Coordinated Computation and Setup of Multi-layer Paths via Inter-layer PCE Communication: Standards, Interoperability and Deployment

Oscar Gonzalez de Dios, Víctor López, Marta Cuaresma, Fernando Muñoz, Telefónica I+D, Spain
Mohit Chamania and Admela Jukan, Technische Universität Carolo-Wilhelmina zu Braunschweig

ABSTRACT

The IETF Path Computation Element (PCE) framework is a prime example of a software-defined network virtualisation artifact, originally envisaged about a decade ago for constraint-based routing in the conventional best-effort Internet. Over the years, however, it evolved into the ultimate tool for multi-technology, multi-domain, multi-layer, and multi-vendor networking. Carriers have already started to implement PCE-based solutions in their combined Internet IP/MPLS and Wavelength Switched Optical Networks (WSONs). However, significant development and testing is still needed to elevate the PCE concepts from theory to practice. This paper addresses standards, interoperability and deployment of coordinated computation and setup of multi-layer paths from a practical perspective and presents its first implementation in a multi-vendor, interoperable PCE testbed. In our testbed, we try to push the limits of multi-vendor interoperability, which current research mostly neglects, and thereby identify missing links in the current research and standards necessary to implement and deploy inter-operable multi-layer systems. We point to the key areas that the research and industrial communities will need to address in making the proposed solutions practical, which in turn will drive the upcoming efforts in software-defined networking with inter-operable PCE-based architectures.

INTRODUCTION

Software-defined networking has become a synonym for networking innovation, embodying a large spectrum of new opportunities and challenges, especially from an operational and control perspective. Among many of the software-defined networking efforts, the IETF Path Computation Element (PCE) [1] stands out as an imminent solution for a variety of technologies where deterministic path computation and traffic engineering is required, including core packet networks, mobile backhaul networks, optical networks, etc. PCE development is an integral part of the work of various standards bodies, not only in IETF but also ETSI and ITU-T, and thus has global scope and widespread acceptance. In regard to PCE design, deployment and evolution, research and industrial communities are today in the process of transitioning PCE from a software-defined concept to an interoperable networking standard. This is particularly the case in the recent IETF initiative on Application-based Network Operations (ABNO) [2], and work in the European Telecommunications Standards Institute (ETSI) on Network Functions Virtualisation (NFV) [3]. In addition to investigating ways to integrate PCE into advanced optical and carrier-grade Ethernet networks, network operators do not shy away for suggesting PCE based architectures for sensor networks as well as for the Internet of Things (IoT) frameworks.

As the PCE framework is maturing as an industry standard, practical and inter-operable PCE solutions are coming of age in combined packet and circuit-switched networks, i.e., a packet-switched IP/MPLS network and a circuit-switched high-bandwidth Wavelength Switched Optical Network (WSON). The coordination of these two networks via PCE is of particular interest due to a profoundly complementary they typically play in network service provision. The IP/MPLS infrastructure is used to support best-effort Internet traffic as well as various guaranteed services, such as Virtual Private Networks (VPNs) and leased lines. The WSON network, on the other hand, is used primarily to provide interconnectivity between different IP/MPLS routers, and only some specialized services with high bandwidth requirements are provisioned directly on WSON infrastructure. Combined however, these two networks can allow for smart network planning features in presence of dynamic traffic, as well as an efficient usage of the entire network infrastructure with high degree of reliability and availability of network services. To this end, the IETF Path Computation Element (PCE) architecture has been proposed to include mechanisms where multiple PCEs can communicate via the Path Computation Element Protocol (PCEP) [4] to enable complex path computations in multi-layer networks [5].

However, significant development, and testing, is ahead of us for the current concepts for multi-layer PCE solutions top advance from theory to practice. It is therefore no surprise that most telecom operators continue to use a rather manual procedure to provision multi-layer connections. This is illustrated in Fig. 1. To setup a VPN service, the Internet operator first attempts to create the service in the IP/MPLS network (1). If sufficient capacity is not available, the operators of the IP/MPLS and WSON networks communicate with each other (typically over emails, or live conversations) to determine the optimal placement for circuits in the WSON network as well as the configuration information required to create a corresponding link in the IP/MPLS network (2). Once this is finalized, the WSON operator provisions the optical circuit (3), after which the IP/MPLS operator performs configurations to setup the requested service (4). The rationale behind the "usefulness" of a manual workflow can be attributed to a few key factors. First, most vendors, especially in the optical network, employ either proprietary protocols or customized extensions to standardized protocols to exchange device specific parameters, which makes multi-vendor interoperability impossible. On the other hand, operators typically shy away from any vendor dependences, and prefer multi-vendor solutions, especially in the case of multi-layer networks. The latter is rather important to note, since "packet" and "transport" departments are traditionally structured as independent organizational entities inside a telecom operator. In other words, due to the significant differences in operations, combined with technology differences in the IP and the optical layers, automated provisioning of multi-layer connections faces challenges which are structural, strategic and technological in nature, all at the same time.

In this article, our goal is to understand whether the current PCE standards can aid or partially replace the manual workflow procedures in multi-layer networks, while respecting the network operator's desire to put a premium on multi-vendor interoperability, administrative separation of the networks, and standards-based solutions.[1] To this end, we implement a multi-layer PCE testbed capable of path computation and provisioning in a network equipped with off-the-shelf IP routers and optical switches. In addition to standardized PCE solutions, the proposed testbed uniquely implements two key management entities, namely the IP Network Control and Management system (NCM), as well as the Virtual Network Topology Manager

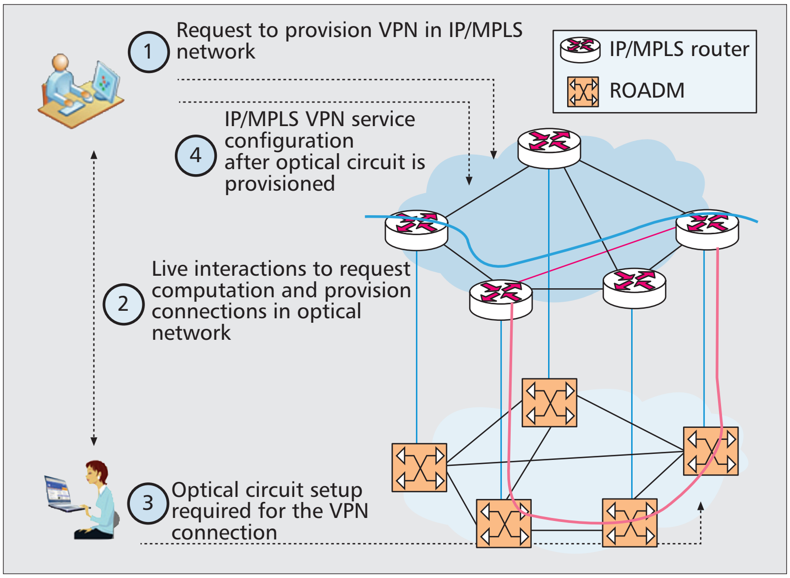

Figure 1. *A typical workflow for multi-layer service provision today.*

(VNTM). Our experiments show that a number of practical requirements have not yet been completely addressed by the standards, as well as a number of open research issues exist critical to the future deployment in commercial networks. These issues include mechanisms for topology discovery, TE information exchange, standardized VNTM interfaces as well as timescales for path computation and provisioning. The experiments show considerable relevance of these issues for upcoming software-defined networking concepts supported through inter-operable PCE-based architectures.

