Optical Switching and Networking 19 (2016) 10–21

Contents lists available at ScienceDirect

Optical Switching and Networking

journal homepage: www.elsevier.com/locate/osn

Schemes for inter-domain lightpath establishment based on PCE architecture

Alisson S.L. Pontes [a], Nelson L.S. da Fonseca [a,*], André C. Drummond [b]

[a] Institute of Computing, University of Campinas, Brazil
[b] Department of Computer Science, University of Brasília, Brazil

ARTICLE INFO

Article history:
Received 20 October 2014
Received in revised form
29 July 2015
Accepted 17 September 2015
Available online 3 October 2015

Keywords:
Inter-domain network
PCE
GMPLS control plane
Routing and wavelength assignment

ABSTRACT

Despite the research advances in intra-domain lightpath provisioning in Wavelength Division Multiplexing networks, efficient and practical schemes for path computation and resource advertisement in multi-domain mesh networks still need to be developed. Most of the solutions proposed in the literature lacks the ability to convey optical network-specific Traffic Engineering information and are based on a periodic message flooding technique. This paper proposes three novel solutions for inter-domain lightpath provisioning in WDM circuit switched mesh networks. Two routing advertisement schemes and two path and wavelength selection criteria for the PCE architecture are proposed. The solutions provide a complete inter-domain routing and wavelength solution. Simulation experiments show the effectiveness of the proposed solutions which significantly reduce the total amount of message exchanged as well as overall call blocking.

1. Introduction

The growing need for traffic engineering in backbone networks has led to the proposition of the PCE architecture, which was standardized by the Internet Engineering Task Force (IETF) [1]. This architecture does not provide a detailed description of all the architectural components, but rather describes a set of building blocks [2] and can be viewed as a first step towards the implementation of a constraint-based multi-domain path computation (traffic engineering). Recent community efforts in open source PCE have enabled innovation in those building blocks precisely that are relevant to a specific PCE application within a network [3], although solutions need to be developed on the top of this architecture for the development of future optical backbone control planes. For a comprehensive survey on the benefits of the PCE architecture refer to [4]. The evaluation of real PCE testbeds have been recently carried out and presented in [5,6].

Indeed, the IETF has specified three approaches for the computation of multi-domain paths: per-domain path computation [7], Backward Recursive PCE-based Computation (BRPC) [8] and Hierarchical Path-Computation Element (H-PCE) [9].

In the per-domain approach, methods for path computation are usually defined based on an auto-discovery mechanism. The complete path is obtained by concatenating segments computed for each domain. However, one of the major drawbacks of this approach is the inability to exploit multiple exit/entry points and the sub-optimal nature of the process which is based generally on outdated information stored in routing information data bases. One of the main issues in path computation is therefore the employment of effective resource advertisement protocols.

* Corresponding author.
E-mail addresses: alisson@lrc.ic.unicamp.br (A.S.L. Pontes), nfonseca@ic.unicamp.br (N.L.S. da Fonseca), andred@unb.br (A.C. Drummond).

http://dx.doi.org/10.1016/j.osn.2015.09.003
1573-4277/© 2015 Published by Elsevier B.V.

In BRPC, the source PCE can specify the sequence of domains to be traversed by using the path computation PCE Communication Protocol (PCEP) [10] to compute the optimal path across the specific sequence of domains. This sequence is either administratively predetermined or discovered by forwarding the request to its neighboring PCE. However, the way to determine it has been left open in the standard [8].

In H-PCE a parent PCE maintains a domain topology map that contains the child domains (domain's PCE) and their interconnections, allowing it to find paths for inter-domain connections.

None of the aforementioned schemes for multi-domain path computation supports wavelength continuity along the path due to lack of information about resources availability. As a result, multi-domain routing schemes do not have all the information needed for the establishment of multi-domain lightpaths. To this end, much effort still needs to be made to have a complete inter-domain service provisioning solution.

To address these mentioned issues, this paper proposes different information advertising and lightpath establishment schemes which have yielded to three novel PCE-based solutions for inter-domain lightpath provisioning in Wavelength Division Multiplexing mesh networks. Our proposals provide new ways of computing the chain of domains and introduce policies that account for the availability of wavelengths. In this complete Routing and Wavelength Assignment (RWA) solution, on-demand advertising schemes and the path computation can maintain confidentiality of intra-domain information. The three proposals differ in relation to the routing advertisement scheme, as well as the criteria for path and wavelength selection to establish a lightpath. Two proposals employ a common advertising scheme and two adopt the same path computation scheme.

Simulation experiments show the effectiveness of the proposed solutions which significantly reduce the total amount of message exchanged as well as overall call blocking.

The remainder of the paper is organized as follows. Section 2 gives an overview of the issues related to path computation and advertisement of resource availability. Section 3 presents our proposed mechanisms for both. Section 4 evaluates the performance of the proposed schemes and conclusions are drawn in Section 5.

2. Related work

Inter-domain routing has been the focus of recent research [11] which spans from IP to optical networking and includes aspects such as physical impairment [12,13] and Quality of Service among others.

Some researchers have analyzed the possibility of adopting the Optical BGP (OBGP) solution [14–16] as the future inter-domain routing protocol for optical networks. The aim of these proposals is to extend BGP so that it can convey and signal optical information between OBGP neighbors. BGP provides means for each domain to obtain and propagate reachability information from neighboring domains and to define routes to other domains. In order to distribute reachability information, different domain border routers disseminate the addresses of all the routers they can reach. When a border router receives each information, it passes the information on to all the routers in its domain. With the information obtained from other BGP routers, each router can define routes to routers in other domains. Francisco et al. [15] provided the first implementation of OBGP (as far as we know), specifying requirements and the necessary extensions of BGP to create OBGP. However, multidomain routing models centered on the exchange of network reachability information are not sufficient for wavelength switched optical networks, it is now widely accepted that neighboring domains should also be able to exchange resource availability information. Indeed, OBGP inherits the main characteristics of BGP [17] as well as its well known problems such as the inability to convey useful Traffic Engineering (TE) information, slow convergence and chattiness [18].

More recently a new mechanism was proposed to collect link state and traffic engineering information and share these informations with external components using the BGP routing protocol. This is achieved using BGP-LS [19], a new BGP Network Layer Reachability Information encoding format. BGP-LS was developed to gather information about the topologies and capabilities of the network in order to fulfill the functions of the PCE architecture when performing inter-domain path computation. BGP-LS enhances the BGP protocol while defining a new BGP NLRI that describes links, nodes and prefixes comprising intra-domain link state information, and while defining a new BGP path attribute (BGP-LS attribute) that carries link, node and prefix properties and attributes, such as the link and prefix metric or auxiliary Router-IDs of nodes.

Casellas et al. [20] extend the BRPC algorithm to handle end-to-end wavelength continuity constraints. In standard BRPC, assuming that BRPC entities know the domain sequence in advance, the algorithm computes the inter-domain path in a reverse way, starting with the destination domain. The destination domain PCE computes a virtual shortest path tree (VSPT). It selects the optimal path from each of the ingress nodes to the destination node, pruning the suboptimal paths from the VSPT before sending it to its own upstream domain PCE. The PCE leaves the wavelength assignment (WA) process to the resource reservation phase and performs routing based on shortest paths and Traffic Engineering (TE) metrics. Casellas et al. propose an extension of BRPC so that the PCE chain could perform both routing and WA. They implemented an intra-domain path computation algorithm that computes the path, the total traffic engineering metric from the ingress to the egress domain node, as well as the set of candidate wavelengths considering the wavelength status of the link. The inter-domain path computation involves the computation of the virtual shortest path tree of the BRPC. Although Casellas' paper claims to compute the end-to-end path performing wavelength assignment, the extension proposed in [20] does not present a solution to the computation of the chain of domains, thus wavelength availability is considered only during the virtual shortest path tree computation procedure. Moreover, the proposal does not provide details about the inter-domain advertising of optical resource information.

