Experimental Validation of Active Frontend - Backend Stateful PCE Operations in Flexgrid Optical Network Re-optimization

R. Martínez(1), Ll. Gifre(2), R. Casellas(1), L. Velasco(2), R. Muñoz(1), R. Vilalta(1)

(1) CTTC, Av. Carl Friedrich Gauss 7, Castelldefels, 08860, Spain, rmartinez@cttc.es
(2) UPC, Optical Communications Group, Jordi Girona 1-3, 08034 Barcelona, Spain, lgifre@ac.upc.edu

Abstract *A PCE-based strategy for LSP re-optimization in flexgrid optical networks is experimentally validated. A front- and back-end stateful PCE architecture is deployed, wherein automatic and coordinated operations (e.g., computation delegation, database synchronization, etc.) are demonstrated using PCEP and BGP-LS protocols.*

Introduction

Flexgrid optical networks provide finer granularity than fixed grid attaining more efficient and flexible use of optical spectrum[1]. Optical spectrum is partitioned into fixed-size nominal central frequencies (NCFs), where Label Switched Paths (LSPs) are set up on a frequency slot (FS) formed by a variable number of contiguous NCFs depending on bit rate and modulation format.

For each LSP request, a centralized Path Computation Element (PCE) executes a Routing and Spectrum Allocation (RSA) algorithm[2] computing both the physical route (nodes and links) and the FS. The PCE requires an updated view of network state (i.e., topology and resources status) collected in the Traffic Engineering Database (TED). This information can be extended in the stateful PCE[3] with the knowledge of the established LSPs (LSP Database, LSPDB). The stateful PCE is thus able to reduce interactions and thus resource contention among concurrent signaled LSPs. The stateful PCE can also be active, to issue network recommendations to allow modifying and rerouting path reservations, for instance, for re-optimization purposes[4].

The operations done in the PCE may be computationally intensive when running the RSA calculation for LSP re-optimization. Specifically, re-optimization allows re-arranging active LSPs in a more efficient way in terms of network resources, especially required under dynamic traffic. To optimize this, the PCE architecture supports redundancy for load-sharing between two (or more) coordinated PCEs: front- and back-end PCEs[4,5,6] (fPCE / bPCE). We consider that fPCE steers the RSA computations for new and restored LSPs, whilst bPCE is dedicated to re-optimization of existing LSPs. This requires coordination and path delegation between both PCEs (achieved using the PCE protocol, PCEP) as well as having the same TED information, which can be automatically attained by the BGP-LS[7] protocol. In BGP-LS, BGP speakers are placed at each PCE, which perform the TED synchronization. Additionally, the LSPDB contents of both PCEs need also to be unique. To ensure this, the PCEP Report (PCRpt) messages are used.

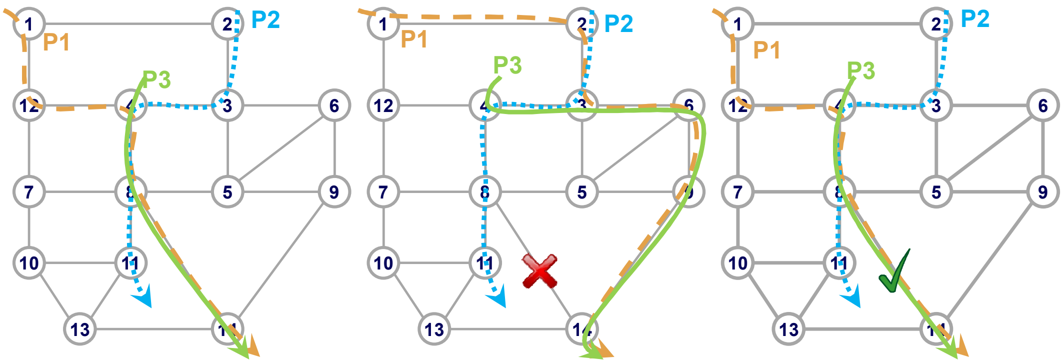
In this paper, we combine the above PCE capabilities (i.e., active stateful, fPCE-bPCE coordination and BGP-LS updating) for providing both PCE path computations for new and restored LSPs and effective LSP re-optimization. The PCE architecture and involved protocols to deal with such objectives are addressed and experimentally validated in a distributed infrastructure connecting CTTC's ADRENALINE testbed (including fPCE) and UPC's bPCE (named PLATON).

