Copyright Notice

Institute of Communication Networks and Computer Engineering
University of Stuttgart
Pfaffenwaldring 47, D-70569 Stuttgart, Germany
Phone: ++49-711-685-68026, Fax: ++49-711-685-67983
Email: mail@ikr.uni-stuttgart.de, http://www.ikr.uni-stuttgart.de

Enhanced BRPC for multi-domain PCE-based path computation in Wavelength Switched Optical Networks under Wavelength Continuity Constraint

Ramon Casellas,[1,*] Ricardo Martínez,[1] Raül Muñoz,[1] and Sebastian Gunreben [2]

[1]Optical Networking Area, Centre Tecnològic de Telecomunicacions de Catalunya, CTTC, Av. Canal Olímpic s/n, 08860 Castelldefels, Barcelona, Spain

[2]Institute of Communication Networks and Computer Engineering, University of Stuttgart, Pfaffenwaldring 4, 70569 Stuttgart, Germany

[*]Corresponding author: ramon.casellas@cttc.es

In the context of the Future Internet, all-optical Wavelength Switched Optical Networks will play an important role in either evolutionary or revolutionary design paradigms. In any paradigm, Dense Wavelength Domain Multiplexing is the most cost-effective technology for the increasing bandwidth capacity. It provides the basis for a core optical transport infrastructure and supports a wide range of heterogeneous services. However, such all-optical networks raise well-known challenges such as the wavelength continuity constraint (WCC). The WCC is hard to address in a multi-domain scenario when provisioning an end-to-end lightpath due to network topology hiding requirements and the limited exchange of information between domains. The IETF is currently standardizing the Path Computation Element (PCE) architecture, a good candidate to perform multi-domain path computation. In such architecture, the approach named

Backwards Recursive Path Computation (BRPC), also under standardization at the IETF, aims at overcoming the limitations of the “Per-Domain” mechanism. However, although BRPC does provide end-to-end shortest paths, it fails to take into account the WCC, which is the main motivation for this work. In this paper, we extend the BRPC algorithm and the companion PCE protocol, in order to address the end-to-end WCC efficiently. We perform a quantitative comparative analysis of the different approaches, experimentally showing the improvements of the conceived solution, which has been evaluated in a GMPLS-controlled network of the ADRENALINE testbed.

OCIS codes: 060.0060, 060.1155, 060.4251, 060.4264, 060.4265.

1. Introduction

In the context of the Future Internet, influenced by the scaling requirements of transport networks, Wavelength-Switched Optical Networks (WSON) will play an important role in either evolutionary or revolutionary approaches. WSONs leverage the advances in both the all-optical switching (e.g., Reconfigurable Optical Add Drop Multiplexer or ROADMs and Optical Cross-Connects or OXCs) and in tunable transceivers, taking full advantage of the huge transport capacity provided by the Dense Wavelength Domain Multiplexing (DWDM) technology. In this architecture, high-bandwidth end-to-end optical connections (i.e., lightpaths) are entirely set up within the optical domain, without optical/electrical transceivers at intermediate nodes. Lightpaths are thus transparent from source to destination with regard to the format and payload of the carried signal. However, all-optical wavelength converters are scarce and expensive resources. Consequently, a cost-efficient strategy in all-optical, transparent networks is to allocate the same wavelength on each link along the computed route for each lightpath. This

imposes and additional restriction, usually referred to as the wavelength continuity constraint (WCC). On the other hand, the introduction of automatic and distributed control planes such as the Generalized Multi-Protocol Label Switching (GMPLS) architecture [1] facilitates the dynamic provisioning of traffic-engineered lightpaths.

Additionally, for scalability or confidentiality reasons, network operators may divide a transparent optical network into domains. In our considered scenario, the transparent optical network uses the Open Shortest Path First – Traffic Engineering (OSPF-TE) [2] as the Interior Gateway Protocol (IGP), and OSPF-TE areas correspond to Traffic Engineering (TE) domains. The TE domain boundary nodes correspond to OSPF-TE Area Border Routers (ABRs).

In the GMPLS architecture, Label Switched Routers or LSRs (i.e., Optical Connection Controller or OCC in WSON) have full topology visibility within their domain boundaries and limited visibility of the other domains, usually as aggregated information (e.g. reachability). Consequently, in traditional source routing approaches, a source OCC is not able to compute, autonomously, an end-to-end inter-domain path with the same control and degree of Traffic Engineering than for an intra-domain path. In this context, two methods are applicable for inter-domain path computation, the per-domain path computation method and the PCE-based path computation method.

- The *per-domain path computation* method: the source OCC determines the next domain and the ingress within that domain. Then, it computes the corresponding path segment to the domain boundary, obtaining a strict Explicit Route Object (ERO) within its own domain, and appending to it (a list of) loose hops for the neighbor domain towards the destination. Next, the path computation moves to the ingress OCC of the next domain, and so forth until the destination domain. During the

signaling phase, the OCC at the boundary domain expands the ERO. As a result, this simple method disables the computation of a shortest inter-domain path from an end-to-end lightpath perspective.

- The *PCE-based path computation* method: this method assumes a domain-chain or succession of TE domains transit from source to destination. The method relies on dedicated Path Computation Elements (PCEs), which collaboratively compute an inter-domain optimum path along the given domain chain. Each PCE is responsible for the path computation within its domain.

To this end, the IETF PCE working group has defined the PCE architecture [3] and a communications protocol (PCEP) [4], so Path Computation Clients (PCCs) such as OCCs may request the computation of an explicitly routed path given a set of constraints. Such an architecture is motivated by the complexity of path computation in large, multi-domain, multi-region, or multi-layer networks, and that of advanced (e.g. protection-enabled) algorithms and heuristics, which may eventually require dedicated computational resources and cooperation between network domains. The new architecture raises new challenges regarding the feasibility and applicability of the PCE in general, and in GMPLS-controlled WSONs in particular. This includes, for example, the WCC, which may significantly degrade performance if not addressed correctly. Finally, the PCE-based method may be the one to be preferred when the end-to-end service needs to be policed (admission control, billing, etc.) as detailed in [5]. This may raise new opportunities due to the coupling control between the PCE and the Network Management System and Business Support Services (NMS-BSS).

Most research efforts on the PCE-based multi-domain path computation seem targeted to the Backwards Recursive Path Computation (BRPC) [6] procedure. The BRPC, under

standardization at the IETF, is currently the one that meets best the operator and the supplier requirements in terms of complexity and network information hiding. It involves the recursive computation, at each PCE of the TE domain chain, of an inverse tree of constrained shortest paths, with one branch for each domain ingress node to the destination, using the inverse tree computed by the downstream domain, as detailed in the next section. Topology confidentiality is reasonable, since no TE information exchange is required between PCEs. Additional topology confidentiality may be achieved by means of the Path Key mechanism, in which a "cookie" or token is sent upstream rather than the actual ERO, which is expanded during the actual path signaling at the concerned domain [7]. The BRPC and the PCE architecture provide an opportunity for advanced TE schemes in the multi-domain setting. A detailed introduction to the PCE and BRPC architectures, along with a simulation based comparative analysis of BRPC and the per-domain method for MPLS/GMPLS networks without WCC appears in [8]. Recently, related works are extending this approach in order to deploy protected label switched paths as in, for example, [9]. In the following sections, we will introduce the problem statement, giving an overview of the standard BRPC and highlighting its limitations when used in WSON with WCC. Second, we detail the algorithm of our devised solution, which extends the standard BRPC. Then, we present the design and implementation of the solution, highlighting the functional architecture and the required control plane extensions. Next, we experimentally evaluate the presented approaches in the ADRENALINE testbed, giving a comparative analysis with numerical results and key performance indicators. Finally, we conclude the paper.