Yannuzzi et al. [18,21] reinforce the idea that the inability to exchange aggregated path-state information is a problem of the current routing models for wavelength switched optical networks, although they propose solutions capable of greatly improving the performance of path-vectors without impacting on key aspects of the protocol (i.e. scalability, convergence properties, and number of routing messages exchanged between domains). The authors proposed an extension of a path-vector protocol supporting the computation and advertisement of path state information between optical domains [21]. However, the solution is still a shortest path algorithm that considers mainly information about reachability to compute inter-domain paths. In [18], Yannuzzi et al. consider a network architecture in which independent circuits physically connect the nodes within the control plane. These nodes, called inter-domain routing agents (IDRA), are similar to PCEs. The authors propose a cost model that reflects the current load in the availability of wavelengths on an inter-domain path, allowing an IDRA to compare routes more accurately. The cost computed and advertised between the IDRAs increases with the load increase as well as with the length of an inter-domain path. This strategy computes the effective number of available wavelengths in the network and this metric is used for path computation. However, the computation of the available wavelengths depends on a slow convergence procedure to achieve scalability.

Another aspect which is not considered by Yannuzzi proposals is the triggering events for updating resource availability, which is critical to the signaling load experienced by the network. In the proposal in [22], the authors introduce a mechanism for pre-reserving inter-domain resources and triggering updates when resource levels reach specific thresholds in order to reduce both inter-domain signaling overhead while and blocking. However, as the proposal is not evaluated in conjunction with a routing protocol the issue of increased signaling load due to triggering events on PCEs remains unsolved.

The work in [23] and [24] adopts a non-optical inter-domain PCE-based network scenario similar to the scenario adopted in this paper. Chen [24] proposes a PCE-based inter-domain path computation scheme for searching an optimal path. However, the path computation does not consider pre-determined domain chains, otherwise it relies in a path computation flooding which may cause more signaling overhead due to exchanging routing information with all possible neighboring domains [9]. Zhang et al. [23] proposed a shortest disjoint working and backup paths considering pre-determined chain of domains. However, the way to determine the chain was not defined. Actually the computation of the chain of domains is still considered an open issue in the standards [8], which motivated the contributions introduced in this paper.

Cugini et al. [25] focused on PCE architecture for flexible optical networks. Both the PCE architecture and PCE communication protocol try to maximize the spectral efficiency. Experimental results show the PCE capability to trigger dynamic rerouting with bit-rate or modulation format adaptation. Moreover, Casellas et al. [26] propose the design and implementation of a GMPLS/PCE control plane for flexible optical networks. The control plane uses a distance adaptive and PCE-based routing and modulation assignment, combined with distributed frequency slot (spectrum) selection. They show the benefits of the knowledge of the status of the slices and the spectral efficiency of the modulation formats to path computation functions.

Zhao et al. [27] present an extension of the BRPC framework to address the quality of transmission (QoT) of intra-domain and inter-domain connections in WDM optical networks. The paper proposes some cross-layer heuristics that properly allocates regenerators in the network to assure the signal quality of the lightpaths. Although this work uses BRPC to perform inter-domain routing, it does not take into consideration the amount of overhead generated by message exchanging on the design of the proposed solution.

Ahmed et al. in [28] emphasized the benefits of adopting the PCE architecture by proposing a dynamic provisioning framework for optimizing the use of network resources as well as the ability to reduce control overhead.

In recent work [29], the design of new inter-domain optical routing protocols is discussed. The paper emphasizes the pervasive deployment of BGP in the current Internet infrastructure, which leads to the current dependence on BGP-based protocols. Moreover, the work discusses the weakness of current BGP extensions for optical networks present in the literature and proposes a set of algorithms to address them, focusing on the QoT issue. The proposed solutions are compared with OBGP.

The work in [30] proposes a hierarchical instance protocol dedicated to provide the PCE with effective domains sequence information on a path-vector protocol (DSP-PCE). The proposed path computation scheme exploits additional attempts along different domain sequences, and the impact on the set-up time is evaluated. Results show that the proposed solution, compared to solutions based on BGP, significantly improves the overall blocking probability. Nevertheless, it cannot be employed to optical networks, due to lack of wavelength availability information.

Reference [9] examines techniques to establish paths when the sequence of domains is not known in advance. The authors focus on the hierarchical PCE architecture. However, this architecture is just applicable to environments with small groups of domains. Applying such hierarchical PCE model to large groups of domains is not considered feasible by the authors. Nevertheless, the H-PCE architecture has been largely studied in the literature. In [31], it is proposed a k random path (KRP)-based inter-domain routing algorithm. In this proposal, the algorithm achieves lower blocking probability, compared with traditional schemes, as the number k of random selected paths increases.

In [32], a hierarchical BGP protocol instance (HBGP-PCE) was proposed to provide the PCE architecture with additional information to be used in multi-domain path computations. The authors claim that the adoption of multiple attempts along different sequences of domains and the use of the BRPC procedure do not provide significant improvement in performance.

Lu et al. [33] propose a domain-level-based routing (DLR) algorithm for a multidomain WDM network with confidentiality constraints of interdomain connectivity and wavelength availability. Instead of determining the traversed domains just based on abstract interdomain information, the proposed algorithm determines the domain sequence based on the decision of multidomain routing results. The DLR algorithm can be implemented in a hierarchical PCE-based routing architecture. Compared with the KRP algorithm, the DLR algorithm achieves lower blocking probability when traffic load is heavy and significantly reduces the synchronization messages.

3. Inter-domain lightpath computation

The schemes proposed here for inter-domain service provisioning use Multiple PCE Path Computation with Inter-PCE Communication Architecture [2], in which at least one PCE per domain can perform inter-domain routing based on information stored in its Traffic Engineering Database (TED). This information is exchanged among the PCEs to compute multi-domain paths.

In these proposals, it is assumed that the PCEs run a path vector with path caching protocol [34], which implies that each PCE has in its TED one or more defined paths (domain chains) to all other domains in the network. Actually, there is, at most, a number of paths to a given destination equal to the number of edge nodes. This implies a certain flexibility in choosing among multiple paths, thus increasing reliability. Given a destination, each PCE participating in a path vector protocol chooses, at any given time, a local-optimal path with respect to the paths last learned from each of its neighbors to reach the destination. If there is more than one local-optimal path, the node deterministically chooses one of them.

Each TED entry contains a list of available wavelengths as well as the output border node linking it to a neighboring domain in the chain of domains (an output border node is a node which has a link to a neighboring domain). When a request for multi-domain lightpath establishment arrives at the source PCE, this PCE chooses from the routes available to reach the destination.

The messages exchanged in the information dissemination schemes include both reachability and resource availability information, which allows the selection of a wavelength for the establishment of an end-to-end lightpath that is being explored during the computation of the domain chain. A scheme called backtracking, similar to the crankback scheme in RSVP [35], resumes the path computation at the previous PCE in the domain chain when TED information leads to an infeasible path. When this happens, computation of a new path can avoid blocked resources. Backtrack events trigger the dissemination of messages for the exchange of reachability and resource availability information.

When PCE p1 receives a signaling routing message from its neighbor p2, it updates its chosen path to the destination domain to become the most preferred of the local-optimal paths with respect to the paths to all other domains, if the chosen path has changed as a result of the update. Similar procedure exists to deal with the failure or addition of a link, or to deal with change of wavelength availability. We assume that for each pair of PCEs p1 and p2, such that p1 is a neighbor of p2, there is a signaling queue to hold the signaling routing messages in transit from p1 to p2. This signaling queue is lossless and behaves according to a first-in-first-out service discipline.

