate professor with this department. His publications include several book chapters, papers in relevant international journals, and more than 50 papers in international refereed conferences. His current research interests are in broadband Internet, and high-speed and optical networks, with emphasis on traffic engineering and control plane.

ALEX SPRINTSON ([spalex@tamu.edu](mailto:spalex@tamu.edu)) is an assistant professor of electrical and computer engineering at Texas A&M University. He received his Ph.D. degree in electrical engineering from the Technion — Israel Institute of Technology, Haifa, in 2003. His research interests lie in the general area of communication networks with a focus on network survivability, QoS routing, network coding, and security of wireless communication networks. He served as a member of the Technical Program Committee for IEEE INFOCOM 2006, 2007, and 2008.

ARIEL ORDA [S'84, M'92, SM'97, F'06] ([ariel@ee.technion.ac.il](mailto:ariel@ee.technion.ac.il)) received B.Sc. (summa cum laude), M.Sc., and D.Sc. degrees in electrical engineering from the Technion — Israel Institute of Technology in 1983, 1985, and 1991, respectively. Since 1994 he has been with the Department of Electrical Engineering at the Technion, where he is the Herman and Gertrude Gross Professor of Communications. His research interests include network routing, survivability, QoS provisioning, wireless networks, the application of game theory to computer networking, and network pricing. He received the Award of the Chief Scientist in the Ministry of Communication in Israel, a Gutwirth Award for Outstanding Distinction, the Research Award of the Association of Computer and Electronic Industries in Israel, the Jacknow Award for Excellence in Teaching, an ICNP '04 Best Paper Award, the Journal of Computer Networks 2004 Best Editor Award, and the 2005 Klein Award for Excellence in Research. He has served as program co-chair of IEEE INFOCOM, and Editor of *IEEE/ACM Transactions on Networking* and the *Journal of Computer Networks*. In 2008 he is the incumbent of the Cor Wit Chair at Delft University of Technology in the Netherlands.