2. Problem Statement

2.A. The Backwards Recursive Inter-domain Path Computation

We will overview the BRPC algorithm with the help of Fig. 1. Let us assume that BRPC knows the domain sequence in advance. The algorithm computes the inter-domain path in a reverse way, starting with the destination domain (i.e., in a backward direction). The destination domain PCE computes a Virtual Shortest Path Tree (VSPT) from the domain ingress nodes (in-ABRs in our context) to the destination node. The destination domain PCE sends the computed VSPT to the upstream PCE. This PCE uses this information to compute its own VSPT: computing first a tree from its domain ingress ABRs that are adjacent to the upstream domain (in-ABRs) to each of its domain egress ABRs adjacent to the downstream domain (out-ABRs). Then it selects the optimal path from each of the in-ABRs to the destination node (through any of the out-ABRs), pruning the sub-optimal paths from the VSPT before sending it to its own upstream domain PCE. In other words, the PCE computes its own VSPT from the in-ABRs using the received paths in the VSPT as *extended TE links*. The upstream domains (i.e., recursive) apply this procedure up to the source domain.

For illustration, consider the example in Fig. 1. In order to compute a path (a-s), the PCE in domain C computes, during the step I, the tree of shortest paths (VSPT), namely (k-n-s), (l-n-p-s) and (m-p-s). The resulting VSPT is sent to the PCE in domain B, which (step II), computes first the set of paths between in-ABRs (nodes f, g) and out-ABRs (nodes k, l, m), obtaining (f-j-k), (f-h-i-m), (g-h-j-k), (g-i-m) and then uses the received information to compute the optimal paths from in-ABRs towards the destination, selecting (f-j-k-n-s) and (g-i-m-p-s). At step III, the PCE recursively applies the same method obtaining the best path from source to destination, namely (a-d-f-j-k-n-s).

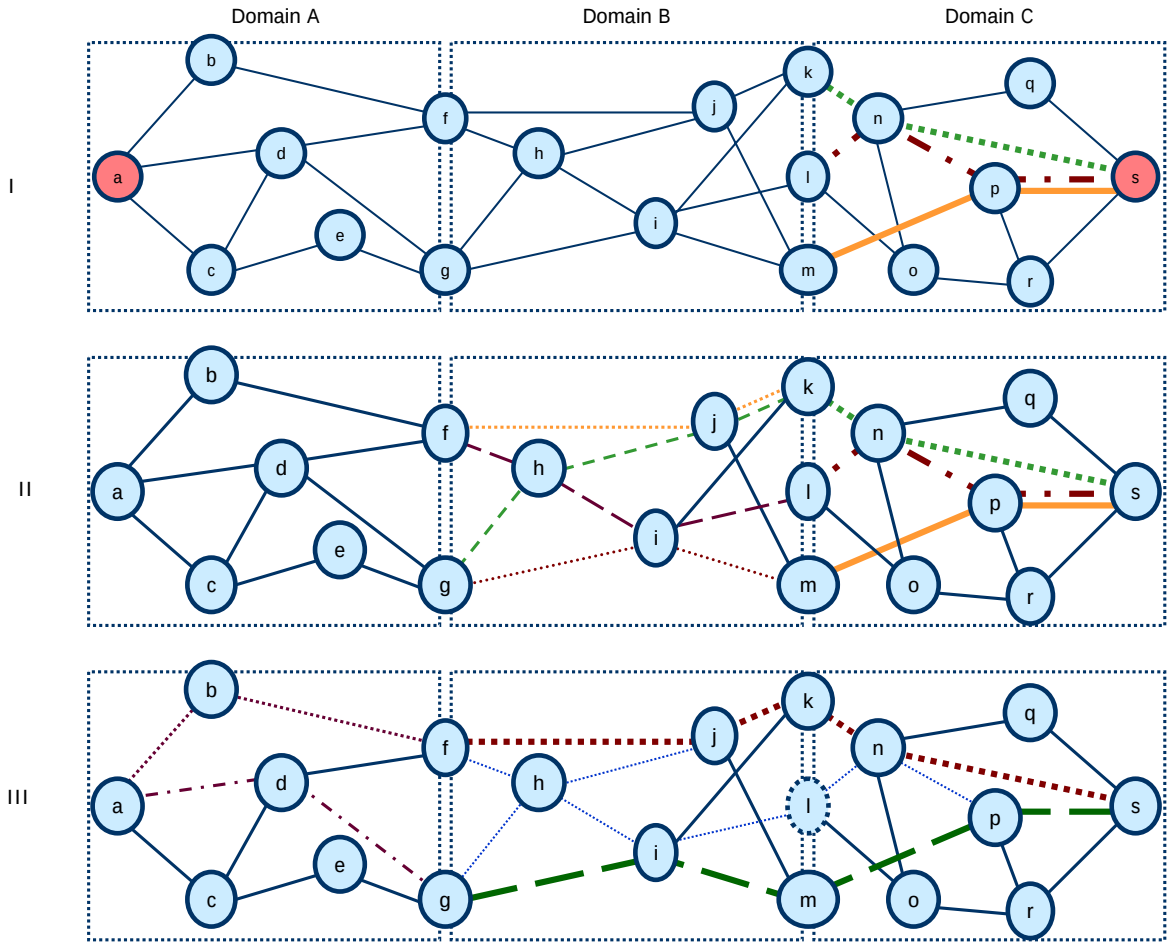

Fig. 1. BRCP procedure for the multi-domain path computation from source node a to destination node s

One of the drawbacks of the BRPC procedure as defined in current internet drafts [6] is that the protocol does not convey enough information in order to perform efficient Wavelength Assignment (WA) in a collaborative setting while insuring wavelength continuity. As mentioned, WCC is a necessary constraint in transparent networks. This motivates the extension of the PCE and BRPC approaches to address this issue.

The purpose of the work in this article is two-fold: first, to extend the BRPC to take into account specific requirements of transparent optical networks in order to conceive algorithms within the BRPC procedure that efficiently address the WCC with distributed (amongst the

PCEs) RWA. Second, to evaluate the performance of these PCE based solutions once deployed in a real GMPLS-enabled network, compared to per-domain path computation.

The next section details our devised and validated solution, which aims at efficiently addressing Routing and Wavelength Assignment in the presence of WCC while retaining the advantages of a BRPC-based optimal path computation.

3. Path Computation and Enhanced BRPC for WSON

3.A. PCE-based Routing and Wavelength Assignment

Using standard OSPF-TE in the per-domain approach, the Routing Controller (RC) of a GMPLS-enabled OCC performs only routing (R) and no Wavelength Assignment (WA) as OSPF-TE only disseminates aggregated unreserved bandwidth information and no information on the wavelength status. The Connection Controllers (CCs) perform the WA in a distributed way along the path: the source node includes a Label Set (LS) object within the Path message of the Resource Reservation Protocol with Traffic Engineering extensions (RSVP-TE) [10-11]. The LS initially contains a set of available wavelengths. Along the path, the CC removes wavelengths from this set if they are not available at the output link. The destination node assigns the allocated wavelength using usually random or first-fit heuristics on the remaining wavelengths of the set. Note that if the Label Set becomes empty, the connection is blocked, indicated by a Path Error. If OSPF-TE is able to disseminate the wavelength status within a domain, the RC can only perform routing and WA (RWA) for the actual domain up to the egress ABR as it lacks the visibility of the downstream domains. Similarly, in the standard BRPC, the PCE performs routing based on shortest paths and TE metrics. It leaves the WA process to the signaling phase.

In our extended BRPC detailed below, that we name EBRCP-WSON, the PCE chain performs both routing and WA, using new PCEP objects to communicate the wavelength status. Note that in all cases, the PCE or RC performs the WA passing the selected wavelength to RSVP-TE as the Suggested Label (SL), a "hint" for the preferred allocated wavelength. If the SL is not available at one hop in the path, the OCC removes the SL from the path message falling back to selecting one amongst the trimmed Label Set (LS) object as detailed above. Alternatively, the operator may enforce the allocated SL a Label Set with only a single label containing the SL value, or by means of explicit label control [12].