The end-to-end lightpath establishment assumption considered in this work imposes the wavelength continuity constraint to the inter-domain connections. Although in Tier-1 transport networks, this scenario may not be considered realistic due to optical signal regeneration in long fiber spans, the wavelength conversion at domain boundaries is not mandatory. Thus, this assumption is valid and leads to the generalization of the solution.

In terms of confidentiality of the information exchanged in multi-carrier networks, the end-to-end lightpath establishment assumption imposes a hard constraint. In order to make wavelength continuity possible, one of two approaches have to be employed: (i) the use of a hierarchical solutions, such as H-PCE; (ii) the use of distributed solutions, such as our proposal, which requires the wavelength availability to be advertised across domains. This trade-off leads to solutions that may compromise confidentiality, but have the potential to be very lightweight in terms of message overhead. Nevertheless, the proposed solution does not advertise any sensitive intra-domain information. Even the wavelengths availability that is transported in inter-domain signaling messages may be only a subset of the whole amount of wavelengths which is effective available internally.

The three proposals presented here differ in relation to the reachability of the informations dissemination message, as well as to the criteria for the choice of the domain chain.

3.1. Source-driven single PCE backtracking (SDSB) scheme

In this proposal, wavelength selection is carried out by the source PCE and domain chain computation and end-to-end path computation is carried in cooperation by the PCEs. Indeed, the backtracking events trigger the dissemination of routing information to a single PCE.

The criterion for domain chain computation determines the choice of the path with the greatest number of available wavelengths. Then, if there is more than one option, the path with the shortest domain chain will be chosen (Algorithm 2). After defining the domain chain, the source PCE randomly selects one wavelength among these available for the selected path. The source PCE signals the local network to allocate the resources necessary to support the call and forwards another PCEP request to the next PCE along the domain chain (lines 1 to 4 of Algorithm 1). The neighbor PCE which received the PCEP request continues the domain chain computation and allocate resources to support the call considering the wavelength chosen by the source PCE and the chosen output border node of the previous domain. The procedure continues until the PCE of the destination domain is reached and an end-to-end lightpath has been established. In the final step, resource reservations are setup by a protocol such as RSVP [36] (lines 12 to 16 of Algorithm 1). The procedure just described is formalized in Algorithm 1.

Algorithm 1. Inter-domain lightpath computation.

Require: Each PCE runs a path vector algorithm with path caching.
Require: Each TED path entry has a list of available wavelengths attached.
Ensure: A multi-domain end-to-end lightpath.

```
1:  Receipt of a call request by a PCE to establish an end-to-end
    lightpath. To select the neighbor to which the PCEP request
    message will be forwarded based on information in TED;
2:  Uses the path selection policy defined by Algorithm 2;
3:  Randomly pick one wavelength from the list of available
    wavelengths of the chosen path;
4:  Reserve resources necessary to support the call from the
    source node to the chosen output border node;
5:  repeat
6:    Receipt of a message by PCE;
7:    if this message is a backtrack message then
8:      Update its TED;
9:    The PCE refers to its TED to select the neighbor to which it
    will forward the PCEP request message;
10:   if There is no path to the destination then
11:     Send back a backtrack message piggybacked on a state
    update message;
12:   Use path selection policy defined by Algorithm 2;
13:   Reserve resources necessary to support the call from the
    input border node to the chosen output border node;
14: until The PCEP request reaches the destination domain OR the
    limit of backtrack events has been reached;
15: if The PCEP request reaches the destination domain then
16:   perform end-to-end path computation and resource reser-
    vation setup;
17: else
18:   The call is blocked;
19: de-allocate unused resources;
```

Algorithm 2. Path selection policy.

```
Require: Given a destination domain and all the possible paths
    (domain chains) to reach it.
Ensure: Output of a single path to the destination domain.
if There is more than one output neighbor domain then
  Choose the one that has a higher number of available wave-
  lengths in the path between it and the destination domain;
  if multiple output neighbor domains were returned then
    Choose the shortest path (fewest domain chain hops);
    if multiple output neighbor domains were returned then
      Choose the path on which the first domain have the
  smallest AS number;
```

The resource availability dissemination procedure used to update the information in the TEDs employees on-demand notification. When a PCEP request message arrives at a PCE and the target wavelength is no longer available in that domain, a backtrack event occurs and a message is sent back to the previous PCE (lines 10 and 11), which must choose an alternative neighbor PCE to try to establish a path. The backtrack message piggybacks an updating message which disseminates the changes in resource availability. An updating message is triggered only by a backtrack event and reaches only the previous domain, which allows the choice of routes based on updated information without needing to flood the entire network with information (lines 7 and 8 of Algorithm 1). In this way, the overhead of periodic, and sometimes unnecessary updates, is avoided. Fig. 1 represents the flow of PCEP request and updating messages when a back-tracking event occurs in the network as shown in Fig. 2. Another backtrack event can follow if no feasible route has been found and this continue until a pre-defined limit backtracks number has been reached. Once this limit is reached, the request is blocked (line 17 and 18 of Algorithm 1).

The flow of control messages can be better understood by means of an example. Let us consider the network shown in Fig. 2, which is composed of five domains, each with a PCE. *PCE*2 is aware of intra and inter-domain links of domain *D*2 and thus knows about the availability of resources for reaching *D*1. Update messages received by *PCE*3 from *PCE*2 are analyzed and stored in the TED of *PCE*3, which now has updated information to reach *D*1. These messages contain information on: (i) the destination domain (*D*1), (ii) the input border nodes from *D*2 and (iii) the wavelengths available between the input border node of *D*2 and the input border node of the destination domain (*D*1) (an input border node is a node which has a link to previous domain in the chain). In the same way, *PCE*3 can furnish updating information to the TED of *PCE*4, thus extending the path leading to *D*1 and updating on available wavelengths from the intersection of wavelengths available on the path *D*1–*D*3 and *D*3–*D*4. The TED of *PCE*5 can be updated in the same way, with information coming from *PCE*3 and from *PCE*4.

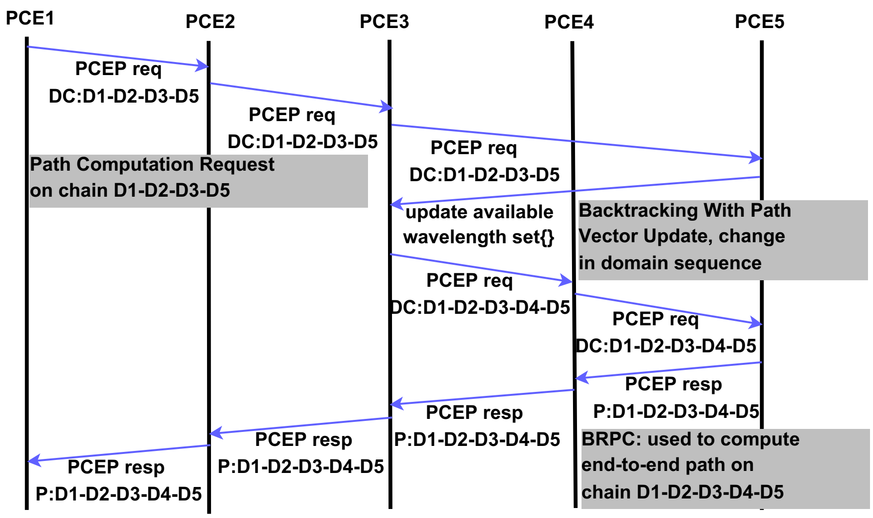

Fig. 1. Flow of PCEP request and update signaling.