The rest of the article is organized as follows. We present the motivation and rationale. The testbed design, implementation and measurements are described. We discuss open issues and upcoming standards, while we conclude the article.

Motivation and Rationale

The PCE architecture consists of two distinct components, namely the Path Computation Element (PCE) and Path Computation Client (PCC). The PCC may be a network device or application that needs to know how to route traffic or manage resources, or it may be a user application with a graphical user interface (GUI) or command line interface (CLI) to request a path from source to destination with desired QoS metrics from the PCE server. A PCE server implements the capability to process path computation requests, and includes a Traffic Engineering Database (TED), containing a network topology information. The specific configuration for path computation in combined IP-optical networks therefore depends on the scope of information available to individual PCE subsystems in every layer as well as the level of cooperation between them [5]. The configurations can generally range from a fully integrated PCEs solution to a coordinated, but separate PCE in every layer.

An *integrated multi-layer PCE* solution, as the

[1] *This article is an extension of the workshop paper [6].*

This is especially critical in optical networks, where due to the analog nature of optical transmission, the relevant TE parameters need to be directly controlled on the actual optical device and are typically exchanged using vendor-specific extensions to control plane protocols.

name suggests, has the complete multi-layer topology information (a common TED) and uses this information to compute paths optimally. Despite its obvious advantages, the use of an integrated control plane is unlikely to happen in the near future. First, an integrated PCE does not conform with the typical organizational and technological separation between the Internet and optical networks. More importantly, however, due to the vendor-specific mechanisms within different technologies, and without standards, gathering critical parameters becomes a challenge, such as topology information or relevant TE parameters required for multi-layer path computation. This is especially critical in optical networks, where due to the analog nature of optical transmission, the relevant TE parameters need to be directly controlled on the actual optical device and are typically exchanged using vendor-specific extensions to control plane protocols.

In this context, a *coordinated* approach seems better suited, where each network layer is treated as a separate control domain, and to facilitate coordination, the client (IP/MPLS) network requests resources from the server (WSON) network. In its simplest form, a configuration like that assumes no direct communication between the PCEs. Instead, the MPLS-PCE identifies a pair of entry points (ingress and egress routers) between the Internet and the optical network and computes the path assuming that an optical circuit can be always be established between them. As there may be multiple possible candidates for entry points for a specific service request, and given that there is no intelligent coordination between the IP/MPLS and WSON PCEs, the computed path is likely to be suboptimal. We therefore propose to implement a slightly richer coordination mechanism, using what we refer to as *inter-PCE communication*, where the IP/MPLS and WSON PCEs exchange information about candidate entry points, path constraints, etc., to compute potentially optimal paths. Though past research has pointed to the potential issues of this configuration, such as additional computation overhead and signaling delays [7], we believe that they are not significant in most settings and that this mechanism is best suited for commercial applications in the near future. Two key features contribute to its suitability:

- The descriptions of multi-layer paths within the standardized PCE protocol via Explicit Route Object (ERO)
- The use of the VNTM controller

An overview of the configuration used in this work is presented in Fig. 2a. Here, the IP/MPLS and the WSON network each have their own PCE. In this architecture, the VNTM was proposed as an architectural concept to re-optimize the topology of the IP/MPLS network by triggering signaling for setup/decommissioning of IP links established using the WSON network [5, 8]. For instance, the VNTM can either suggest the provisioning of the IP links, or can support the establishment of optical circuits in response to specific service requests [9]. The information required for performing these operations is quite similar to what is available in the PCE, and in general, the functions of the VNTM and the PCE could be consolidated into a single controller module. However, the PCE standards do not specifically include any functions of provisioning, and cannot be used to actually trigger a traffic re-routing in the network, which would then be the role of VNTM [5]. This design choice is ensures that the PCE can be used irrespective of the actual provisioning mechanism in place be it via control plane, network management systems, or manual configuration.

The NCM is used to orchestrate interactions between the different components within the setup. In a real deployment, the NCM can be a module within the operator's Network Management System (NMS) or can even be deployed directly on a router. Upon arrival of a new connection request, the NCM requests a path for the same from the MPLS-PCE (1). In case a path cannot be computed in the IP/MPLS network, the MPLS-PCE identifies a set of candidate entry points and requests the computation of optical circuits between them from the WSON-PCE (1a). The WSON-PCE computes the necessary paths and returns the information to the MPLS-PCE, which constructs the multi-layer ERO and returns the computed path to the NCM (1b). In case the computed path contains multi-layer path segments, the NCM sends a request to the VNTM Uto setup the necessary optical circuit segments (2). The VNTM can use one of many different options available for provisioning optical circuits, and once the setup of the necessary optical circuits is complete, the NCM can initiate signaling to setup MPLS paths in the IP/MPLS network. It should be noted the standard interfaces/protocols for communication with the VNTM have not been yet defined. RFC5623 [5]suggests the use of ad-hoc mechanisms, including the PCEP as a possible protocol for communication with the VNTM, but does not define any message formats for the same. This is the primary motivation for the use of the NCM as an orchestrator in our setup, as integrating PCEs and VNTMs from different vendors using non-standard interfaces is not feasible in real networks. In our setup, we also demonstrate that the PCE protocol is a promising solution for communication with the VNTM, which if widely adopted could pave the way for a better coupling of the two components.