3.B. Intra-domain Path Computation Algorithm

The implemented path computation algorithm within a TE domain (OSPF-TE area) uses the modified Dijkstra shortest paths algorithm. The algorithm computes the path (ERO), the total TE metric from source *s* to destination *d* and the set of candidate (available) wavelengths. During the execution of the algorithm, and when considering a TE-link for relaxation (i.e. to be accepted in the shortest path tree), in addition to standard constraints (e.g., Shared Risk Link Groups or SRLGs, administrative groups, unreserved bandwidth…) the set of candidate wavelengths towards *d* trimmed considering the wavelength status in the TE link. If the set is empty, the link is not considered. More precisely, the modified shortest path algorithm is as follows:

At step 1, the initialization phase, the distance to all nodes is set to infinity; the set of available wavelengths is set to the empty set and the predecessor of the node to unknown.

At step 2, an ordered set named PendingNodeSet (also known as Q-set) containing the set of ordered nodes to visit is initialized with the source node.

At step 3, the algorithm keeps on iterating as long as there are pending nodes to visit and for each visited node u and considered outgoing link e, it checks the constraints including the

WCC. If the link is relaxed, the cost to the neighbor node v is updated, the set of available wavelengths to the node v is set to the available ones to node u (Su) minus the used wavelengths, and the node v is added to the PendingNodeSet. Fig. 2 gives the pseudo-code of the described algorithm.

```
1. Initialise all vertices in graph:
       Distance from source node is Infinity
       Prececessor of the node is Unknown

2. Visit the source node
       Distance from source node is 0
       Wavelengths from source is AllWavelengths
       Append source node to PendingNodeSet (Dijkstra Q-set) at distance 0

3. While PendingNodeSet is not Empty (Remaining Nodes)
       Extract next Node u from PendingNodeSet
       For all out edges (e) of Node u
              Consider Next node through u, v
              Set new_distance = distance (source -> u) + distance (u -> v)
              If new_distance <= known distance (source -> v)
                     Construct the set S of available wavelengths = Su
                     Remove form S all unavailable wavelengths in link e: u -> v
                     If S is empty discard link
                     If new_distance = known distance
                            If cardinal S < cardinal Sv discard link
                     Relax link e
                            Update PendingNodeSet with v reachable with new_distance
                            Set precedessor of v to u
                            Set Sv to S
```

Fig. 2. Pseudo-code for Intra-domain Path Computation Algorithm

3.C. Proposed Inter-domain Path Computation Algorithm with Wavelength information

The application of the EBRPC-WSON reduces to the computation of the VSPT from one or more source nodes to one or more destination nodes, and the trimming of suboptimal paths, while meeting the WCC. Source nodes are the domain ingress ABRs (denoted by in-ABRs) in the transit domains or the actual lightpath source node in the source domain. Destination nodes are the egress ABRs (denoted by out-ABRs) for transit domains or the lightpath destination node in the destination domain. For each of the source nodes, denoted by *ns*, the purpose of the algorithm is to find a (ordered) set of extended paths from ns to the end destination, denoted by *nd*, and to keep the optimal path ns → nd. In this context, an extended path p1 is better than an extended path p2 (denoted p1 < p2) if the TE metric of p1 is lower that the one of p2 and, in case of equal TE metric, if the number of end-to-end available wavelengths in p1 is equal or greater than the number in p2. Fig. 3 shows the pseudo code for the proposed algorithm in the transit (source and destination domains are particular cases). Note that all transit nodes perform the WA step (optionally), as a means to convey information on the preferred wavelength to the upstream PCE, which may discard it. .

INPUT: TE database in the domain, in-ABRs, out-ABR, PCRep message from the downstream PCE.

OUTPUT: Optimal path from each in-ABR to final destination, including available end-to-end wavelengths

```
For each source node ns in all in-ABRs in transit domain
```

```
    For each path p in each response of the PCRep message to the destination
```

- The out-ABR is the first subobject of the ERO attribute of the path p
- Set current domain destination node the out-ABR
- If ns = out-ABR, the extended path is the one given, proceed to next path

- Compute path from ns to out-ABR, using intra-domain algorithm

With set of paths: ns to out-ABR1, to out-ABR2, out-ABRn

- Obtain extended path (ep) ns to nd by:

  ERO merging: ns → out-ABRi + out-ABRi→nd becomes merged ns→nd

  Additive computation of TE metric and hop count

  Intersection of avail. wavelengths (LS)
  ns → out-ABR and out-ABR→dest.

  Constructin of of the resulting label set object.

- Perform WA for ep (optional in transit domain, mandatory in ingress domain) and construct the suggested (preferred) label.

- Select from ep set optimal path using TE metric and wavelength set size

Fig. 3. Pseudo-code for the proposed Inter-domain Path Computation Algorithm with Wavelength information

4. Design and development of a PCE

4.A. Path Computation Element functional architecture

The implementation of our deployed PCE involves a single, multi-threaded and asynchronous process (Fig. 4). One or more dedicated threads are responsible for updating the traffic engineering database (TE updaters), and another thread from a thread pool is responsible for the actual path computation, using a writer/readers lock. Upon acceptance of a connection, the Finite State Machine drives the PCEP protocol. Dynamic shared libraries provide pluggable algorithms, following an algorithm API (Application Programming Interface). The API allows abstracted access to the underlying TE database in form of a directed graph. Further, it allows the request for path computation to other PCE peers for cooperative path computation as in the BRPC. Our PCE has been successfully used within a single domain for Shared Path Protection (see [13,14]).

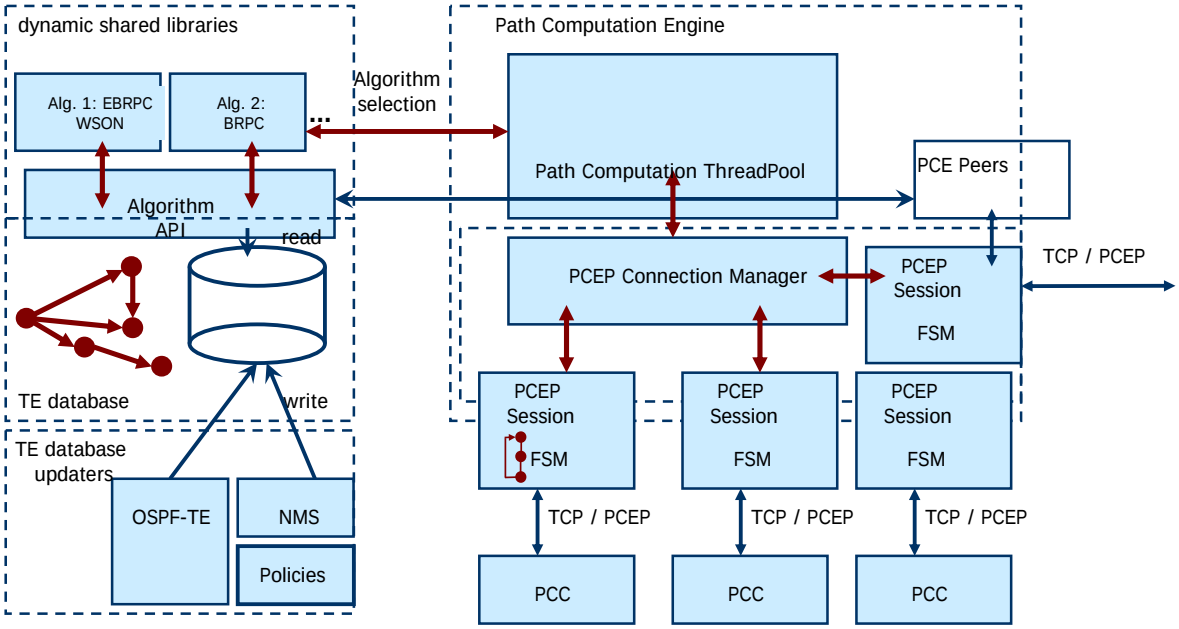

Fig. 4. Functional Architecture of the deployed Path Computation Element

4.B. PCE deployment model

Our PCE deployment model relies on a single PCE per OSPF area, co-located in an OCC. The deployed synchronization mechanism with the Traffic Engineering database (TEDB), although coupled to a Routing Controller (RC) is non-intrusive. By sniffing OSPF-TE traffic, the PCE constructs a dedicated (i.e. not shared) database by means of stateful inspection of TE Link sub-TLVs contained within OSPF-TE Link State Updates, passively reusing the OSPF-TE dissemination mechanism and not requiring the creation of an additional listener adjacency. The PCE performs ABR discovery by parsing OSPF-TE type 3 Summary-LSAs, which announce reachability information towards both the source and the destination.