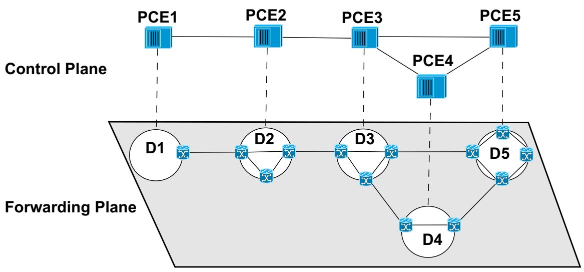

Fig. 2. PCE-based multi-domain scenario.

At the end of this process, *PCE*5 knows the topology (Fig. 3) and the available resources to reach *D*1, which includes two possible paths to reach it, one via *D*3 (shown as a dashed line) and other via *D*4 (shown as a continuous line). Moreover, it has no information about the internals of any other domain, which guarantees confidentiality, an asset for providers of a competitive businesses.

The proposed approach provides a lightweight solution by combining path computation and resource advertisements. As a result, the approach resolves some of the main limitations of the BGP-based solution, such as: (i) the lack of ability to convey useful traffic engineering information, achieved by the use of the PCE architecture; (ii) lack of multiple routes, made possible by the path caching scheme which allows the source PCE to pick more than one path if desirable; and (iii) slow convergence and chattiness, addressed by the backtrack messages which carries update information triggered only if necessary during the establishment of an end-to-end lightpath.

3.2. Destination-driven single PCE backtracking (DDSB) scheme

In the previous proposal, wavelength assignment is performed by the source PCE, which uses information contained in its TED. The selected wavelength can however no longer be available along the domain chain which leads to the blocking of the request.

A solution is proposed in this section to consider effective wavelength availability along the domain chain by using information available during the path computation procedure. In the proposed destination-driven wavelength assignment scheme, the wavelength assignment is not in charge exclusively of the source PCE anymore. It is performed cooperatively by all the PCEs in the domain chain. The wavelength is assigned in the destination PCE, according to the wavelengths effectively available during the domain chain computation. Instead of selecting the domain chain with the greatest availability of wavelengths, a PCE selects the domain chain which will lead to the minimal reduction in the set of candidate wavelengths to reach the destination PCE. As in the SDSB scheme, the source PCE forwards a PCEP request to the next PCE along the domain chain, but in this current proposal the source PCE informs the next PCE about the candidate wavelengths available for the establishment of the lightpath. The neighbor PCE which received the PCEP request performs the same procedure. This procedure is repeated until the PCE of the destination domain is reached thus establishing an end-to-end lightpath (Algorithm 3).

Algorithm 3. Inter-domain lightpath computation.

Require: Each PCE runs a path vector algorithm with path caching.
Require: Each TED path entry has a list of available wavelengths attached.
Require: The initial value of candidate wavelengths array is empty.
Ensure: A multi-domain end-to-end lightpath.

```
repeat
    PCE receives a message;
    if It is a backtrack message then
        Update its TED;
    else
        The PCE refers to its TED to select to which neighbor it will
    forward the PCEP request message;
        if There is no path to the destination then
            Send back a backtrack message piggybacked with a state
    update message;
        Use path selection policy (Algorithm 4);
until The PCEP request reaches the destination domain OR n
    backtracks have accurred;
if The PCEP request reaches the destination domain then
    Randomly pick one wavelength from the array of candidate
    wavelengths;
    perform end-to-end path computation and resource reser-
    vation setup;
else
    The call is blocked;
```

Algorithm 4. Path selection policy.

Require: An array of candidate wavelengths.
Require: Given a destination domain and all the possible paths (domain chains) to reach it.
Ensure: Outputs a single path to the destination domain.

```
if multiple output neighbor domains were returned then
    Choose the one which minimally reduces the set of available
    wavelengths to the destination.
    if multiple output neighbor domains were returned then
        Choose the shortest path (fewest domain chain hops);
        if multiple output neighbor domains were returned then
            Choose the path on which the first domain have the
    smallest AS number;
```

When a PCEP request message arrives at a PCE and there is no wavelength in the set informed in the PCEP message, a backtrack event occurs and a message is sent back to the previous PCE, which must choose an alternative neighbor PCE to retry the procedure. The flow of PCEP requests and update signaling is the same as that in Fig. 1.

In Section 4, it will be shown that DDSB produces less blocking and overhead than the SDSB approach. This is mainly due to the flexibility obtained by leaving the wavelength selection to the end of the path establishment procedure. For this reason, the information dissemination scheme presented in the next section will be evaluated only jointly with the destination-driven approach for wavelength selection.

3.3. Destination-driven Multiple PCE Backtracking (DDMB) scheme

The proposal in this section disseminates updated information to the PCEs no further than *n* hops from the PCE at which the path computation failed. It is a hybrid scheme combining the usual convergence based routing

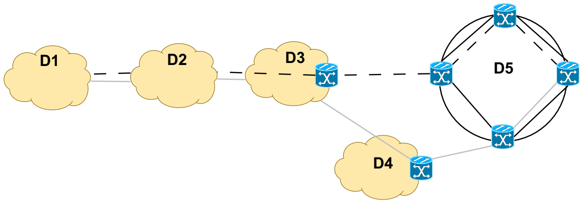

Fig. 3. Inter-domain view of Domain 5 PCE.

protocols with the neighbor-only advertising presented in SDSB and DDSB, which aims to seek a trade-off between routing overhead and accuracy of information on network resources availability. In this scheme, when a PCE receives a backtracking message it spams all PCEs in its neighborhood with Update Request messages. For a PCE *p* its *n*-neighborhood represents the set of all nodes within *n* hops of distance from *p* (not including *p* itself). An *n*-neighborhood is an updated zone of a PCE *p* if during the updating procedure *p* spams all PCEs within a distance of *n* hops, i.e., each PCE has accurate information about the nodes in its updated zone.

For instance, let us consider Fig. 4, which illustrates a network composed of 5 domains. Consider that a path computation request arrives at *PCE*1 to reach an Optical Cross Connect (OXC) of domain *D*5. The wavelength assignment and path computation procedures are the same as DDSB. When a PCEP request arrives at *PCE*1, it must first choose available routes to the destination. The first criterion in this selection determines that paths with the greatest number of available wavelengths should be chosen. If there is more than one option, then the path with the shortest domain chain should be chosen. Then, *PCE*1 forwards a PCEP request to *PCE*2, the next PCE along the domain chain, and informs it about the candidate wavelengths available for establishing the lightpath. *PCE*2 then performs the same procedure, considering the set of wavelengths available and the output border node of domain *D*1.

Now, consider that *PCE*2 forwards the call directly to *PCE*5, but *PCE*5 does not have sufficient resources to complete the request through $OXC_{5,1}$. So *PCE*5 sends back a backtracking message with an update message. When *PCE*2 receives the backtrack message from *PCE*5 it sends an update request message to all nodes within *n* hops, and receives update responses. The main difference between DDMB and DDSB is the resource advertisement procedure. In DDSB, *PCE*2 would compute an alternative path (through *D*3 or *D*4) based on the updated information received from *PCE*5 and the information contained in its TED about *D*3 or *D*4. In the DDMB approach, on the other hand, *PCE*2 can compute an alternative path based on information received from all PCEs in its *n*-neighborhood (except to those PCEs reached via *PCE*5). Overlapping updated zones will be created, resulting in reasonably consistent information between nodes belonging to adjacent zones.

Fig. 5 represents the flow of PCEP request/response during a backtracking event after a call request on the network in Fig. 4. When *PCE*2 receives a backtrack message from *PCE*5 it sends update request messages to the *n*-neighboring PCEs and receives update responses. In this simple example, *n* is equal to one, *D*3 and *D*4 are in the set of the 1-neighboring of *D*2.

In Section 4, the effectiveness of the DDMB proposal is evaluated considering different numbers of advertising hops and the results are compared to those obtained with the DDSB.