The signaling interactions between the different components responsible for provisioning a multi-layer path are shown in Fig. 2b. Upon the arrival of a service request, the NCM sends a request for path computation to the MPLS-PCE using the PCE protocol [4]. In case an active session does not exist between the NCM and the MPLS-PCE, the NCM initiates the session setup with the MPLS-PCE, which involves an exchange of *Open* and *Keepalive* messages, after which the NCM sends a *Path Computation Request* to the MPLS-PCE that contains information about the service endpoints and additional constraints on the computed path. The MPLS-PCE attempts to compute a path within the MPLS network, and if found, returns the computed path to the NCM. However, if a path is not found, the MPLS PCE communicates with the WSON-PCE to compute a multi-layer path. To communicate with the

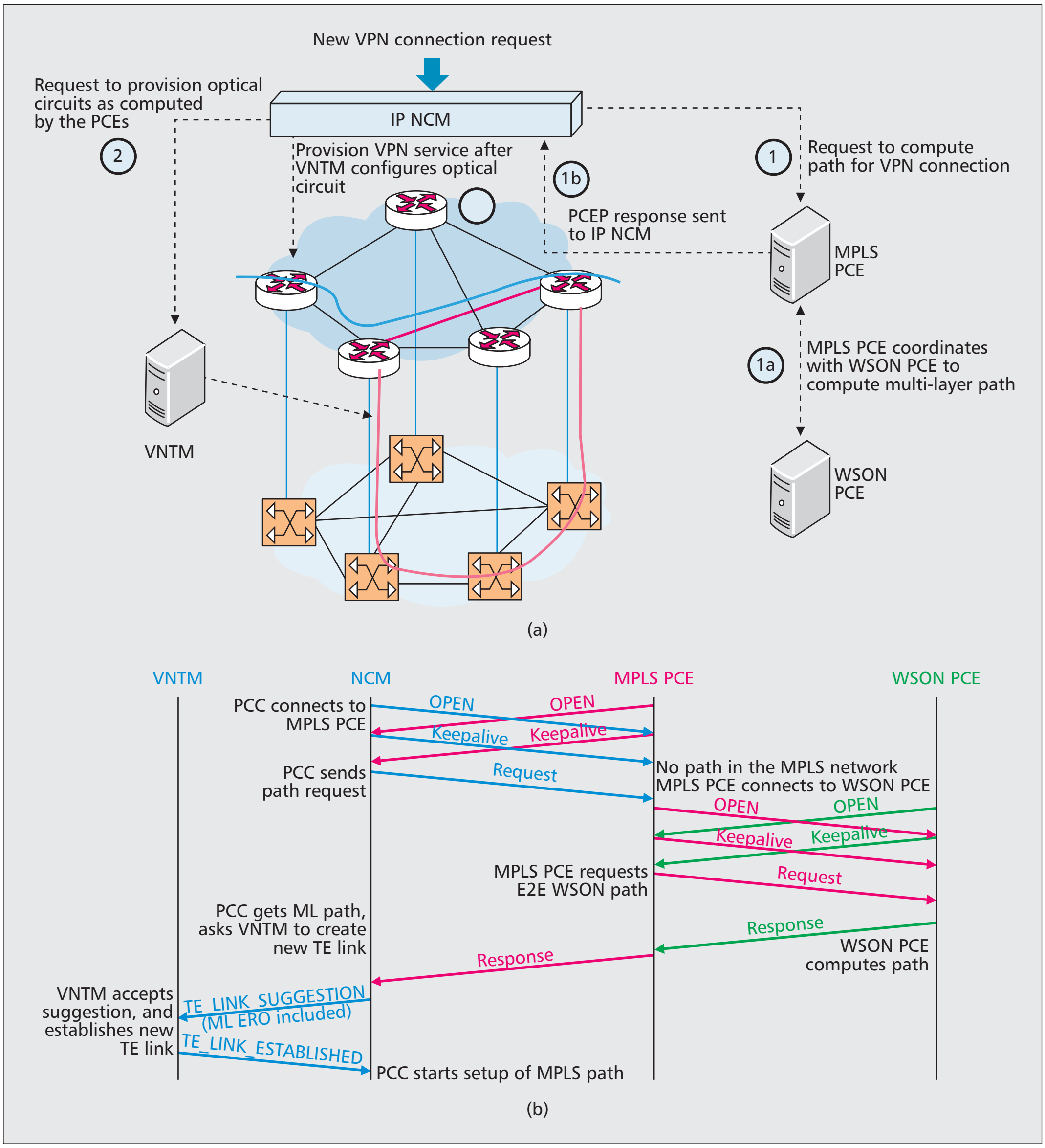

Figure 2. *Automated framework for multi-layer connection provisioning: a) multi-layer service provisioning with PCE and VNTM; and b) signaling flow for interaction between various components in the proposed architecture.*

The VNTM can use one of many different options available for provisioning optical circuits, and once the setup of the necessary optical circuits is complete, the NCM can initiate signaling to setup MPLS paths in the IP/MPLS network.

WSON-PCE, the MPLS PCE can re-use an active PCEP session or must initiate a new session by exchanging Open and Keepalive messages. The MPLS-PCE then sends a path computation request with the entry points as computed by the MPLS-PCE to the WSON-PCE.

As specified in RFC5440, the PCEP uses the Explicit Route Object (ERO) to define paths in the network. The ERO defines a path as a sequence of ERO sub-objects: for example a sequence of *IP Address* sub-objects is used to define the path in the IP/MPLS network, while a sequence of unnumbered interface sub-objects followed by the IP address of the destination switch is used to define the path in the optical network. Since the ERO specification in [4] cannot be used to describe a path which contains multiple path segments in different layers, a few extensions to the same were proposed in [10]. This new ERO sub-object (*SERVER_LAYER_INFO* object) was proposed as a delimiter to identify the start/end of each new connection in the server (optical) layer. In the path computation process presented in Fig. 2b, the WSON-PCE sends the computed paths between the entry points as an ERO in the *Path Computation Response* message to the MPLS-PCE, which then computes the end-to-end path and generates a multi-layer ERO (ML-ERO). The ML-ERO consists of path segments in the IP/MPLS and the WSON network separated using the *SERVER_LAYER_INFO* sub-object. This ERO is included within a Path Computation *Response* message to the NCM to complete the path computation process.

In case a multi-layer path is found, the NCM requests the VNTM to provision the optical circuits indicated in the ML-ERO. As specified, we propose the PCEP protocol for communication between the NCM and the VNTM. We define a new message (*TE_LINK_SUGGESTION*) which

In the testbed, the MPLS-PCE is deployed in Germany, while the VNTM, WSON-PCE, and the NCM are deployed on individual virtual machines on a single server at Telefonica premises in Madrid. Both PCEs are implemented based on IETF standards, and thus tested on interoperability.

includes the ML-ERO to send a request to the VNTM to provision the optical segments of a multi-layer path. The VNTM uses the optical segments as computed in the ERO to provision an optical circuit. Once the provisioning of the optical circuit is completed, the VNTM sends a *TE_LINK_ESTABLISHED* message to the NCM, which then initiates the provisioning of the MPLS circuit in the IP/MPLS network. The VNTM uses the User Network Interface (UNI) standard available on commercial routers to initiate signaling for provisioning of optical circuits, based on the current RSVP-TE standard extensions [11].