Note that, without loss of generality, the same EBRCP-WSON procedure could be applied even if the PCE obtains topology information and wavelength status by other means (e.g. via the NMS in the Management Plane).

4.C. Control Plane Extensions

In this work, we have extended the deployed control plane protocols, namely OSPF-TE (IGP TE routing), RSVP-TE (signaling) and PCEP (path computation protocol) to better address WSON requirements. This involves either proprietary extensions or related ones proposed by relevant state of art and normalization efforts at the IETF Common Control and Measurement Plane (CCAMP) and PCE working groups. In particular, the bitmap-based OSPF-TE wavelength information dissemination and specific PCEP extensions for optical networks [15,17]. The following sections detail the implemented extensions and their purpose.

4.D. Routing Extensions

OSPF-TE routing extensions are mandatory if, as in the case of our PCE deployment model, the PCE obtains the TE topology and resource status by stateful inspection of TE Link State Advertisements (LSAs) included in the OSPF-TE update messages. In this sense, to provide better granularity regarding the state of the individual wavelengths of a TE link, OSPF-TE disseminates bitmap encoded wavelength status (for exact details and other related extensions, see [15]) as shown in Fig. 5.

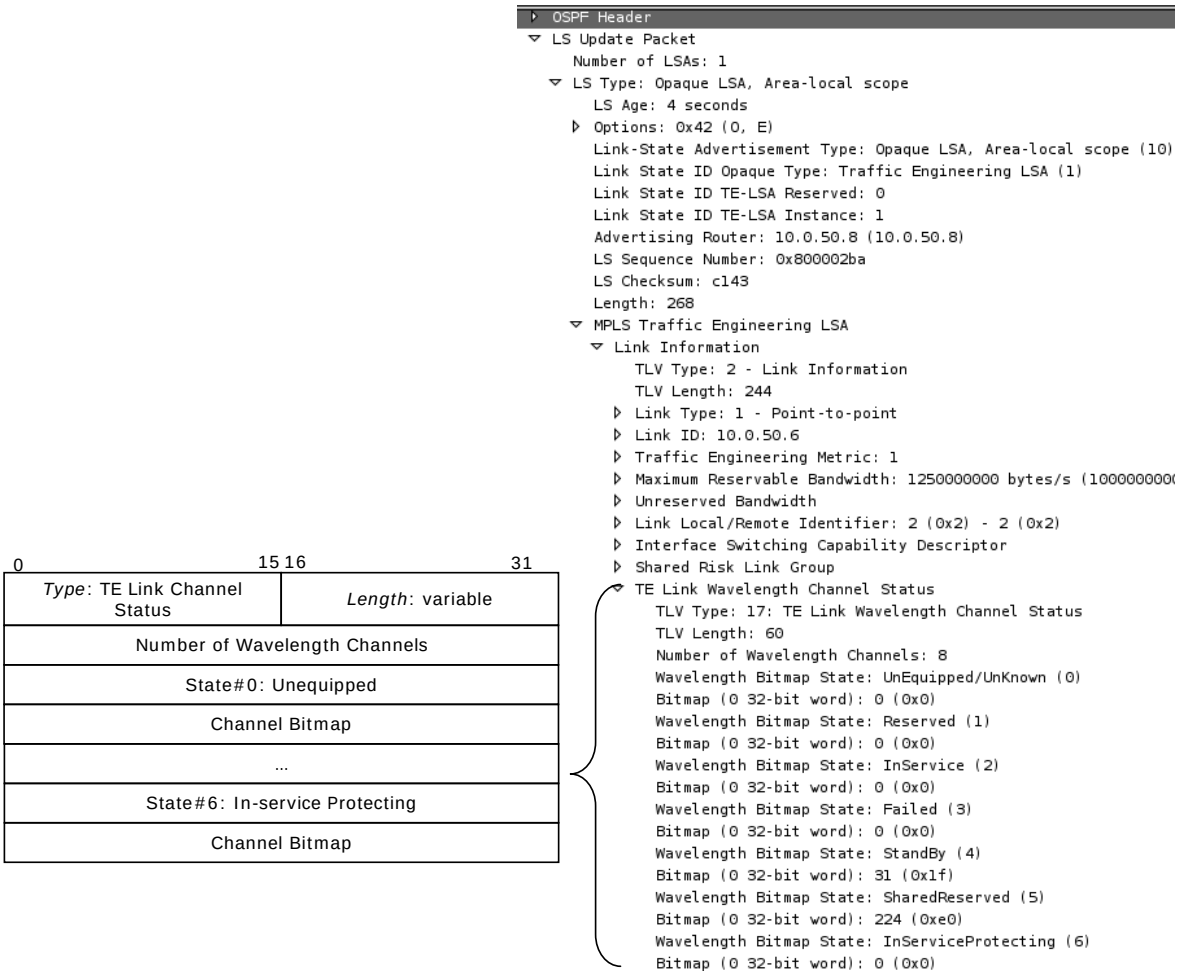

Fig. 5. OSPF-TE extensions for the dissemination of wavelength status (bitmap encoded) [15]

PCEP protocol and related extensions

The PCEP protocol [4], currently in the standardization process by the IETF, enables a Path Computation Client (PCC) or a PCE in another domain, to request Path Computation Services from the PCE. The PCEP uses a TCP connection and, after an initial session handshake (Fig. 6), the PCC may request a Path Computation using a Path Computation Request (PCReq) message. In the case of a successful path computation, the PCE replies with a Path Computation Reply (PCRep) including the path ERO and its attributes or a NoPath object otherwise. If either the PCC or the PCE does not desire to keep the connection open, ends up the PCEP session with a Close message.

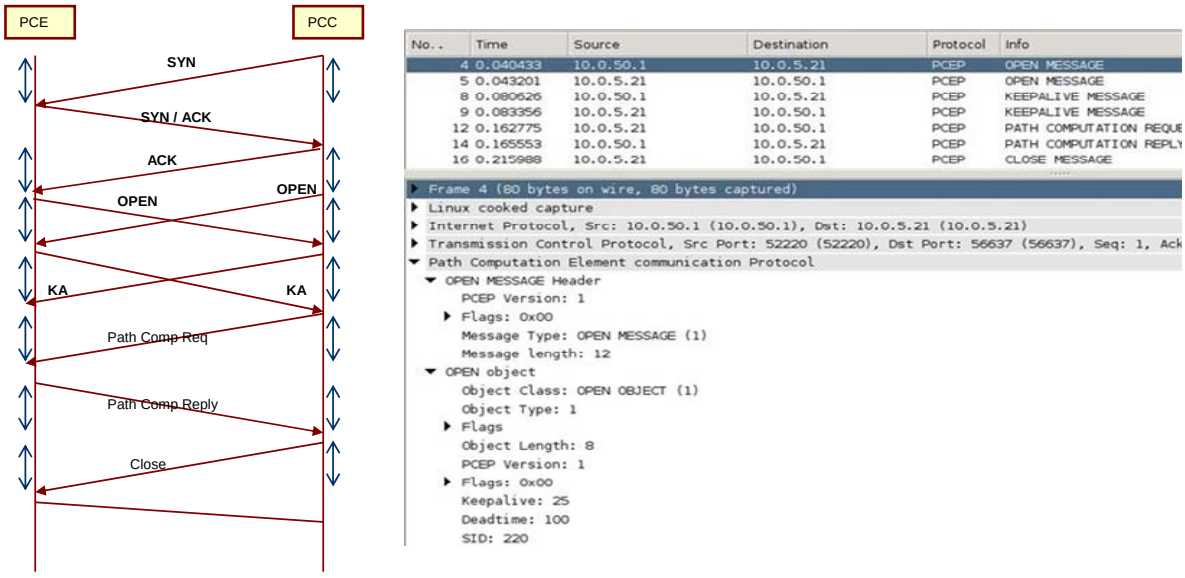

Fig. 6. PCEP handshake and Path Computation

The extensions to the PCEP protocol involve the following concepts: first, the addition of new Type Length Value (TLV) sub-objects to allow the request of optical lightpaths and the specification of the desired wavelength allocation policies loosely based on [17]. Second, the definition of new PCEP objects to convey, in the corresponding PCRep message, the ERO object as a path attribute and the set of available wavelengths for the path to guarantee the WCC. These objects are the Label Set, the suggested label and, for bidirectional connections, the upstream label. Fig. 7 and Fig. 8 show the format of the two new PCEP objects. For the time being, we only implemented the all-inclusive label set. The SL and the UL objects share the same object format. Additionally, Fig. 8 shows an example of the resulting PCEP message exchange, including a Path Computation Request and a Reply.