4. Performance evaluation

In this section, the effectiveness of the proposals are evaluated and compared with each other. Moreover, results given by the OBGP protocol [14] are also considered for benchmark purpose, as in [29]. First, the effectiveness of the SDSB and DDSB proposals (Sections 3.1 and 3.2) are assessed and compared to OBGP. Then the proposal with better performance is compared to DDMB (Section 3.3) which considers different numbers of advertising hops.

Topologies used in the simulations were the NOBEL-EU (Fig. 6) and the NEWYORK (Fig. 7) mesh topologies, with descriptions available in the library of test instances for Survivable fixed telecommunication Network Design (SNDlib) [37]. The NOBEL-EU topology was originally defined in the COST 266 European project [38], and it has been used for protocol evaluation. It is composed of 28 domains and 41 inter-domain links, with a mean domain connectivity of 2.93. Nodes were chosen to include some of the main Internet exchange points. The NEWYORK network represents a telecommunication network in the greater New York area, but its origin is not known due to non-disclosure agreements. It has 16 domains and 49 inter-domain links, resulting in a mean domain connectivity of 6.12. These two topologies were chosen since the number of nodes and connectivity vary.

The nodes in each domain are fully connected and there are as many inter-domain links as the number of nodes at the border. In this way, blocking due to unavailability of intra-domain paths are avoided. The Dijkstra algorithm is used to define intra-domain paths.

Connection requests are uniformly distributed among all pair of nodes in the network. The arrival rate of calls and their duration follow, respectively, a Poisson and a negative exponential distribution. All the links have just one fiber and each fiber has 60 wavelengths. The network load varied from 100 to 900 Erlang. The load increases, it is due to a higher arrival rate and the mean holding time is set to 600 s. Each simulation involved 100,000 connection requests and confidence intervals at a 95% confidence level were derived using the independent replication method. At least 10 independent replications were generated for each experiment. The blocking probability and signaling overhead were assessed in the simulations.

Signaling overhead considers all control signaling generated by the inter-domain routing protocol to discover and maintain resource availability. For the schemes proposed, these messages are triggered by backtrack events. The *MinRouteAdvertisementIntervalTimer* parameter in the OBGP protocol, that determines the amount of time that must elapse between two BGP advertisements, was set to 30 s, as suggested in [17].

Figs. 8 and 9 compare the performance of SDSB, DDSB and OBGP. They show the blocking probability and the number of messages exchanged (overhead) as a function of the load for the NOBEL-EU topology. The blocking probability produced by OBGP is roughly equivalent to the blocking probability generated by SDSB but, after 120 Erlangs, the curve moves away reaching a 20% lower blocking probability under a load of 200 Erlang. Nevertheless, the DDSB approach produces almost no blocking despite of the network load. The results show a clear advantage for the adoption of the PCE-based approach with the destination-driven wavelength assignment (DDSB), which produced almost no blocking since the availability set of wavelengths is known at the time the end-to-end path is established. On the other hand, the source-driven approach (SDSB) presented a slightly worse performance than OBGP in terms of call blocking.

The proposed scheme balances the load by choosing inter-domain links which have the greatest number of wavelengths available. Such balance avoids the formation of unnecessary bottleneck which increases the blocking probability. Moreover, by updating the routing tables only during backtracking leads to more stable routing tables.

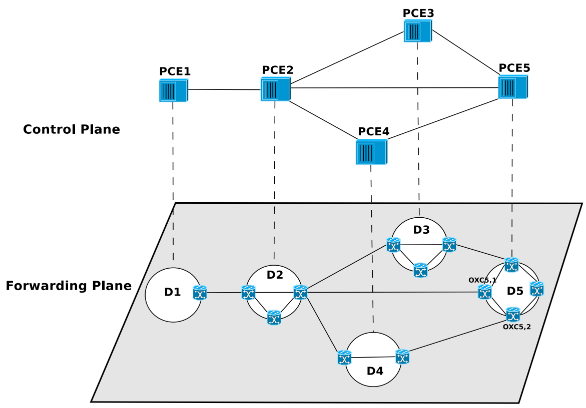

Fig. 4. PCE-based multi-domain scenario.

Fig. 9 shows that the difference in the number of signaling messages sent is quite striking. OBGP generates a few orders of magnitude more messages than does SDSB and DDSB. In the OBGP protocol, whenever a request cannot be forwarded to the next domain during the path establishment procedure, it is blocked and updating messages flood the network. In Fig. 10, which shows results for SDSD and DDSB alone, it can be seen that the number of messages generated by DDSB is almost zero even for a load of 200 Erlangs. Moreover, the SDSB approach produces a considerably greater number of messages than does the DDSB, since it generates a greater number of backtracking events.

Figs. 11 and 12 show the blocking probability and the number of messages exchanged (overhead) as a function of the load for the NEWYORK topology. The greater degree of connectivity of this topology leads to a lower blocking probability values than was found for the NOBEL-EU topology. Nonetheless, OBGP produces blocking probability slightly greater than those given by SDSB and one orders of magnitude greater than those generated by DDSB. This show that the proposed approaches are able to take advantage of the greater degree of connectivity of this topology by exploring alternative paths to avoid blocking.

The number of messages sent by OBGP is a few orders of magnitude greater than that generated by SDSB and even greater than that sent by DDSB. This huge difference is mainly due to the on-demand updating scheme of our proposal which is in clear contrast with the flooding updating scheme of the OBGP.

In SDSB, the source PCE decides on the choice of wavelength to establish a lightpath based on information obtained by the resource advertisement scheme, this does not necessarily provide complete updated information on resource availability due to scalability constraints. On the other hand, the destination selection scheme considers the actual wavelength availability during path computation, which gives the destination PCE a more precise view of the resources available.

The signaling overhead generated by the destination selection solution was low although for source selection it

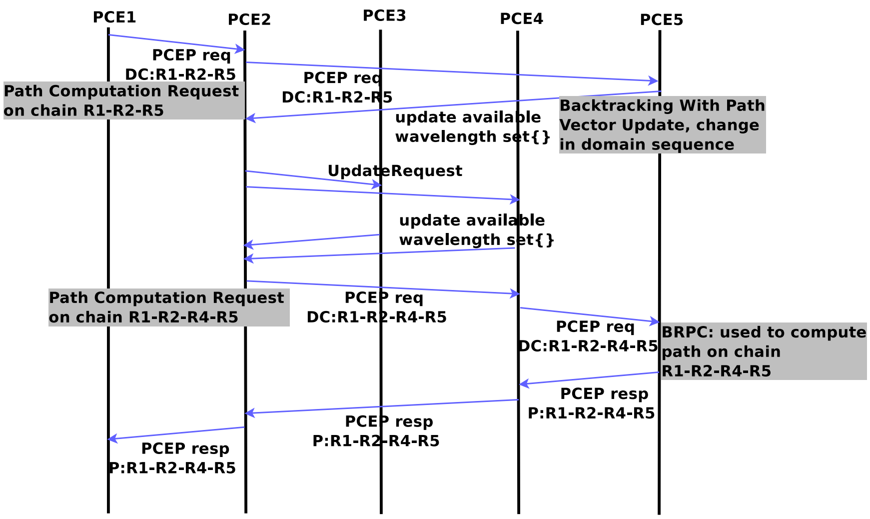

Fig. 5. Flow of PCEP request and update signaling.

Fig. 6. Figure representing the NOBEL-EU network topology.

increased sharply as load increased. As a consequence of the on-demand resource advertisement scheme, signaling overhead is inversely proportional to the path computation procedure, so the more efficient the path computation scheme, the lower the signaling overhead.

