

SDN-based Service Orchestration of OpenFlow and GMPLS flexi-grid networks with a stateful hierarchical PCE

R. Casellas(1), R. Muñoz(1), R. Martínez(1), R. Vilalta(1), L. Liu(2), T. Tsuritani(3), I. Morita(3),
V. López(4), O. González de Dios (4), J.P. Fernández-Palacios(4)

(1) CTTC, Optical Networking Area, Av. Carl Friedrich Gauss n7 Barcelona, Spain, ramon.casellas@cttc.es

(2) University of California, Davis. One Shields Ave, Davis, CA 95616, USA

(3)KDDI R&D Labs, Fujimino-shi, Saitama 356-8502, Japan

(4) Telefónica I+D, GTCO, Don Ramon de la Cruz 82-84, Madrid, Spain

Abstract: We report the experimental evaluation of the end to end connectivity provisioning in which OpenFlow and GMPLS control planes interwork by means of a stateful PCE that coordinates the provisioning. The hierarchical nature of the PCE allows seamless multi-domain operation.

OCIS codes: (060.4250); (060.4251).

1. Introduction

Cloud applications such as data backup or virtual machine (VM) migrations involve an increasing amount of traffic between data centers (DC), and require efficient network architectures in terms of cost, energy consumption and reliability for end-to-end connectivity provisioning, combining flexible, fine-grained and adaptive traffic control in a very dynamic context regarding forwarding entries and policies (intra-DC) with long haul, potentially multi-carrier aggregated transport (inter-DC). For the former, the application of Software Defined Networking (SDN) fulfills such requirements whereas, for the latter, a GMPLS/PCE control plane, with its maturity and multi-domain support, accounts for existing deployments and return on investment. Consequently, the interworking between GMPLS and SDN/OpenFlow needs to be solved. In [1], a GMPLS domain was modelled as an OpenFlow node. More generically, such interworking can be seen as the orchestration (i.e., the automated configuration, coordination and management of complex systems, commonly with transactional semantics with rollback capabilities) of heterogeneous control plane technologies, and it is becoming an application domain of SDN. The use of SDN in optical transport networks is worth investigating, not because of the flexible, dynamic and decoupled centralized control (it is a common model of operation, given the specifics of the optical technology such as the wavelength continuity constraint or the need to account for physical impairments), but because of the potential simplification and better integration with operation and business support systems by means of open and standard interfaces and the use of existing functional entities.

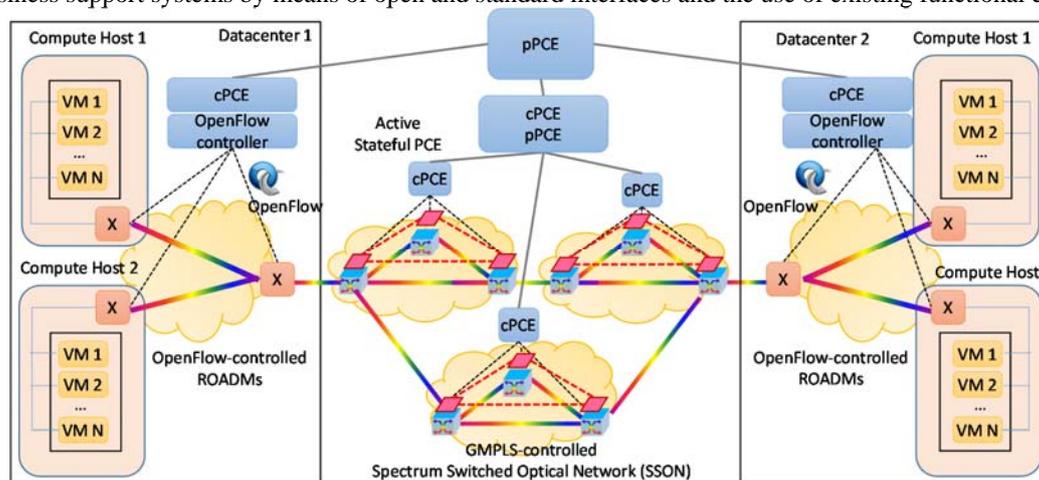


Fig. 1 DC interconnection as use case for the applicability of stateful H-PCE as orchestration in heterogeneous (OpenFlow/GMPLS) domains.