Experimental Testbed, Implementation and Measurements

This section goes *in media res* and describes the experimental testbed setup capable of replacing the manual workflow procedures shown earlier, while honoring the operator's premium on multi-vendor interoperability, administrative separation of the networks, and standard-based inter-operable solutions. A detailed overview of the testbed used for our interoperability study is presented in Fig. 3a. The testbed includes 4 control and management components, namely the IP Network Control and Management (NCM) unit (172.16.1.1), the VNTM (172.16.1.3), the WSON-PCE (172.16.1.2) and the MPLS-PCE (172.16.3.1). For the purpose of demonstration of a path setup, a traffic generator and sink are also incorporated on two of the routers, as shown in the figure.

In the testbed, the MPLS-PCE is deployed in Germany, while the VNTM, WSON-PCE, and the NCM are deployed on individual virtual machines on a single server at Telefonica premises in Madrid. Both PCEs are implemented based on IETF standards, and thus tested on interoperability. The WSON-PCE uses a proprietary implementation, while the MPLS-PCE is developed as open-source [12]. The Traffic Engineering Databases (TEDs) are located with the corresponding modules on the same machines, and can be updated using OSPF Link-State advertisement messages. The VNTM is designed to receive requests over PCEP and also implements two new messages (TE_LINK_SUGGESTION} and TE_LINK_ESTABLISHED}). Based on the requests, the VNTM initiates signaling for the setup of optical circuits in the WSON network using UNI. The NCM is implemented to orchestrate operations, and communicate with the MPLS PCE and the VNTM using the PCEP protocol.

A detailed description of the configuration used for the network equipment is presented in Fig. 3b. The testbed includes equipments from different vendors, i.e., 3 IP/MPLS Juniper MX240 routers and 4 ADVA ROADMs. Connections between the IP/MPLS routers are established using optical circuits in the WSON network. A separate network is established for control plane communication between the network devices and management components in the testbed. The control planes of the IP/MPLS and the WSON networks are separated via the use of independent subnets, with IP/MPLS routers using the 192.168.8.0/24 subnet and the WSON switches using the 172.16.1.0/24 subnet. Figure 3b also shows the addressing scheme used for interfaces in the IP/MPLS and the WSON network. Each interface in the IP network is associated with a control plane address (e.g. *ge* – 2/1/8 on *M*X240 – 1 has address 21.21.21.2) and client transponders on the ROADMs that connect to the interfaces on IP routers are also associated with a unique IP address (e.g. *Shelf* 1 – 11 on *ROADM* – 1 has IP address 21.21.21.1). On the other hand, the server transponders, used to interconnect ROADMs use unnumbered interface addressing with a unique transponder identified by the IP address of the ROADM and the sequence number assigned to it.

The testbed is focused towards demonstrating multi-layer path computation and highlighting open issues with respect to standards in the same. Therefore, for the purpose of this testing, we use a simple algorithm for multi-layer path computation. Here, in case the MPLS PCE is unable to find a path in the IP/MPLS network, it requests an end-to-end optical path from the WSON network. In case the end-to-end path is found, a multi-layer ERO is returned to the NCM, otherwise the path computation is terminated. Also, while MPLS paths are computed by the PCE setup (hence MPLS-PCE), and the routers used are MPLS capable, for the purpose of this demonstration, instead of configuring MPLS paths we configure *native IP routing* and *static routing rules* on all routers along the computed path to forward IP traffic.

Path Computation

The signaling exchange for path computation in the setup has been outlined earlier, and Fig. 4 shows the wireshark trace of the messages exchanged between the NCM, the MPLS PCE and the WSON PCE. The first sequence of *Open* and *Keepalive* messages indicates the establishment of a PCEP session between the NCM (172.16.1.1) and the MPLS-PCE (172.16.3.1), after which the NCM sends a *Path Computation Request* to the MPLS-PCE. In case the MPLS-PCE cannot compute a path in the MPLS network, it opens a new session with the WSON PCE (172.16.1.2) as indicated by the next sequence of O*pen* and *Keepalive* message exchanges. The MPLS-PCE then sends a *Path Computation Request* to the WSON-PCE, which responds with the computed path as an ERO inside a *Path Computation Reply Message* to the MPLS-PCE. The MPLS-PCE then computes a ML-ERO and includes it in a *Path Computation Reply Message* which is then sent to the NCM. Note here that in case active PCEP sessions already exist between any PCE peers (NCM - MPLS-PCE or MPLS-PCE — WSON-PCE), the exchange of *Open* and *Keepaliv*e messages is not necessary.

The details of the path computation request and response messages exchanged between the NCM, MPLS-PCE and WSON-PCE highlight some specific issues and challenges in multi-layer