Figs. 13 and 14 compare the performance of DDSB and DDMB considering 2, 3 and 4 advertising hops. They show the blocking probability and the overhead as a function of the load for the NOBEL-EU topology. For the sake of comparison, results for a centralized algorithm (optimal path computation) are also presented. The blocking probability experienced by all the evaluated algorithms stay close to zero for network loads lower than 400 Erlang. As the load increases the blocking probability also increases evincing a small difference among the algorithms results, which are smaller as the number of advertising hops increases. There is a less than 2% decrease in blocking probability when advertisement is sent 4 hops away than when sent to a PCE one hop distant. The great increase in signaling when considering non-neighboring PCEs counteracts the potential gain decrease of blocking due to instability of the content of TEDs.

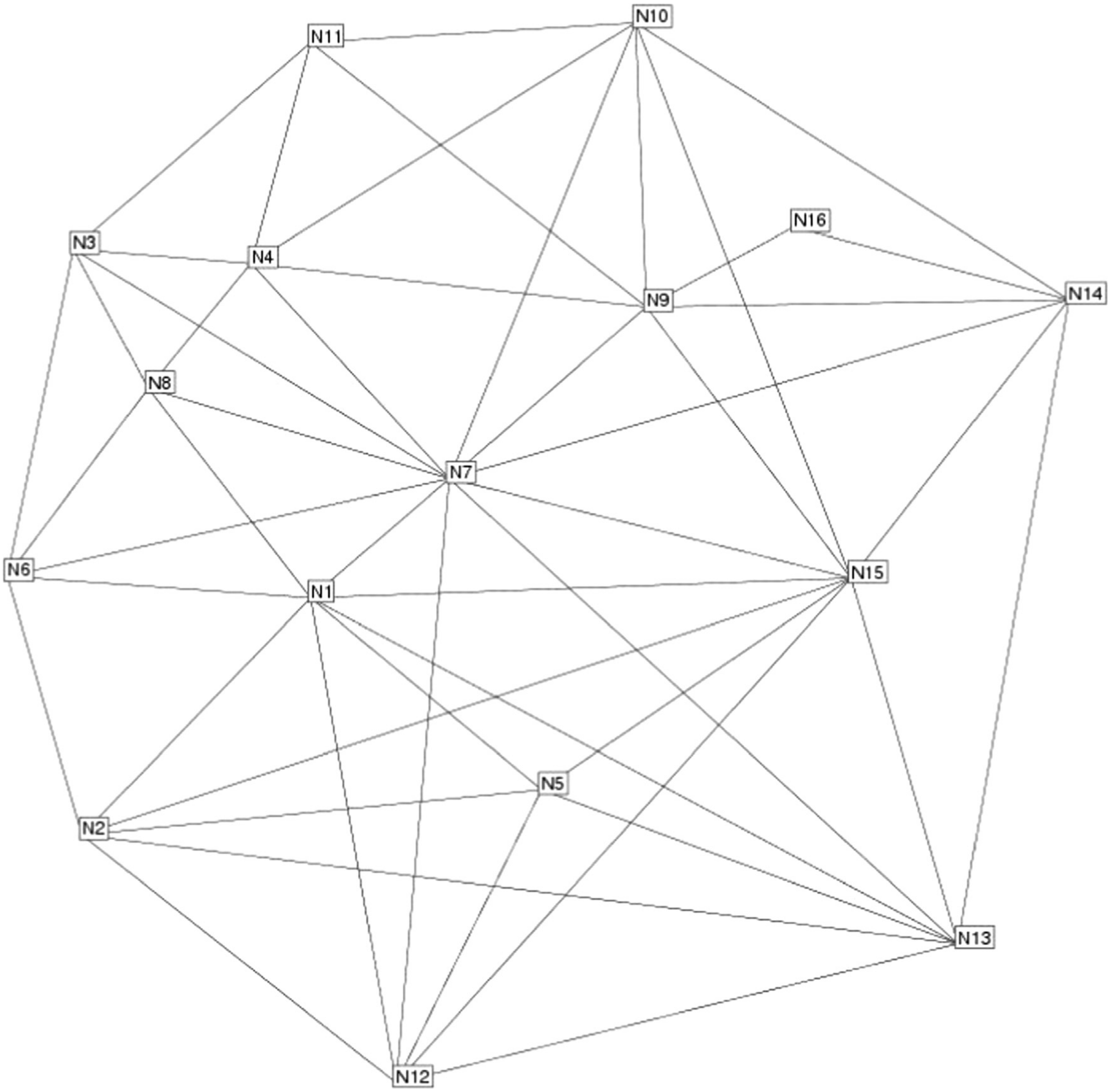

Fig. 7. Graph representing the NEWYORK network topology.

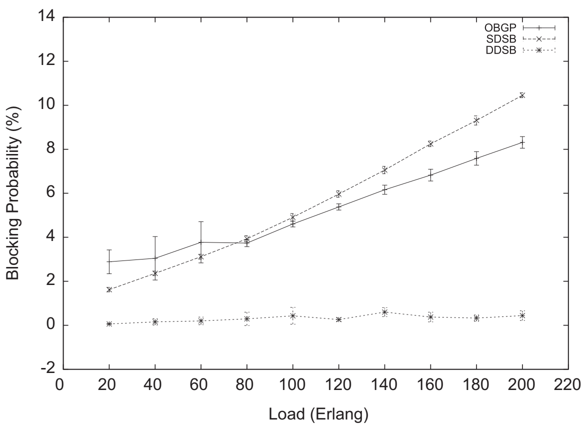

Fig. 8. Blocking probability as a function of the load for the NOBEL-EU topology.

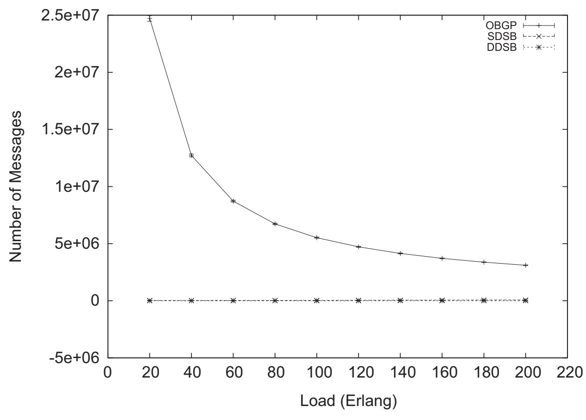

Fig. 9. Routing advertisements as a function of the load for the NOBEL-EU topology.

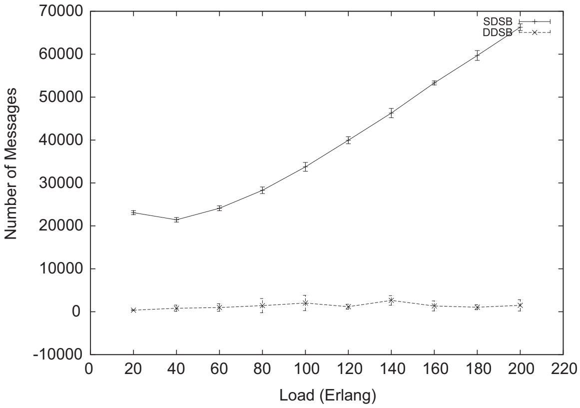

Fig. 10. Routing advertisements as a function of the load for the NOBEL-EU topology. Detailing the results for SDSB and DDSB.

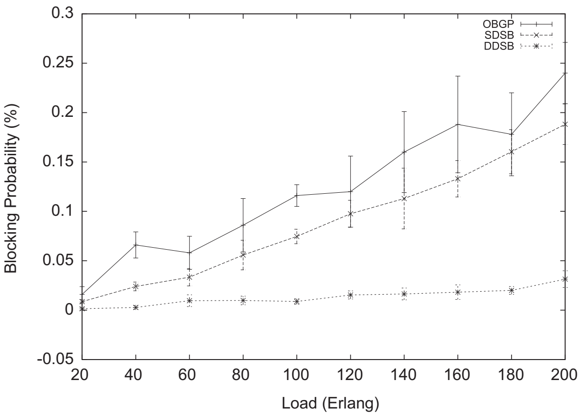

Fig. 11. Blocking probability as a function of the load for the NEWYORK topology.