After Failure Repair re-Optimization (AFRO)

Figure 1 depicts the LSP re-optimization process when a link fails and afterwards is repaired. In Fig. 1a, three flexgrid optical LSPs (P1, P2 and P3) were computed by the fPCE and established (using RSVP-TE signaling): P1 (path 1-12-4-8-14), P2 (path 2-3-4-8-11) and P3 (path 4-8-14).

Let's assume that link 8-14 fails (Fig. 1.b). This affects P1 and P2 LSPs, which need to be restored. Each failed link is flushed from both PCCs' and fPCE's TED for the subsequent path computations. Nevertheless, bPCE's TED requires a dedicated updating, which herein is done using the BGP-LS protocol. Indeed, the fPCE speaker sends a BGP Update (BGPUpd) message to its peer bPCE to inform that the failed link is unavailable in the bPCE's TED.

For the LSP restoration, the ingress LSP controllers (Path Computation Clients, PCCs) request a new path computation to the fPCE via the PCEP Request (PCReq) message. Like regular LSP provisioning, restoration paths are computed using a RSA algorithm which computes the shortest path ensuring sufficient continuous and available spectrum on each link for the FS and avoiding failed links. The stateful PCE capability avoids assigning the same FS to



a) Network Before Failure b) Network After Failure & Restoration c) Network After Link Repair & Re-Optimization

Fig. 1: Example of LSP re-optimization after link failure repair

a received bulk of backup path computations. This does lower potential resource contention when signaling LSPs.

In the example (Fig. 1 b), computed backup P1 and P3 LSPs result larger routes (P1: 1-2-3-6-9-14 and P3: 4-3-6-9-14,), i.e. more hops, than their failed working paths.

Once LSPs are restored, the fPCE updates the bPCE's LSPDB using the PCRpt message. The fPCE/ bPCE protocol exchange for TEDs and LSPDBs synchronization are described in Fig. 2.. Nevertheless, to update bPCE's TED, it is required that fPCE BGP-LS speaker sends a BGP Update (BGPUpd) message to its peer bPCE. Such a message informs that the failed link is unreachable in the bPCE's TED.

After a failed TE link is repaired, the fPCE's TED is again updated. The fPCE performs two tasks: first, it sends a BGPUpd to the bPCE to notify that a link is now reachable (step 1 in Fig. 2); second, the fPCE seeks for a candidate set of LSPs eligible to be re-optimized (step 2).

In this work, the re-optimization of all the existing LSPs is requested. This is done by a PCReq message including a Global Concurrent Optimization (GCO) request. The PCReq includes a list of LSPs in a Synchronization VECtor (SVEC) object. For each path computation request in the SVEC, there is a Request Parameter (RP) object. The RP carries a Request-ID (32 bit integer) and the symbolic path name of the LSP within the LSPDB as a TLV. This symbolic name is always passed from fPCE to bPCE when a new LSP has been set up and identifies each LSP. Consequently, the bPCE is able to unambiguously identify the LSPs to be re-optimized using both the RP and the LSPDB.

Next, the bPCE runs the re-optimization algorithm for each LSP (step 3). The applied algorithm is the AFRO mechanism[5], which re-routes the LSPs on shorter routes to enhance the spectrum efficiency. The path computation result is then sent to the fPCE via a PCEP Response (PCRep) message (step 4). In this message, for each re-optimized LSP an Explicit Route Object (ERO) is carried with the route and assigned FS.