<attribute-list>::=[<LSPA>][<BANDWIDTH>][<metric-list>][<IRO>]**[<LABELSET>][<labels>]**
<labels>::=<suggested label>[<upstream label>]

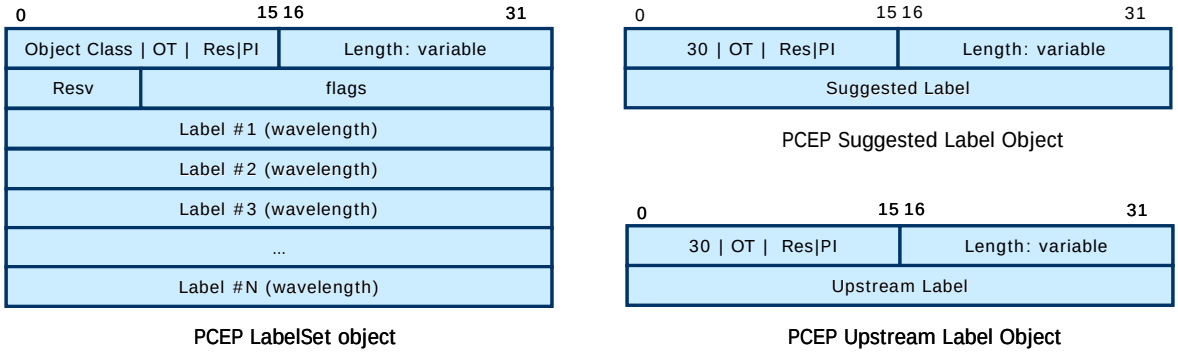

Fig. 7. PCEP objects for the LabelSet, Suggested Label and Upstream Label.

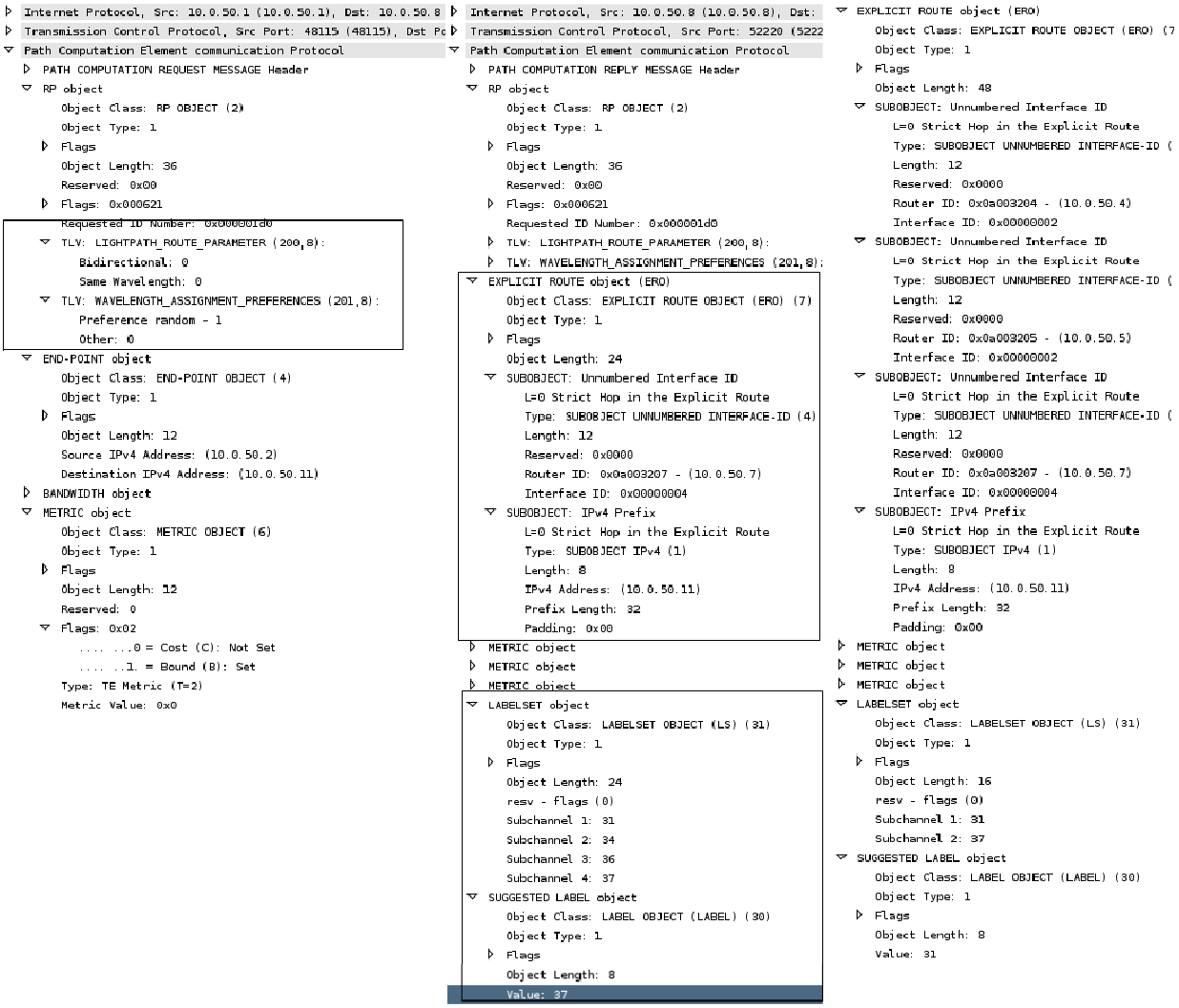

Fig. 8. Wireshark capture of the implemented extended PCEP and BRPC for WSON, showing the request for a Unidirectional unprotected lightpath with Random WA between nodes 10.0.50.2 and .11 (PCEs

are .1 and .8) and the reply with one path from ABR .7 (7-11, 4 free wavelengths) and from ABR .4 (4-5-7-11, 2 free wavelengths)

5. Experimental Multi-domain evaluation Scenario

5.A. ADRENALINE Testbed: Main Features

The ADRENALINE (All-optical Dynamic REliable Network hAndLINg IP/Ethernet Gigabit traffic with QoS) [16] test-bed is a GMPLS-based WSON network developed at CTTC premises. The ADRENALINE transport plane is composed of an all-optical Dense Wavelength Division Multiplexing mesh network with two colour-less Reconfigurable Optical Add Drop Multiplexer nodes and two Optical Cross Connect nodes, providing reconfigurable (in space and in frequency) end-to-end lightpaths, transparent to the format and payload of client signals. ADRENALINE deploys a total of 610 km of G.652 and G.655 optical fiber divided in 5 bidirectional links In these fibers optical amplifiers (Erbium-Doped Fiber Amplifiers or EDFAs) compensate power losses during optical transmission and switching at C-band. Each optical node is equipped with an Optical Connection Controller (OCC) for implementing a distributed GMPLS-based distributed control plane. From a research point of view, one of the focus and goals of the ADRENALINE testbed is the performance evaluation of GMPLS-based traffic engineering algorithms and schemes. For this purpose, we added a new set of 42 GMPLS-enabled controllers without associated optical hardware (i.e., the optical hardware is emulated). This set of GMPLS controllers introduces a new degree of flexibility in topology configuration, without restrictions regarding the targeted optical network topology or regarding the resources per link (e.g., number of available wavelengths, fibers, etc). Thus, the GMPLS controllers can be inter-connected following any devised topology, by means of Ethernet point-to-point channels

carried over emulated optical links. The proposed solution allows the specification of control link parameters for realistic QoS constraints (fixed and variable packet delays, packet losses, bandwidth limitations, etc.). In particular, and in order to provide a flexible framework for topology reconfiguration, the IP Control Channels (IPCC) in the DCN are implemented in terms of point-to-point IP interfaces with optional GRE or IPIP tunneling over Ethernet interfaces. For this purpose, it uses virtual local area networks (IEEE 802.1q VLANs), configured both in the layer 2 Ethernet switches and in the GMPLS-enabled controllers within the testbed. This approach enables the deployment of arbitrary layer 2 interconnections between network nodes absolutely decoupled of the physical infrastructure.