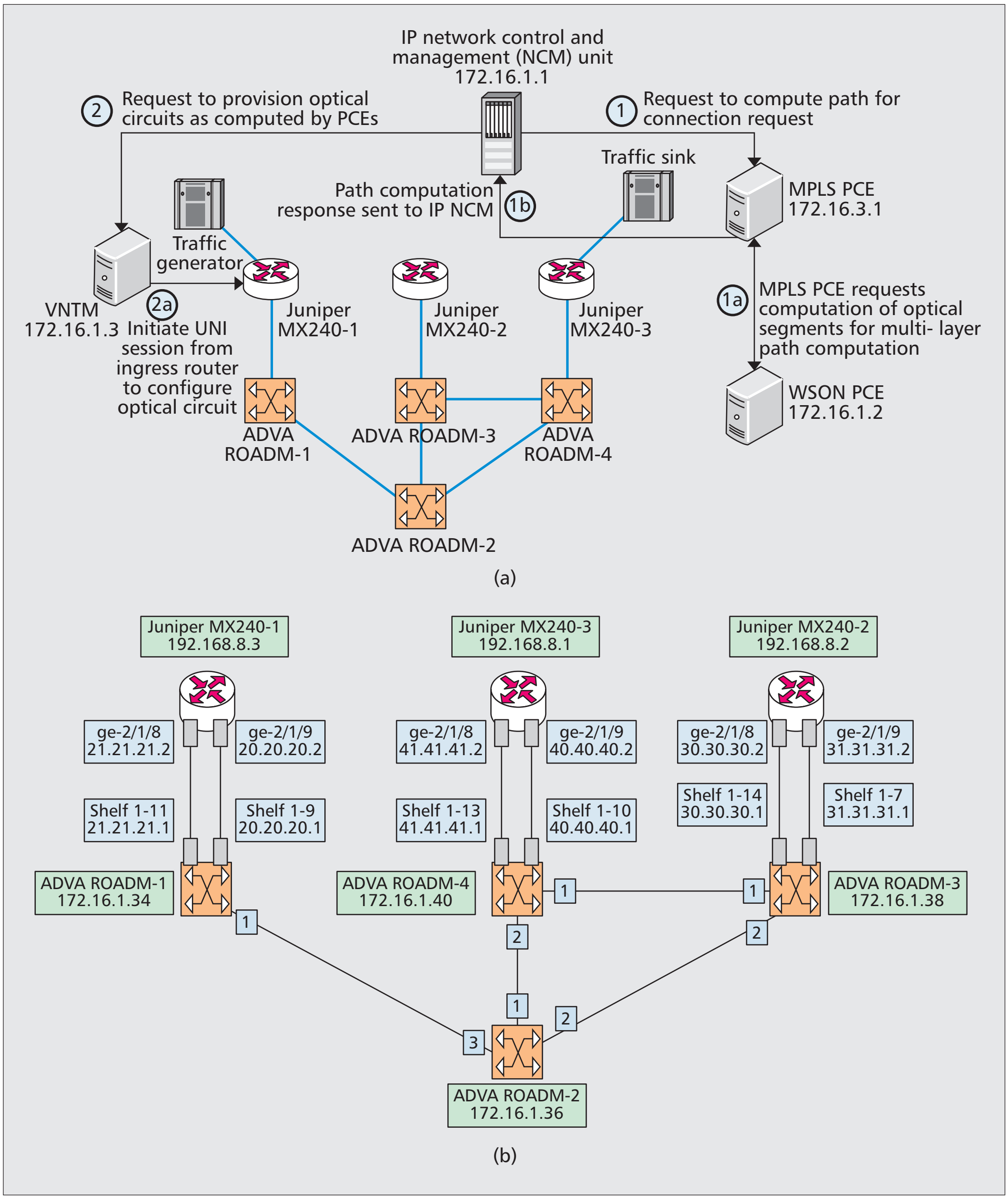

Figure 3. *Overview of the Testbed setup: a) Overview of testbed components; and b) addressing details.*

In our implementation, information about the interconnection of routers and corresponding optical switches as well as the information about free interfaces is initialized manually in the MPLS-PCE TED and is updated using ad-hoc scripts during run-time.

path computation. In the example presented in Fig. 4, the Path Computation Request from the NCM to the MPLS-PCE requests a path from router *MX240* – 1 (192.168.8.3) to *MX240* – 3 (192.168.8.1). In this setup, a path is not found in the MPLS network and the MPLS-PCE attempts to compute an end-to-end optical circuit between the two routers. Therefore, the path computation request sent from the MPLS-PCE to the WSON-PCE requests a path from *ROADM* – 1 (172.16.1.34) to *ROADM* – 4 (172.16.1.40). To facilitate this translation, the TE database in the MPLS-PCE must have the information about the inter-layer associations between the IP/MPLS interfaces and the corresponding client interfaces on the optical devices. For example, in this scenario, the TED in the MPLS-PCE should contain inter-layer association information indicating that MX240-1 (192.168.8.3) is connected to *ROADM* – 1 (172.16.1.34) over ge-2/1/8 and *Shelf*1 – 11 interfaces (Fig. 3b). However, the discovery of this information is non-trivial, as the Link Management Protocol (LMP) does not work between IP/MPLS routers and ROADMs on commercial equipment used in this test. In our implementation, information about the interconnection of routers and corresponding optical switches as well as the information about free interfaces is initialized manually in the MPLS-PCE TED and is updated using ad-hoc scripts during run-time.

The path computation response from the WSON-PCE to the MPLS-PCE shows the description of an optical circuit in the WSON network. The server transponders in the WSON network use unnumbered interface addressing in the testbed, and the ERO in the response consists of a sequence of server transponders used