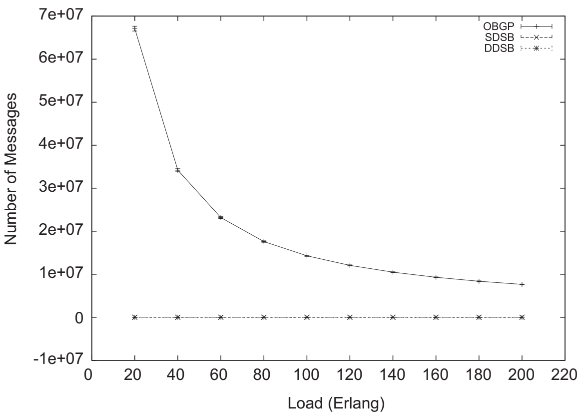

Fig. 12. Routing advertisements as a function of the load for the NEWYORK topology.

Increasing the hop distance for updating PCE TEDs strongly impacts the signaling overhead. Updating PCEs more than one hop away yields the signaling overhead to exchange of $10+6$ messages. The negative impact of this increase on blocking probability was also observed in the other proposals in this paper as well as in OBGP.

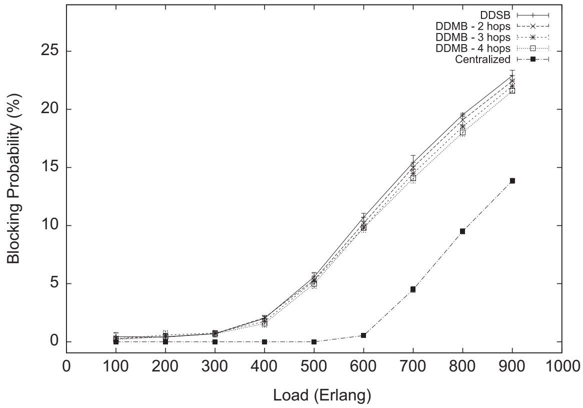

Fig. 13. Blocking probability as a function of the load for DDSB and DDMB considering 2, 3 and 4 hops for the NOBEL-EU topology.

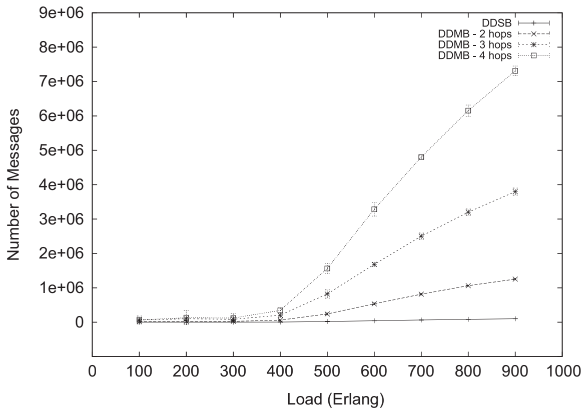

Fig. 14. Routing advertisements as a function of the load for DDSB and DDMB considering 2, 3 and 4 hops for the NOBEL-EU topology.

5. Conclusions

This paper proposes three novel solutions for inter-domain lightpath provisioning in Wavelength Division Multiplexing circuit switched mesh networks based on two routing advertisement schemes and two path and wavelength selection criteria.

The effectiveness of the proposed schemes are assessed and compared to that of OBGP. It was demonstrated that the destination-driven wavelength assignment and corresponding path computation police produce lower blocking and a huge decrease of signaling overhead. Moreover, advertising information to a higher number of PCEs as in the DDMB approach did not lower the blocking probability, especially for highly connected networks.

The destination-driven proposal preserves intra-domain confidential information. On-demand dissemination of information on reachability and resources availability allows a lightpath establishment procedure which includes backtracking as a solution for alternative attempts of those which failed. An RWA algorithm balances the load based on the number of available wavelengths per path which avoids the formation of bottlenecks and consequently decreases blocking. The proposals provide a complete inter-domain routing and wavelength solution.

Acknowledgments

We would like to thank Prof Admela Jukan for her valuable comments. This work was sponsored by FAPESP (grant no 2010/03020-4 and 08/57857-2) and CNPq (grant no 305236/2014-8).