2. SDN based orchestration of Openflow/GMPLS domains

In this paper, we focus on such *control orchestration*, defining a SDN architecture and interfaces that addresses the considered DC-interconnection and GMPLS/OpenFlow interworking. For this, a centralized “controller of controllers or orchestrator” (as opposed to a mesh of controllers) handles the automation of business-level decisions, working at a higher, abstracted level and covering inter-domain aspects. Specific per-domain controllers map the abstracted control plane functions into the specifics of the underlying control plane technology, such as network resource discovery, topology management, connectivity provisioning and monitoring. Our approach relies on a set of stateful PCEs [2] arranged in a hierarchical PCE (H-PCE) manner [3]. The H-PCE architecture, where a single parent PCE (pPCE) is responsible for domain selection, using aggregated domains topology, and child PCEs are responsible for segment expansion (i.e., computation) in their respective domains, has been proposed for the purposes of path-computation in multi-domain optical networks [4], as well as multi-layer networks [5].

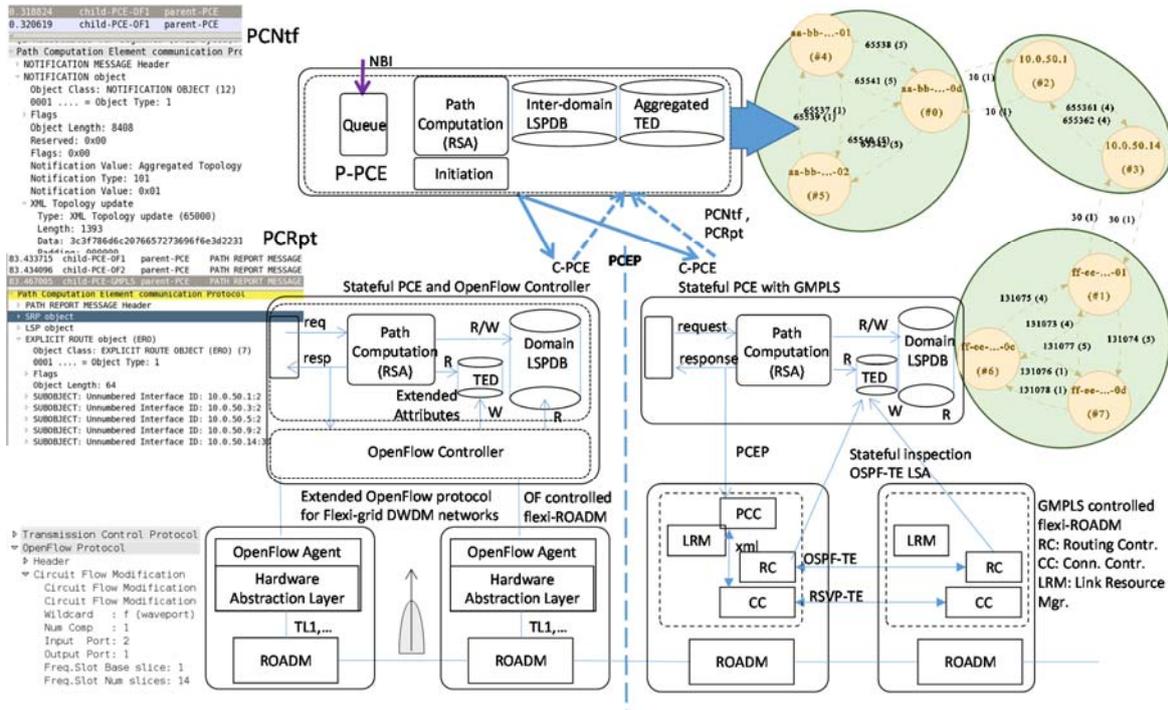


Fig. 2 Block diagram of system. The parent PCE performs orchestration, domain selection and requests segment expansion and instantiation. Children announce reachability and logical intra-domain links using PCEP notifications allowing an aggregated TED using generic identifiers.