5.B. Network Topology

Fig. 9 shows the deployed network topology, consisting of 3 TE domains that correspond to OSPF-TE routing areas (areas 1, 0 and 2). There are 2 ABR between areas 1 and 0 (nodes 7 and 4) and 3 ABRs between areas 0 and 2 (nodes 9, 10, and 12). PCEs are located in nodes 1, 8 and 14. All control plane links have equal propagation delay set to 3ms. TE links have 8 wavelengths on each direction.

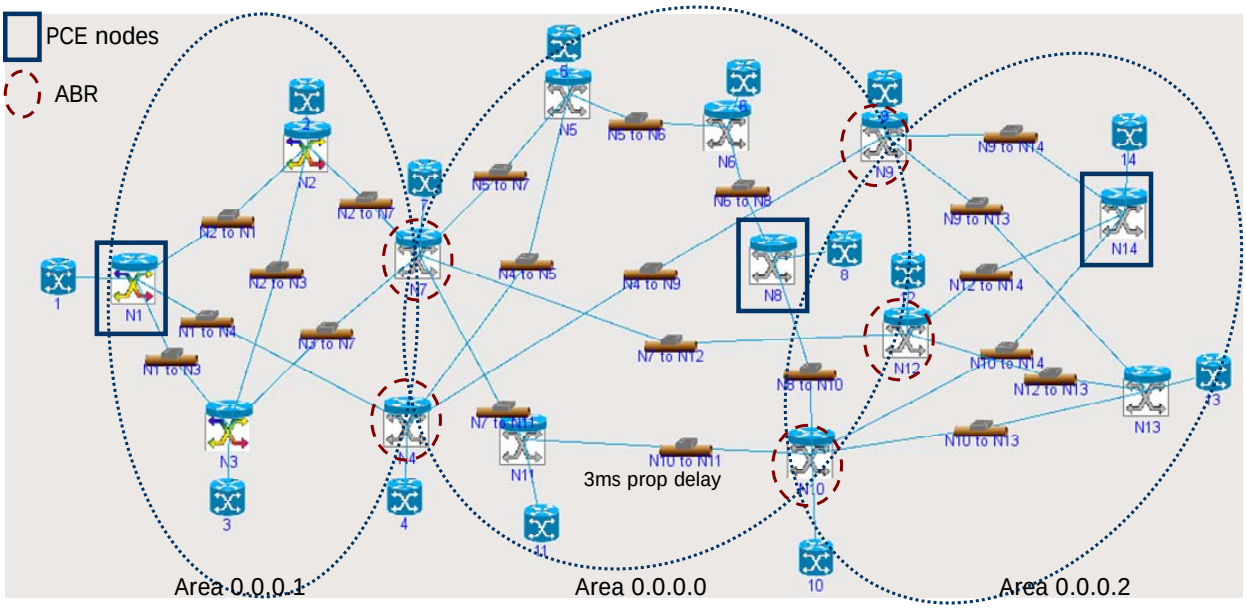

Fig. 9. Evaluated Network Topology, scenario containing 14 nodes in 3 areas.

6. Experimental Results

In order to evaluate the aforementioned solutions, we compute the blocking probability and setup delay at a range of offered traffic (8-72 Erlangs). Each key performance indicator value is computed requesting 10^4 lightpaths, with a negative exponential holding time of average 120 seconds (the inter arrival process in a Poisson process, with a rate depending on the offered load). The ingress and egress node pairs are chosen randomly; following a uniform distribution between all distinct node pairs that either: *a) belong to different areas or b) are a node and an ABR (e.g. pair 1-14 and pair 1-7 have equal probability).* Wavelength Allocation, either by the RC, the CC or the PCE is random.

6.A Blocking Probability

Fig. 11 plots the obtained blocking probability (BP) for the 3 considered cases, namely the per-domain path computation (PD), the standard BRPC (BRPC) and our conceived method (EBRPC-WSON). The PD approach, constrained within a domain to a source/ingress and/or destination/egress node, is only able to use the TE information and wavelength status to meet the WCC within a given domain, selecting, during the ERO expansion, longer paths if needed. However, the lack of visibility of other domains and the suboptimal choice of boundary nodes are the main causes of lower performance. BRPC leverages the multiplicity of ABRs between adjacent domains when compared to the PD method, showing an improvement of around 12% at 72 Erlangs of offered traffic (16,2% BP in BRPC, and 18,4 % in PD). However, it is still subject to failures given the difficulty to meet the WCC along the path since the wavelength status information is lost during successive transmissions of the VSPT to the upstream domain. In consequence, paths are selected based on the TE metric as long as there is, on a per link basis, enough unreserved bandwidth (i.e., at least one available wavelength). Although the PCE insures the selection of optimal paths within its own domain, upstream PCEs are not able to use the wavelength status information thus the network tends to use always the same feasible shortest paths with enough unreserved bandwidth. Crossing of one or two domain boundaries mitigates this effect in our network topology, but could show a performance decrease if the domain chain becomes longer.

The EBRCP-WSON significantly reduces the blocking probability, not only by finding the optimal path as BRPC does, but also by insuring that the conveyed label set and the preferred wavelengths (suggested label) are available end-to-end. It is thus likely to meet the WCC along the path barred routing with outdated information or contentions. At 72 Erlangs, the EBRPC

shows a 14.8% BP, almost a 20% improvement over PD, and around 8.6% improvement over the standard BRPC.