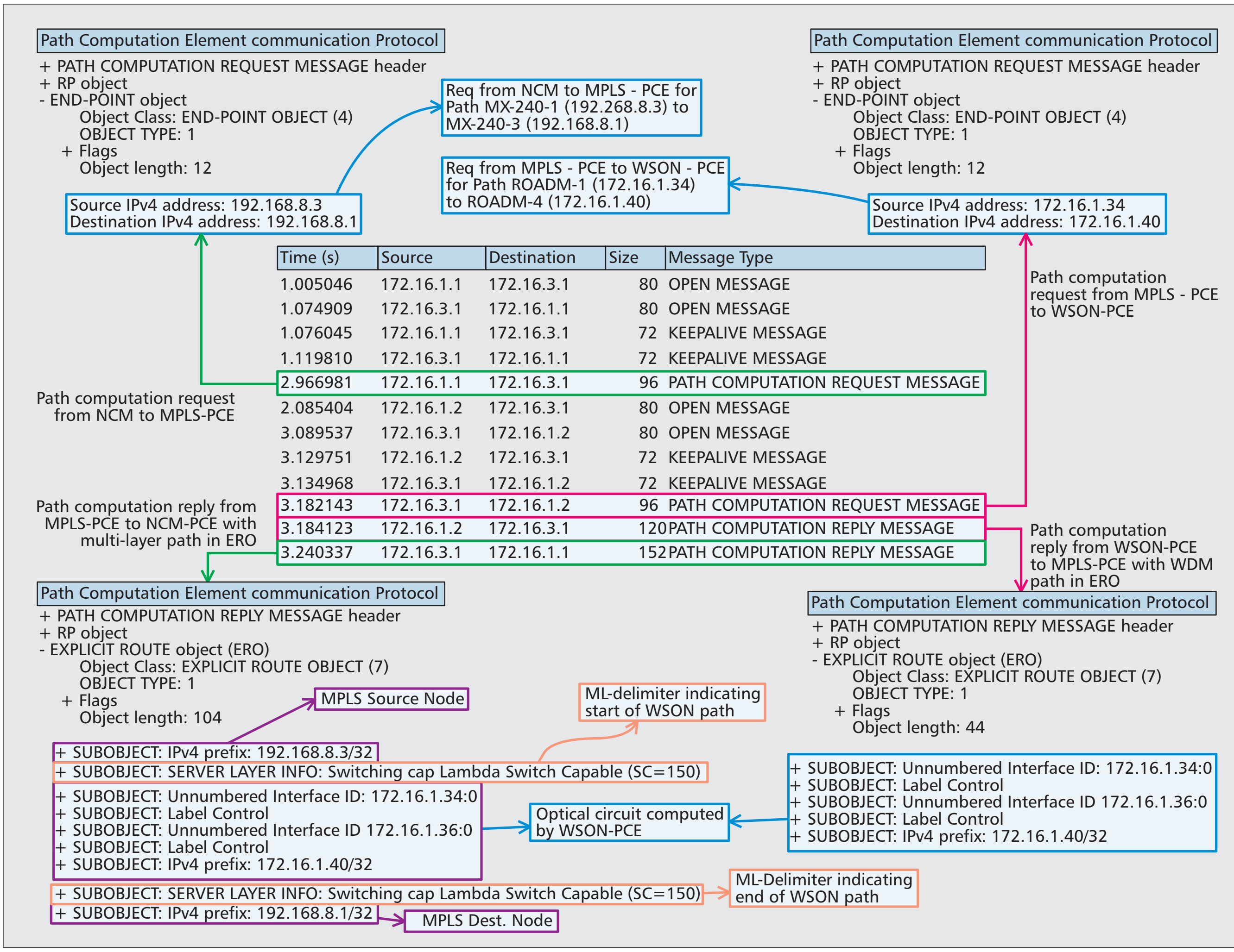

Figure 4. *Wireshark snapshots of the PCEP signaling observed.*

from the source to the penultimate ROADM in the network, while the final ERO sub-object is the IP address of the destination ROADM. The MPLS-PCE, upon receiving this response, creates a multi-layer ERO and includes it in the response to the NCM. The trace shows the use of the *SERVER_LAYER_INFO* object to indicate the demarcation between the path segments in the IP/MPLS and the optical network. In the response, a *SERVER_LAYER_INFO* object is included after the IP address of the source router 192.168.8.3, which indicates that the optical circuit is computed begins from this router. The second *SERVER_LAYER_INFO* object is included after the optical segment and is followed by an IP address object 192.168.8.1 which is the destination router for the optical circuit.

To assess the average time required for path computation in our implementation, we used the Atlanta virtual network topology from [13] as the base topology for both the MPLS and the WSON network, and emulated} path computation based on that topology. In the emulator, both PCEs were stateless, and the NCM was modified to generate path computation request messages between random source/destination pairs in the IP network which are sent to the MPLS-PCE. To enforce multi-layer path computation, the capacities of all MPLS links are set initially to 0. Just like previously, WSON-PCE and the NCM were deployed on the same physical machine (in Spain), while the MPLS-PCE was deployed in Germany, and the average Round Trip Time (RTT) (measured using ping traces) between the two sites was 44.29 ms (averaged over 1000 RTT's with a standard deviation < 0.15 ms). The emulation was used to measure the path computation times, and active PCEP sessions were established between the NCM, MPLS-PCE and WSON-PCE, and re-used during this test. The PCEP sessions, once established, can be re-used for different path computation requests and new PCEP sessions are not required for every connection request. Path computation requests were sent sequentially, i.e., after the completion of the previous request from the NCM. The average multi-layer path computation time within this setup was found to be 97.84 ms (σ = 1.02 ms). Given that the signaling involves two round trips, each with a delay of 44.29 ms, the total processing times within the two PCEs was ~ 9 ms. The total residence time within the MPLS-PCE was ~ 2 ms which included a negligible path computation time (< 1 ms), and the comparatively higher processing time on the WSON-PCE was

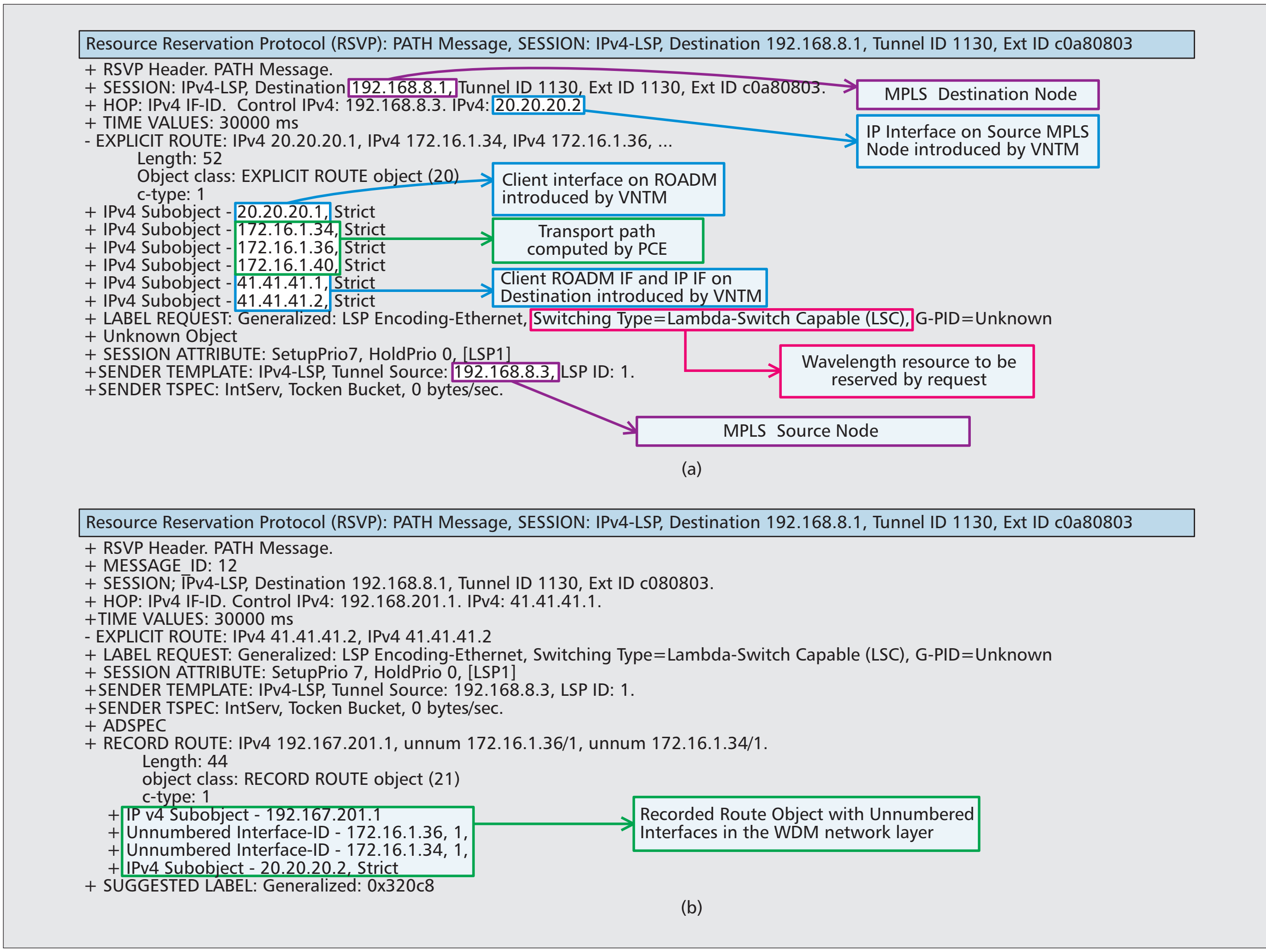

Figure 5. *UNI message exchange: a) UNI path message from Router 1; and b) UNI Resv message from Router 3.*

observed primarily due to the higher complexity of the wavelength routing algorithm, which computes the route using K-Shortest Path algorithm and finds free wavelengths using the First-Fit mechanism.