3. Network reference model and control plane architecture

Our main contribution (Fig. 2) combines and extends the H-PCE architecture with stateful and OpenFlow [6]-related capabilities. A stateful PCE [7] (sPCE) is a PCE that is able to consider the network status in terms of links and nodes (the Traffic Engineering Database or TED) and the status of active connections (the database of Label Switched Paths or LSPDB). A sPCE is said to be *active* if it is able to recommend or modify/affect the state of existing connections and have *instantiation capabilities* [8], given the ability to trigger, upon request or autonomously, the establishment of connections. Thus, the pPCE orchestrates the provisioning of services with generalized identifiers (e.g. covering both GMPLS and OpenFlow) and each cPCE acts as a middleware for each domain control layer, either in tight cooperation with an OpenFlow controller [9] e.g., within the DC in a single domain with a large number of short lived flows requiring multiple dynamic routing rules and low latency, or delegating the actual establishment and release of connections to an underlying GMPLS/PCE control plane [10], e.g., for the core optical network, with typically tens of long haul, high capacity (beyond 100 Gbps) and low latency traffic, backwards compatible with carriers that deploy GMPLS-based transport services that upgrade their network to flexi-grid. All cPCEs have PCEP as their Northbound interface (NBI). PCEP is also the NBI for a GMPLS controlled edge node path computation client (PCC) module, PCEP is used for provisioning, rerouting, delegating and reporting. OpenFlow is used as Southbound interface (SBI) for cPCEs in such domains. The overall architecture is implemented and deployed in a lab-trial.

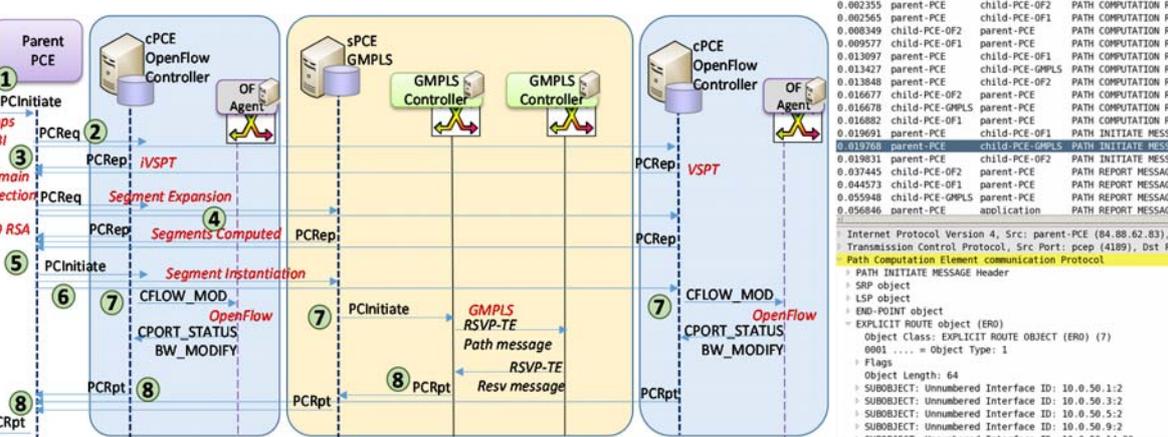


Fig. 3 Message flow for the establishment of an end-to-end connection. Wireshark capture showing a given instance with details regarding the GMPLS segment. The PCinitiate message contains the SRP, LSP, ENDPOINT and ERO objects [8]

4. Proposed Control Plane Procedures and Protocol Extensions

Fig.3 shows the main procedure for setup. Following a NBI request (1), the pPCE requests the virtual shortest path trees (VSPT) from the source and destination nodes to their border nodes (2) and uses the VPSTs and virtual links to perform domain selection (3) requesting the segment expansion (4). Each cPCE performs RSA and returns usable path and frequency slots (4). The pPCE combines the segments into an end-to-end ERO (5) and requests the establishment of the segments with PCinitiate messages (6). OpenFlow domains use a modified OpenFlow to program the ROADMs and GMPLS domains delegate the establishment to the underlying control plane (7). Finally, PCRpt messages communicate establishment (8) within the GMPLS domain, cPCE to pPCE and pPCE to NBI. PCEP has been extended for initiation, delegation and topology discovery of GMPLS and OpenFlow networks. The ENDPOINTS object supports both numbered and unnumbered interfaces and OpenFlow datapath ids. Explicit Route Objects (EROs) defined for the RSVP-TE protocol and used within PCEP to convey path routes can now include sub-objects that refer to OpenFlow entities (nodes as datapath ids and interfaces). PCEP has also been extended to support Routing and Spectrum Assignment (RSA) to convey selected modulation format, FEC and frequency slots and to announce reachability and aggregated topology dissemination, notably in the cPCE → pPCE direction, the latter using XML encoding (alternatively, BGP-LS can be also deployed for such purpose [11]). GMPLS OSPF-TE and RSVP-TE protocols have been extended with a new switching capability that corresponds to the ITU-T media layer for flexi-grid DWDM. This involves disseminating the status of nominal central frequencies (referred to as slices) and the support for the signaling of media channels with variable frequency-slot sizes. Finally, the OpenFlow protocol circuit switching addendum was extended to support capability discovery, the configuration of cross-connections and the reporting of the network topology.