References

[1] The internet engineering task force. URL ⟨http://www.ietf.org/⟩, August 2011.
[2] A. Farrel, J.-P. Vasseur, J. Ash, A path computation element (PCE)-based architecture, RFC 4655 (Informational). URL ⟨http://www.ietf.org/rfc/rfc4655.txt⟩, August 2006.
[3] M. Drogon, A. Jukan, M, Chamania, An open-source path computation element (PCE) emulator: design, implementation and performance, J. Lightwave Technol., 30 (4), http://dx.doi.org/10.1109/JLT.2011.2170402.
[4] F. Paolucci, F. Cugini, A. Giorgetti, N. Sambo, P. Castoldi, A survey on the path computation element (PCE) architecture, IEEE Commun. Surv. Tutor. 15 (4) (2013) 1819–1841, http://dx.doi.org/10.1109/SURV.2013.011413.00087.
[5] Ó.G. de Dios, V. Lopez-Alvarez, M. Cuaresma, F. Muñoz, M. Chamania, A. Jukan, Coordinated computation and setup of multi-layer paths via inter-layer PCE communication: standards, interoperability and deployment, IEEE Commun. Mag. 51 (12) (2013) 144–154, http://dx.doi.org/10.1109/MCOM.2013.6685770.
[6] M.C. Saturio, V. López, Ó.G. de Dios, F.M. del Nuevo, J.P. Fernández-Palacios, Implementation and assessment of pre-reservation mechanism for PCE environments, J. Netw. Syst. Manag. 22 (3) (2014) 488–508, http://dx.doi.org/10.1007/s10922-013-9296-y.
[7] J. Vasseur, A. Ayyangar, R. Zhang, A per-domain path computation method for establishing inter-domain traffic engineering (TE) label switched paths (LSPs), RFC 5152 (Standards Track). URL ⟨http://www.ietf.org/rfc/rfc5152.txt⟩, February 2008.
[8] J. Vasseur, R. Zhang, N. Bitar, J.L. Roux, A backward-recursive PCE-based computation (BRPC) procedure to compute shortest constrained inter-domain traffic engineering label switched paths, RFC 5441 (Standards Track). URL ⟨http://www.ietf.org/rfc/rfc5441.txt⟩, April 2009.
[9] D. King, A. Farrel, The application of the path computation element architecture to the determination of a sequence of domains in MPLS and GMPLS, RFC 6805, November 2012.
[10] J. Vasseur, J.L. Roux, Path computation element (PCE) communication protocol (PCEP), RFC 5440. URL ⟨http://www.ietf.org/rfc/rfc5440.txt⟩, March 2009.
[11] M. Chamania, A. Jukan, A survey of inter-domain peering and provisioning solutions for the next generation optical networks, IEEE Commun. Surv. Tutor. 11 (2009) 33–51.
[12] G. Bernstein, Y. Lee, A. Gavler, J. Martensson, Modeling WDM wavelength switching systems for use in GMPLS and automated path computation, IEEE/OSA J. Opt. Commun. Netw. 1 (1) (2009) 187–195, http://dx.doi.org/10.1364/JOCN.1.000187.
[13] H. Song, M. Brandt-Pearce, Range of influence and impact of physical impairments in long-haul DWDM systems, J. Lightwave Technol. 31 (6) (2013) 846–854, http://dx.doi.org/10.1109/JLT.2012.2235409.
[14] M. Blanchet, F. Viagenie, B. Canarie, Optical BGP (OBGP): InterAS lightpath provisioning, Internet-Draft draft-parent-obgp-01, Internet Engineering Task Force, work in progress. URL ⟨http://tools.ietf.org/id/draft-parent-obgp-01.txt⟩, March 2001.
[15] M. Francisco, L. Pezoulas, C. Huang, I. Larnbadaris, End-to-end signaling and routing for optical IP networks, in: IEEE International Conference on Communications, vol. 5, 2002, pp. 2870–2875, http://dx.doi.org/10.1109/ICC.2002.997365.
[16] N.-D. Tran, W.-F. Wang, Experimenting OBGP for inter-domain optical switching, in: International Conference on Advanced Technologies for Communications, ATC 2008, 2008, pp. 120–123, http://dx.doi.org/10.1109/ATC.2008.4760534.
[17] Y. Rekhter, T. Li, S. Hares, A border gateway protocol 4 (BGP-4), RFC 4271 (Standards Track). URL ⟨http://www.ietf.org/rfc/rfc4271.txt⟩, January 2006.
[18] M. Yannuzzi, X. Masip-Bruin, G. Fabrego, S. Sanchez-Lopez, A. Sprintson, A. Orda, Toward a new route control model for multidomain optical networks, IEEE Commun. Mag. 46 (6) (2008) 104–111, http://dx.doi.org/10.1109/MCOM.2008.4539473.
[19] H. Gredler, J. Medved, S. Previdi, A. Farrel, S. Ray, North-bound distribution of link-state and TE information using BGP, draft-ietf-idr-ls-distribution-04, November 2013.
[20] R. Casellas, R. Martínez, R. Muñoz, S. Gunreben, Enhanced backwards recursive path computation for multi-area wavelength switched optical networks under wavelength continuity constraint, J. Opt. Commun. Netw. 1 (2) (2009) A180–A193, http://dx.doi.org/10.1364/JOCN.1.00A180. URL ⟨http://jocn.osa.org/abstract.cfm?URI=jocn-1-2-A180⟩.
[21] M. Yannuzzi, X. Masip-Bruin, S. Sánchez-López, E. Marín-Tordera, Obgp+: an improved path-vector protocol for multi-domain optical networks, Opt. Switch. Netw. 6 (2) (2009) 111–119, http://dx.doi.org/10.1016/j.osn.2008.12.001. recent trends on optical network design and modeling—selected topics from ONDM 2008. URL ⟨http://www.sciencedirect.com/science/article/B7GX5-4VCNP8T-1/2/617a86908232171c5a19c42a5ec2d4f2⟩.
[22] X. Chen, A. Jukan, F. Rambach, M. Hoffman, M. Chamania, An adaptive inter-domain PCE framework to improve resource utilization and reduce inter-domain signaling, Opt. Switch. Netw. 6 (4) (2009) 259–267, http://dx.doi.org/10.1016/j.osn.2009.08.007.
[23] Q. Zhang, M. Hasan, X. Wang, P. Palacharla, M. Sekiya, Survivable path computation in PCE-based multi-domain networks, IEEE/OSA J. Opt. Commun. Netw. 4 (6) (2012) 457–467, http://dx.doi.org/10.1364/JOCN.4.000457.
[24] H. Chen, A forward-search P2P TE LSP inter-domain path computation, Internet-Draft draft-chen-pce-forward-search-p2p-path-computation-04, Internet Engineering Task Force, work in progress. URL ⟨http://tools.ietf.org/html/draft-chen-pce-forward-search-p2p-path-computation-04⟩, March 2012.
[25] F. Cugini, G. Meloni, F. Paolucci, N. Sambo, M. Secondini, L. Gerardi, L. Poti, P. Castoldi, Demonstration of flexible optical network based on path computation element, J. Lightwave Technol. 30 (5) (2012) 727–733, http://dx.doi.org/10.1109/JLT.2011.2180361.
[26] R. Casellas, R. Munoz, J. Fabrega, M. Moreolo, R. Martinez, L. Liu, T. Tsuritani, I. Morita, GMPLS/PCE control of flexi-grid DWDM optical networks using co-ofdm transmission [invited], IEEE/OSA J. Opt. Commun. Netw. 4 (11) (2012) B1–B10, http://dx.doi.org/10.1364/JOCN.4.0000B1.
[27] J. Zhao, S. Subramaniam, M. Brandt-Pearce, Intradomain and inter-domain QoT-aware RWA for translucent optical networks, IEEE/OSA J. Opt. Commun. Netw. 6 (6) (2014) 536–548, http://dx.doi.org/10.1364/JOCN.6.000536.
[28] J. Ahmed, C. Cavdar, P. Monti, L. Wosinska, A dynamic bulk provisioning framework for concurrent optimization in PCE-based WDM networks, J. Lightwave Technol. 30 (14) (2012) 2229–2239, http://dx.doi.org/10.1109/JLT.2012.2195296.
[29] E. Marín-Tordera, X. Masip-Bruin, M. Yannuzzi, R. Serral-Gracià, Multi-domain optical routing: is there life beyond extending BGP? Opt. Switch. Netw. 11 (2014) 119–135, http://dx.doi.org/10.1016/j.osn.2013.05.005.
[30] D. Siracusa, S. Grita, G. Maier, A. Pattavina, F. Paolucci, F. Cugini, P. Castoldi, Domain sequence protocol (DSP) for PCE-based multi-domain traffic engineering, J. Opt. Commun. Netw. 4 (11) (2012) 876–884, http://dx.doi.org/10.1364/JOCN.4.000876. URL ⟨http://jocn.osa.org/abstract.cfm?URI=jocn-4-11-876 ⟩.
[31] S. Shang, N. Hua, L. Wang, R. Lu, X. Zheng, H. Zhang, A hierarchical path computation element (PCE)-based k-random-paths routing algorithm in multi-domain {WDM} networks, Opt. Switch. Netw. 8 (4) (2011) 235–241, http://dx.doi.org/10.1016/j.osn.2011.06.006. optical network architectures and management. URL ⟨http://www.sciencedirect.com/science/article/pii/S1573427711000543 ⟩.
[32] L. Buzzi, M. Bardellini, D. Siracusa, G. Maier, F. Paolucci, F. Cugini, L. Valcarenghi, P. Castoldi, Hierarchical border gateway protocol (HBGP) for PCE-based multi-domain traffic engineering, in: 2010 IEEE International Conference on Communications (ICC), 2010, pp. 1–6, http://dx.doi.org/10.1109/ICC.2010.5502390.
[33] R. Lu, X. Zheng, N. Hua, Domain-level-based routing algorithm in multidomain WDM optical network, J. Opt. Commun. Netw. 5 (5) (2013) 532–543, http://dx.doi.org/10.1364/JOCN.5.000532. URL ⟨http://jocn.osa.org/abstract.cfm?URI=jocn-5-5-532⟩.
[34] D. Medhi, K. Ramasamy, Network Routing: Algorithms, Protocols and Architectures, Morgan Kaufmann Publishers, 2007 (an imprint of Elsevier)., San Francisco.
[35] A. Farrel, A. Satyanarayana, A. Iwata, N. Fujita, G. Ash, Crankback signaling extensions for MPLS and GMPLS RSVP-TE, RFC 4920 (Standards Track). URL ⟨http://www.ietf.org/rfc/rfc4920.txt⟩, July 2007.
[36] D. Awduche, L. Berger, T. Li, V. Srinivasan, G. Swallow, RSVP-TE: extensions to RSVP for LSP tunnels, RFC 3209 (Standards Track). URL ⟨http://www.ietf.org/rfc/rfc3209.txt ⟩, December 2001.
[37] Survivable fixed telecommunication network design library. URL ⟨http://sndlib.zib.de⟩.
[38] Project cost 266. URL ⟨http://www.ikr.uni-stuttgart.de/Content/ResearchProjects/View/Full.html?cost266⟩.