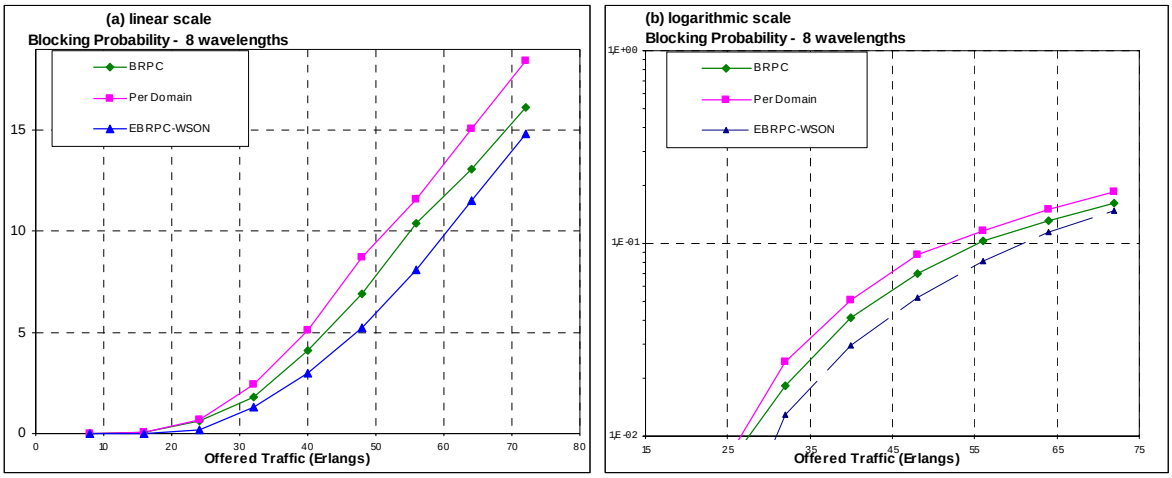

Fig. 10. Obtained blocking probability for the PD, BRPC and EBRPC-WSON

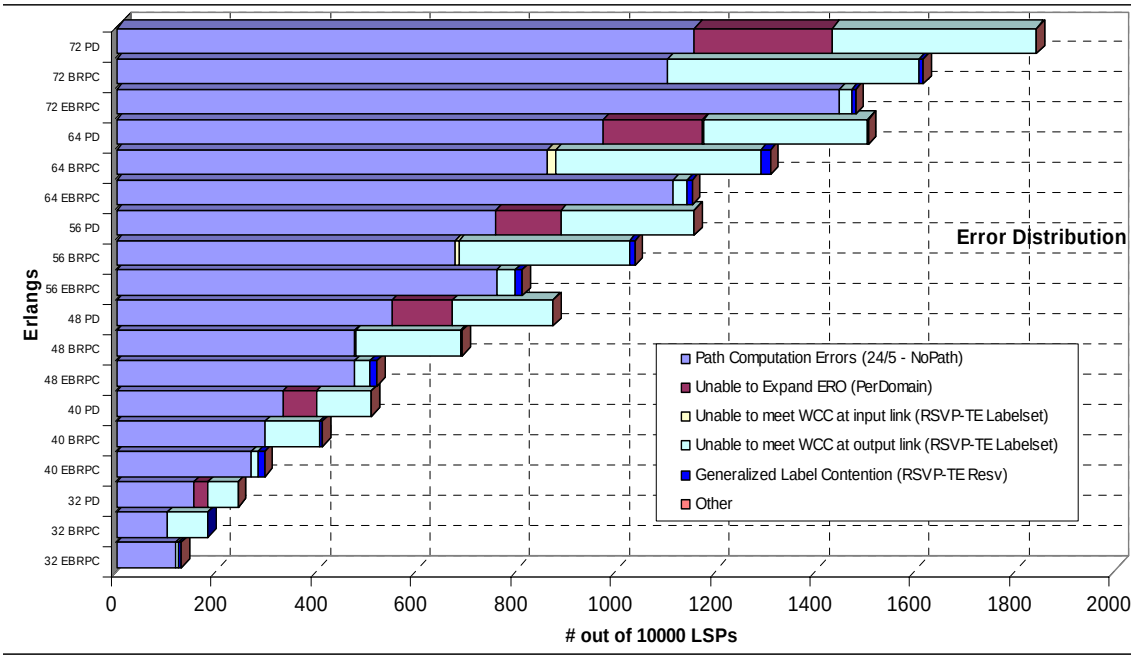

Fig. 11. Path Error decomposition for PD, BRPC, EBRPC-WSON, for 32-72 Erlangs

Fig. 11 shows the decomposition of the Path Errors by their nature. For the 3 approaches and for a traffic range (32-72 Erlangs) we plot the obtained the amount of errors in the 10^4 connections. We classified the errors as:

a) Path computation errors at the LSP source or PCE,

b) Failure to expand ERO expansions during signaling,

c) Failure to guarantee the WCC at the input port during signaling (i.e., the LS became empty),

d) Failure to guarantee the WCC at the output port, and

e) Error due to Generalized Label contention during Resv message processing.

The PD shows a proportion of Path Errors due to the Connection Controller being unable to perform ERO Expansion (no feasible path with continuous wavelengths from the domain ingress node to the domain egress node towards the destination). This cause is not present in the PCE based methods, since no signaling occurs until an end-to-end ERO is available. Considering EBRPC-WSON, there are almost no failures due to (c) and (d) with the exception of a (proportionally) low number of occurrences due to outdated information or contention, one of the benefits of this approach. On the contrary, for the BRPC and PD, many connections fail due to the lack of knowledge of end-to-end available wavelengths (e.g., around 1/3 and 1/4 respectively at 72 Erlangs).

Regarding the aforementioned outdated information, in general, the network performance shows a notable decrease when the inter-arrival times are below a given threshold, due to several factors. First, let us define OSPF-TE convergence time as the time passed between the moment of the generation of a new TE-LSA and its installation in the TE database of every OCC in the network or TE domain (including the PCE). For illustration purposes, considering a domain of

around 10 nodes, the flooding of a TE LSA can take up to ~1 second. Consequently, the higher the OSPF-TE flooding dissemination delay and the higher convergence time, the more likely a path computation will use outdated information.

Second, the PCE-based approaches are very sensible to traffic dynamics, especially when the deployment method involves a centralized PCE serving requests within a domain while obtaining TE data from OSPF-TE. Since the TE update flooding happens proportionally to the arrival and departure rate, the PCE is more likely to perform path computation based on outdated information. Moreover, although we do not have numerically evaluated it, we have observed that increasing the number of concurrent connections (proportional to the number of PCCs and especially at short timescales) will eventually force the PCE operating system kernel to drop or to delay PCEP sessions.

On the contrary, the per-domain approach seems slightly more robust to an increase of traffic dynamics, since segment path computations (e.g. ERO expansions) happen at the actual path setup in a distributed manner (ingress node within the domain).

6.B. Lightpath Setup Delay

The Lightpath Setup delay reflects the necessary time to establish a lightpath, and includes the path computation latency and the corresponding signaling delay. As expected, PCE-based path computation presents a higher setup delay than the per-domain approach (cfr. Fig. 12). The setup delay is one order of magnitude higher in PCE based solutions than in the PD solution. The former shows a constant average of about 190 ms while the latter shows a constant average of ~24 ms. In both cases, for the considered scenario and traffic properties, the CPU and memory of the OCCs and the PCE are not performance bottlenecks.

For PCE-based approaches, the path computation latency includes the TCP initial handshake (including the socket queuing for the “connect / accept” sequence in the PCE operating system), the PCEP handshake, and the actual path computation. Moreover, the latter involves subsequent requests to downstream domains and random delays to access the database, since OSPF-TE may concurrently update it.

Considering the worst-case scenario, in which TCP and PCEP handshakes happen on a per request basis, we conclude that, to a first approximation, the sum of round-trip propagation delays involved in the PCEP handshakes between all concerned PCEP speakers is which mostly determines the overall path computation latency.

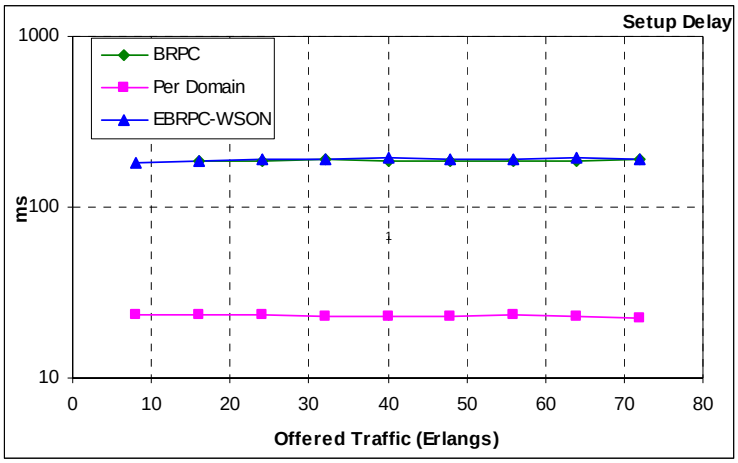

Fig. 12. Path Setup Delay (including Path Computation) in milliseconds (logarithmic scale)

Fig. 13 shows the histogram of both the per-domain approach and the EBRPC-WSON for 72 Erlangs. Fig. 13.a plots the per-domain histogram, and presents spikes that grossly correspond to the round-trip propagation delays (i.e., hop count for homogeneous links) between the source and destination. Fig. 13.b plots the EBRPC-WSON histogram, which presents two main components that correspond to computation involving two and three PCEs (i.e., for LSPs crossing, respectively, two and three areas).