At the fPCE, the new received ERO for each LSP is compared to the one stored in the LSPDB. If an ERO update is detected, the fPCE sends a PCEP Update (PCUpd) message to the corresponding PCC (step 5). This indicates the PCC to re-establish the LSP using the new computed ERO. Finally, the fPCE sends PCRpt messages to the bPCE to actually update the LSPDB repository with the current LSP information. The complete re-optimization is summarized in the workflow of Fig.3.

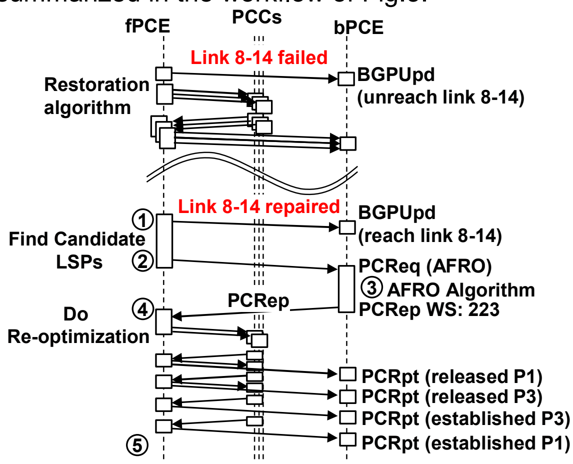


Fig. 2 fPCE/bPCE: PCEP messages and TED/LSPDB synchronization for restoration / re-optimization sequence.

Fig. 1.c depicts the re-optimization of P1 and P3

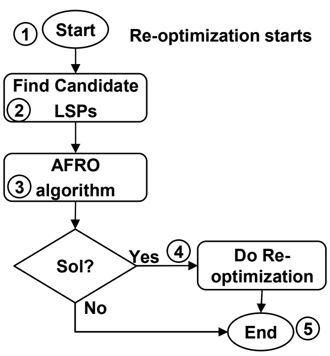


Fig. 3 Workflow for the re-optimization process

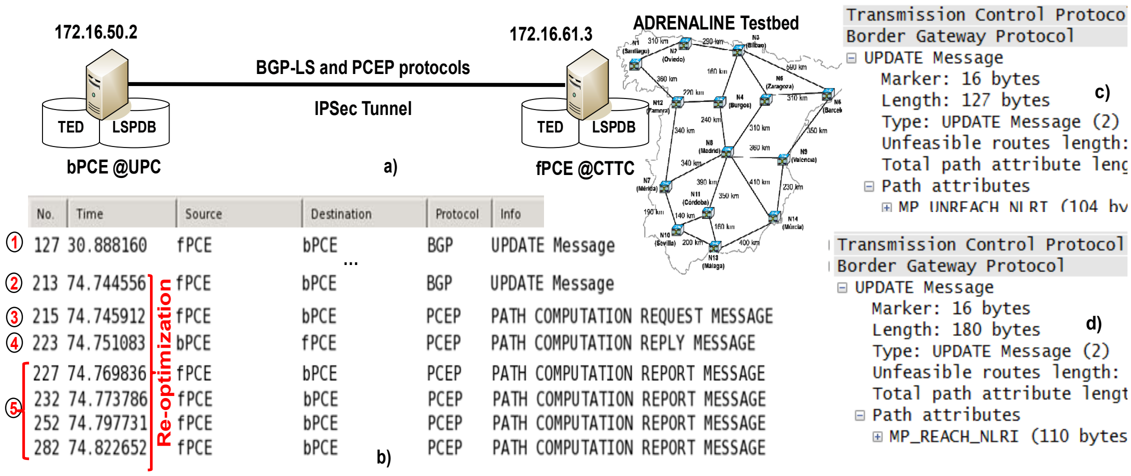


Fig. 4 a) GMPLS controllers at ADRENALINE testbed and fPCE (CTTC) – PLATON bPCE (UPC) peer connection; b) BGP-LS and PCEP exchanged messages; c) BGP-LS Upd message (link down); d) BGP-LS Upd message (link up)

LSPs. Observe that the re-optimized routes match with the computed routes before the link failed. Indeed, the re-optimization is devised to re-route LSPs, accommodating the paths on shorter routes than those obtained during the restoration process. Consequently, this leads to achieve better spectrum efficiency.