Path Provisioning with VNTM

After the multi-layer path computation by the PCEs, the NCM requests the VNTM to setup the optical circuit segments. In our implementation, we use the PCE protocol to exchange request/response messages between the NCM and the VNTM. As shown in Fig. 5a, after the computation of a path, the NCM sends a TE_LINK_SUGGESTION message which includes the multi-layer ERO to request the VNTM to setup the necessary optical path segments. For every optical segment that must be established, in our implementation, the VNTM connects to the ingress IP router, and using the Command Line Interface (CLI), initiates a UNI session (Fig. 5). This in turn initiates an RSVP session in the optical network to setup a circuit. Once the optical circuit is established, the VNTM notifies the NCM of the same using a *TE Link Established* message. At this point, the NCM initiates the configuration of the IP/MPLS routers so that traffic from the traffic generator can reach the traffic sink.

Figure 5a shows the wireshark trace from the RSVP path message sent from the source router MX240-1 (192.168.8.3) to MX240-3 (192.168.8.1). It is critical to note here that the route indicated in the ERO in the RSVP path message also includes information about the inter-layer links between the IP/MPLS routers and the WSON equipment which was not present in the computed multi-layer ERO as presented in Fig. 4. In our implementation, this information is introduced by the VNTM. The implemented VNTM contains a TE database similar to that of the MPLS-PCE, with all the information about the IP network topology, the available interfaces on IP routers and the inter-layer associations, as well as management information to access the different routers over CLI. As specified before, this information is not readily available in multi-layer multi-vendor networks and is initialized manually and updated using ad-hoc scripts. Using this information, the VNTM introduces specific interface control plane IP addresses for the IP interfaces as well as the corresponding client transponders on the ROADMs. As seen in Fig. 5a, the ERO indicates the use of the IP interface with IP address 20.20.20.2 on MX240-1 which is connected to 20.20.20.1 on ROADM-1 at the source, and the IP interface with IP address 41.41.41.2 on

One of the major challenges in multi-vendor environments involves mechanism for topology discovery. While this information was manually populated and updated using ad-hoc scripts in our implementation, automatic mechanisms are needed for the same.

MX240-3 which is connected to 41.41.41.1 on ROADM-3 is used at the destination. The RSVP ERO also uses IP addresses for intermediate WSON hops which is different than the sequence of unnumbered interfaces as computed by the WSON-PCE and this transformation is also performed by the VNTM. The trace for the response of the RSVP Path message is shown in Fig. 5b and it indicates the actual route used for provisioning that also includes the IP addresses of the transponders used to provision the actual optical circuit.

We measured the provisioning times for optical circuits on our testbed as the time taken to establish a UNI session from the source to the destination router. On average, the time to establish the UNI session was found to be 54 seconds, which is significant. The large path setup times are observed due to the known physical constraints of current ROADMs, which require a significant time to configure the physical layer properties on optical devices, necessary to stabilize the lightpath amplification and noise levels. As the optical technology evolves, these time requirements are bound to decrease, yet will always have to be factored into any software-defined network solution as an intrinsic physical constraint.

Open Issues and Upcoming Standards

Table 1 summarizes the requirements and system design choices we made in our implementation, as well as the resulting open issues with respect to the mechanisms used. From this table, it is clear that a number of practical requirements have not been completely addressed by standards and academic research in the area, as well as a number of open issues exist critical to the deployment in commercial networks.

One of the major challenges in multi-vendor environments involves mechanism for topology discovery. While this information was manually populated and updated using ad-hoc scripts in our implementation, automatic mechanisms are needed for the same, especially in large networks to eliminate errors arising from incomplete/inaccurate updates. Traffic Engineering (TE) information exchange between the two networks is also essential for operation in commercial networks, and can be addressed in a number of ways. The proposed Application Layer Traffic Optimization (ALTO) protocol [14] includes the capability to provide abstract topology information to an overlay application, which can be used to provide the IP/MPLS network an abstract view of the WSON topology in order to optimize path computation. The challenge can also be addressed by developing extensions to the GMPLS UNI to exchange TE information between the IP/MPLS and WSON networks. The mechanisms to populate and maintain TEDs (e.g., IGP protocols, standard interfaces to Network Management Systems (NMS)) are also non-trivial from the implementation perspective, and need to be evaluated in real systems, as the relevant information, such as available IP interfaces, may not be available by any single mechanism.

A few protocol specific issues can be highlighted, specifically in the context of the VNTM. The implementation used the PCE protocol for communication with the VNTM and we found it to be a suitable and relatively simple protocol for all practical purposes. The standardization of the protocol to the VNTM, especially by re-using the PCE protocol will significantly improve multi-vendor interoperability and will also open avenues for multiple interaction configurations between the PCE and the VNTM. Another possibility in this regard could also be the use of OpenFlow [15] which is gaining traction in optical transport networks and can be used to provision optical circuits.

We identified some timing issues, which in general should not be that critical in the context of service provisioning, as the timescales for provisioning a connection request is significantly lower than the inter-arrival time between service requests for commercial operators. However, a few increasingly important practical scenarios, such as dynamic multi-layer restoration may need to compute and provision paths in a much shorter time interval. For multi-layer restoration to become reality, these issues need to be studied within the context of a specific application to define limits on computation and provisioning times.

Finally, we demonstrated multi-layer (vertical) interoperability. Horizontal multi-vendor interoperability (i.e., in the same layer), also poses significant challenges, especially in the optical domain. The proposed model of cooperating PCEs can still be applied in this context, akin to a multi-domain scenario. The demonstration also presented the orchestration of various components including the PCE and the VNTM using an NCM, but the framework need to address additional issues such as policy management, security, OAM etc. under a standard framework that can be extended to multiple network scenarios. A proposal for the same is currently being developed within the IETF under the Application-based Network Operations (ANBO) architecture [2].