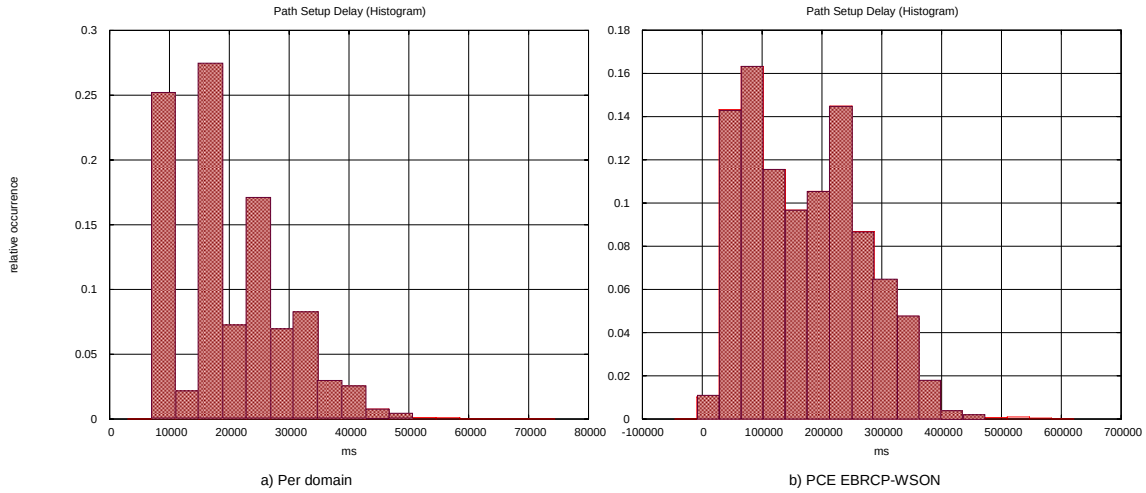

Fig. 13. Path Setup Delay histogram in ms for both the per-domain and PCE based approaches

The obtained results validate the applicability of the PCE although it implies a notable increase in LSP setup delay compared to the per-domain distributed source routing. Note that, in our case, the time of the actual execution of the algorithm in the PCE itself is in the order of a few milliseconds.

This increase may be relatively unimportant at low loads as traffic dynamicity is orders of magnitude slower than the OSPF-TE dissemination delay and related convergence times which corresponds to today's WSON deployment models. However, at higher traffic loads and with dynamic traffic conditions, the BRPC-based methods may suffer the effects of outdated information, since the optimal computed path may not remain optimal at the time of the actual LSP signaling. Nonetheless, note that it is possible to reduce the PCE-based computation latency (and in consequence the path setup delay) by means of persistent connections, in which the PCEP session is left in the UP state and the whole handshake is performed only once.

7. Conclusions

The Path Computation Element enables advanced path computation solutions for multi-domain traffic engineered Label Switched Paths. In particular, the PCE may play an important role in network architectures of the Future Internet context. We focused on the mechanisms for lightpath provisioning in multi-domain all-optical networks. We implemented, deployed and experimentally validated three different path computation algorithms: the per-domain, the PCE-based BRPC and our proposed PCE-based EBRPC approach that efficiently addresses the WCC. We quantitatively evaluated the key performance indicators such as the blocking probability and the setup delay. As expected, the per-domain method shows on average the smallest path setup delay, providing robustness in front of very high traffic dynamics. However, it also shows the highest blocking probability because it is constrained to a given entry and exit boundary nodes and thus it is unable to find the shortest feasible end-to-end path and because it has limited visibility to take into account the WCC during the ERO expansion. We have also shown that the well-known BRPC, conceived to allow the computation of an end-to-end (shortest) path in the presence of multiple ABR nodes in the network, fails to capture the WCC constraint present in all optical transparent networks since the information regarding wavelength availability is lost in the VSPT processing.

The devised extended BRPC, motivated by the specific requirements of WSON under WCC, minimizes blocking due to WCC while still computing optimal end-to-end paths, outperforming the other two without significantly impacting scalability (in terms of additional control plane required bandwidth or latency).

Both PCE-based approaches come at the cost of a higher path setup delay due to the increased path computation latency, and at the cost of additional path computation entities and control plane extensions. It is noteworthy that the proposed extension to BRPC is relatively

negligible in terms of path computation (the PCE executes the path computation algorithm in a few milliseconds), and only extends the PCEP Reply message marginally, having no noticeable impact on a dedicated control network of 100Mbps Ethernet-based control channels. Finally, let us point out that the approach only requires extending OSPF-TE with wavelength related information if this is the selected means to obtain wavelength status in the PCE deployment method.

Acknowledgements

The work in this paper has been partially funded by Spanish Ministry of Science and Innovation, via the RESPLANDOR project (TEC2006-12910/TCM), within the BONE (Building the Future Optical Network in Europe) project, a Network of Excellence funded by the European Commission through the 7th ICT-Framework Programme.

1. E. Mannie, Ed., “Generalized Multi-Protocol Label Switching (GMPLS) Architecture”, RFC 3945, Oct. 2004
2. K. Kompella and Y. Rekhter, “OSPF Extensions in Support of Generalized Multi-Protocol Label Switching”, IETF RFC 4203, Oct. 2005.
3. A. Farrel, A. Vasseur, J. Ash, “A path computation element (PCE) based architecture", RFC 4655, Aug. 2006.
4. J.-P. Vasseur, J.-L. Le Roux, “Path Computation Element (PCE) Communication Protocol (PCEP)", work in progress, I-D, draft-ietf-pce-pcep-15, Sept.2008.
5. R. Douville, et al., “A service plane over the PCE architecture for automatic multi-domain connection-oriented services,” IEEE Communication Magazine, vol. 46, no. 6, pp. 94 102, Jun. 2008.

6. J.-P. Vasseur et al., “A backward recursive PCE-based computation (BRPC) procedure to compute shortest constrained inter domain traffic engineering label switched paths", work in progress, I-D, draft-ietf-pce-brpc-09, April 2008.
7. R. Bradford, J.-P. Vasseur, A. Farrel, “Preserving topology confidentiality in inter domain path computation using a key based mechanism", work in progress, I-D, draft-ietf-pce-path-key-03, May 2008.
8. Sukrit Dasgupta, Jaudelice C. de Oliveira, Jean-Philippe Vasseur, “Path-Computation-Element-Based Architecture for Interdomain MPLS/GMPLS Traffic Engineering: Overview and Performance”, IEEE Network 30-45 (2007)
9. F. Paolucci, F. Cugini, L. Valcarenghi, P. Castoldi “Enhancing Backward Recursive PCE-based Computation (BRPC) for Inter-Domain Protected LSP Provisioning”, OFC/NFOEC 2008.
10. D. Awduche, et al., “RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, Dec. 2001.
11. A. Farrel, A. Ayyangar, JP. Vasseur, “Inter-Domain MPLS and GMPLS Traffic Engineering - Resource Reservation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 5151, Feb, 2008.
12. Berger, L., “GMPLS Signaling Procedure for Egress Control”, RFC 4003, February 2005.
13. R. Casellas, R. Martínez, R. Muñoz, “Design, implementation and validation within ADRENALINE® testbed of a Path Computation Element for Wavelength Switched Optical Networks”, in Proc. 4th International Conference on IP over Optical (iPOP2008), Tokyo (Japan), June 2008.

14. R. Casellas, R. Muñoz, R. Martínez, “A Path Computation Element for Shared Path Protection in GMPLS-enabled Wavelength Switched Optical Networks”, in Proc. 34th European Conference on Optical Communications, ECOC2008, Brussels (Belgium) September 20-25 2008.
15. R. Martínez, R. Casellas, R. Muñoz, Experimental evaluation of GMPLS enhanced routing for differentiated survivability in all-optical networks, OSA Journal of Optical Networking. Vol. 7, No.5, pp. 496-512, May 2008.
16. R. Munoz, C. Pinart, R. Martinez, J. Sorribes, G. Junyent, M. Maier, A. Amrani "The ADRENALINE Test Bed: Integrating GMPLS, XML and SNMP in transparent DWDM networks", IEEE Communications Magazine, N.8 V.43 40-48 (2005).
17. Y. Lee, et al., “PCEP Requirements and Extensions for WSON Routing and Wavelength Assignment”, work in progress, I-D, June 27, 2008