Experimental validation

The experimental validation (Fig. 4a) is carried out in the GMPLS flexgrid optical network of ADRENALINE testbed with a peer connection (IPSEC tunnel) between a stateful fPCE (at CTTC) and PLATON bPCE (at UPC). The fPCE is co-located with a GMPLS controller to capture exchanged OSPF-TE messages and construct the TED.

Fig. 4b shows relevant BGP-LS and PCEP messages exchanged between the fPCE (IP: 172.16.61.13) and bPCE (IP: 172.16.50.2) during the LSP re-optimization. To do so, we reproduce the network and re-optimization scenario described in Fig. 1.

BGPUpd message labeled by (1) in Fig 4 b is sent by the fPCE to bPCE to notify that link 8-14 failed. The message (Fig. 4c) allows bPCE to set the link to unreachable in its TED.

After a link is repaired and fPCE realizes about that, such a PCE sends a BGPUpd message, labeled by (2) in Fig. 4d, to the bPCE setting the link to reachable. As mentioned, link repair event triggers the re-optimization process. The fPCE computes the set of candidate LSPs to be re-optimized which are forwarded to the bPCE (PCReq message, labelled by (3)). When the bPCE receives this message, the AFRO algorithm is run. The solution is sent back to the fPCE in the PCRep message (labelled by (4)).

Upon receiving the PCRep message, the fPCE updates (re-routes) the LSPs in the network. Finally, when all re-optimized LSPs are re-established, fPCE sends PCRpt messages (labelled by (5)) to the bPCE to update the LSPDB. In the example, the re-optimization time computed between the reception of the BGPUpd message at bPCE (for link reachable) and the last PCRpt message takes around 78 ms.

Conclusions

LSP re-optimization, triggered e.g. by a link repair, is an intensive and complex path computation process leading to improve resource utilization under dynamic traffic. To deal with this, an appealing strategy leverages the load-sharing performed by two coordinated PCEs (namely, fPCE and bPCE): fPCE delegates the LSP re-optimization path computation to a dedicated bPCE. This work experimentally validates such a PCE architecture deploying fPCE (at CTTC) and bPCE (at UPC) to re-optimize flexgrid LSPs. The required set of automatic operations between both PCEs (i.e., path computation delegation, algorithm execution, TED synchronization and LSP reporting) has been successfully demonstrated using the PCEP and BGP-LS protocols interactions.

Acknowledgements

This work was partially funded by the EC's FP7 (IP IDEALIST project – 317999), and by Spanish MINECO (FARO and ELASTIC projects, TEC2012-38119 and TEC-2011-27310, resp.).

References

[1] M. Jinno, et al., "Spectrum-efficient and scalable elastic optical path network: architecture, benefits, and enabling technologies", IEEE Com. Mag, 47, 2009.

[2] L. Velasco et al., "Solving Routing and Spectrum Allocation Related Optimization Problems: from Off-Line to In-Operation Flexgrid Network Planning," IEEE/OSA JLT, 2014.

[3] E. Crabbe, IETF draft-ietf-pce-stateful-pce, 2013.

[4] R. Muñoz, et al. "PCE: what is It, how does it work and what are its limitations? IEEE/OSA JLT, 32, 2014.

[5] L. Velasco et al., "First Experimental Demonstration of ABNO-driven In-Operation Flexgrid Network Re-Optimization," PDP in IEEE/OSA OFC, 2014

[6] R. Casellas et al., "Applications and Status of Path Computation Element", IEEE/OSA JOCN, 5, 2013.

[7] H. Gredler, IETF draft-ietf-idr-ls-distribution, 2013