5. Experimental Performance Evaluation

A testbed has been deployed (w/ emulated optical HW) with three domains: source and destination domains are controlled by an integrated cPCE/OpenFlow controller [9]. The transport domain is controlled by GMPLS/PCE, communicating with the PCC located at each GMPLS controller. A cPCE is located at each domain. Topology PCNtf with XML encoded virtual link mesh is announced to the pPCE every 60s. Each domain has 14 nodes/ROADMs. Flexi-grid TE links are assumed to have 128 usable nominal central frequencies. Given the inherent parallelization of OpenFlow, Setup Delays (between sending a PCinitiate and receiving a PCRpt) are mainly determined by H-PCE latency (computation and instantiation) and setup delay within the GMPLS domain which emulates a nationwide Japanese topology (Fig. 4, left). Since the OF domains are attached to a pair of nodes, setup delay within the GMPLS is quite constant and ~30ms as seen by the cPCE.

To evaluate the control plane overhead, we setup and release 100 connections with random source and destination nodes in remote OpenFlow domains, with 1s between operations. Test lasts 229 seconds, with request 1 starting at T=22s. The pPCE processes 4835 TCP segments (1584502 bytes, Tx=2486 segments, Rx=2369, avg. packet size is 326.35 bytes, with 16.7 packets/s), corresponding to 2135 PCEP messages (Open, KeepAlives, PCRreq, PCRrep, PCinitiate, PCRpt and PCNtf topology notifications). 51% of PCEP messages have 80-159 bytes, and 42% 640-1279. PCRpt and PCNtf messages are large, encoding available central frequencies and topology info.

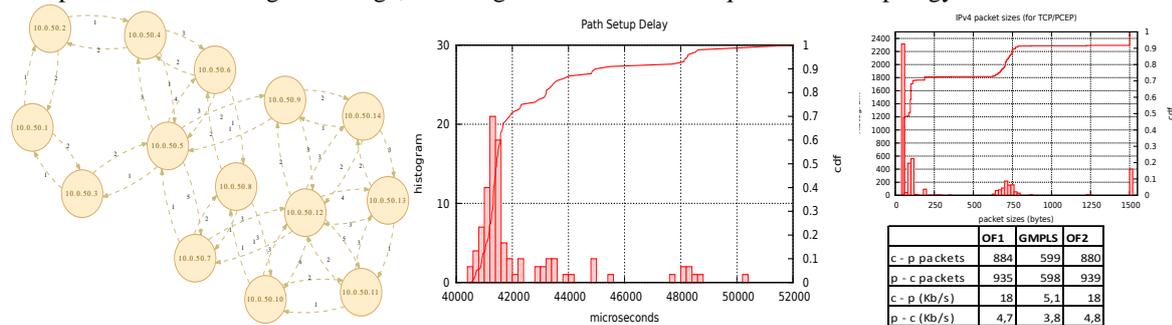


Fig. 4 GMPLS/PCE transport network topology (left). Setup delay histogram (center). PCEP control plane overhead (right) for 100 reqs

6. Conclusions

We have designed and validated a SDN orchestration mechanism that is based on a hierarchical stateful PCE extended to support OpenFlow. This solution relies on a mature, stable and feature complete PCEP protocol with extensions for topology management and discovery as well as for instantiation capabilities. The testbed deployment assesses its applicability for optical networks and is an industry-ready solution to the GMPLS/OpenFlow interworking.

7. Acknowledgments

Spanish MINECO project FARO (TEC2012-38119), and by the EU-Japan FP7 project STRAUSS.

8. References

- [1] S. Azodolmolky Opt Exp V19 N26, 2011
- [2] IETF RFC4655, PCE Architecture, 2006
- [3] IETF RFC6805, H-PCE, 2012
- [4] R. Casellas, Opt. Expr V19 N26, 2011
- [5] R. Casellas, Opt. Expr V20 N28, 2012
- [6] ONF <https://www.opennetworking.org>
- [7] Ed. Crabbe, draft-ietf-pce-stateful, 2013
- [8] Ed. Crabbe, draft-crabbe-pce-pce-initiated-lsp, 2013
- [9] R. Casellas, JOCN V5 N10 Oct13
- [10] R. Casellas, P.5.3 ECOC2013
- [11] M. Cuasmas, We.4.E.3, ECOC2013