Summary and the Outlook

In the coming years, it is highly likely that future software-defined networking solutions across the whole telecommunications and networking spectrum will find a use for path computation function as a central building-block for the work done. We expect considerable advances in these efforts to be enabled by inter-operable PCE-based architectures. In our contribution, we implemented a PCE testbed capable of multi-layer path computation and provisioning in a network equipped with commercial IP routers and optical switches. We designed the testbed in line with the operators' requirement to put a premium on multi-vendor interoperability, administrative separation of networks, and all-standardized solutions. Our experiments have shown that that a number of practical requirements have not been completely addressed by the current standards and academic research in the area. These issues include mechanisms for topology discovery, TE information exchange,

	Requirements	System Design and Implementation	Open Issues and Solutions
Topology Information	**Inventory management** • Information about available IP/MPLS router interfaces and optical interfaces	• Available interface information not disseminated in IGP • Manual update in TED	• Standard interface needed to NMS/inventory management to populate TED
Topology Information	Inter-Layer Connectivity • IP/MPLS TED: Information about interface connectivity between routers and switches • VNTM TED: Inter-layer associations between IP router interfaces and client interfaces on optical switches	• The connectivity information is not available and is configured manually (one time) in the TEDs • Inter-layer associations are not available and are maintained in a separate database	• Standard control/management mechanisms, akin to Link Management Protocol (LMP), to discover neighboring interfaces between the IP and the optical network equipments in a multivendor setting
Topology Information	**TE Information Exchange** • From IP/MPLS to WSON networks • From WSON to IP/MPLS about optical circuits established (e.g., Shared Risk Link Groups (SRLG))	• Algorithms choose entry-points to WSON based purely on IP/MPLS topology • TE Information for computed and established paths not available	• Use of upcoming standards such as ALTO [14] to share abstract WSON topology with IP/MPLS PCE • Extensions to UNI specification to provide TE information of established circuits
Architecture and Protocols	**PCE Protocol** • Extensions to define multi-layer paths for the ERO	• SERVER_LAYER_INFO object in the ERO to identify start/end of path segment in different layers	• Further extensions required to the ERO to for description of complex path definitions (e.g., working and protection paths)
Architecture and Protocols	**VNTM** • Standard protocol for communication with VNTM • Choice of provisioning mechanism for VNTM to initiate Optical Circuit setup in multi-vendor network	• Standard protocol unavailable, PCEP used for communicating with the VNTM • CLI used to connect to IP routers and initiate UNI session from the routers to establish optical circuits	• Analyze the suitability of the PCEP protocol for VNTM, while addressing issues of discovery, policy frameworks and notifications • Analyze standard frameworks such as OpenFlow[15] that can be used for provisioning optical circuits
Architecture and Protocols	**Timing Issues** • Reduce large path computation and circuit provisioning times that limit usability for time critical services • Reduce large provisioning delays that can cause race conditions	• Computation algorithm to use only one round between the MPLS and WSON PCE • Explicit (custom) notifications for updating information on interfaces used in IP network to reduce topology update times	• Methods to exchange topology information with improved path computation algorithms and reduced inter-layer communication • Analysis of stateful PCEs and algorithms to evaluate expected provisioning times, especially considering timing constraints of existing technologies
Architecture and Protocols	**Component Separation** • Ideal architectural separation that best enables automated (application-driven) provisioning in multi-vendor/technology/layer networks	• Independent PCEs used for each layers during path computation • Path computation and provisioning functionalities separated (PCE and VNTM) • Orchestration of process performed by the NCM	• Integration of other critical components such as policy framework and AAA within a standard architectural framework such as ANBO [2]

Table 1. *Summary of the various requirements, implementation specifications and open issues.*

standardized VNTM interfaces as well as timescales for provisioning. As we have demonstrated, the innovation in path computation solutions can be significantly expedited if the research and industrial communities join efforts to make the proposed solutions inter-operable and practical.

Acknowledgements

This work was supported by the European Commission through the ONE project in the Seventh Framework Program (FP7), contract number INFSO-ICT-258300. We also thank Adrian Farrel for insightful discussions on future PCE-based architectures. We thank anonymous reviewers, Guest Editors and the EiC for useful comments and support in the final stage of the article.

References

[1]A. Farrel, J. P. Vasseur, and J. Ash, "A Path Computation Element-Based Architecture," IETF RFC 4655, Aug. 2006, http://tools.ietf.org/html/rfc4655.
[2] D. King, A. Farrel, "A PCE-based Architecture for Application-based Network Operations," IETF Internet Draft, Dec 2012, http://tools.ietf.org/html/draft-farrkingel-pce-abno-architecture-00.
[3] European Telecommunications Standards Institute White paper, "Network Functions Virtualisation — An Introduction, Benefits, Enablers, Challenges & Call for Action," Oct. 2012, http://portal.etsi.org/NFV/NFV_White_Paper.pdf.
[4] J. Vasseur and J. L. Roux, "Path Computation Element Communication Protocol," IETF RFC 5440, Mar. 2009, http://tools.ietf.org/html/rfc5440.
[5] E. Oki et al., "Framework for PCE-based Inter-Layer MPLS and GMPLS Traffic Engineering," IETF RFC 5623, Sept. 2009,. http://tools.ietf.org/html/rfc5623.
[6] M. Chamania *et al.*, "Coordinated Computation of Multi-layer Paths via Inter-layer PCE Communication:

Standards, Interoperability and Deployment," *IEEE Int'l. Conf. Communication ICC 2012, Wksp. Telecommunications: From Research to Standards.*
[7] S. Gunreben and F. Rambach, "Assessment and Performance Evaluation of PCE-based Inter-Layer Traffic Engineering," *Int'l. Conf. Optical Network Design and Modeling*, 2008, Mar. 2008, pp. 1–6.
[8] K. Shiomotoi *et al.*, "Requirements for GMPLS-Based Multi-Region and Multi-Layer Networks (MRN/MLN)," IETF RFC 5212, July 2008, http://tools.ietf.org/html/rfc5212.
[9] T. Miyamura *et al*., "Demonstration of PCE-Controlled Dynamic Traffic Engineering for GMPLS-based Multilayer Service Network," *IEE Conference Publications*, vol. 2005, no. CP502, 2005, pp. v2–269–v2–269, available: http://link.aip.org/link/abstract/IEECPS/v2005/iCP502/pv2-269/s1.
[10] E. Oki *et al.*, "Extensions to the Path Computation Element communication Protocol (PCEP) for Inter-Layer MPLS and GMPLS Traffic Engineering," IETF draft-ietf-pce-inter-layer-ext-06.txt, Jan. 2012, http://tools.ietf.org/html/draft-ietf-pce-inter-layer-ext-06.
[11] D. Papadimitriou *et al.*, "Generalized MPLS (GMPLS) Protocol Extensions for Multi-Layer and Multi-Region Networks (MLN/MRN)," IETF RFC 6001, Oct. 2010, http://tools.ietf.org/html/rfc6001.
[12] M. Chamania, M. Drogon, and A. Jukan, "An Open-Source Path Computation Element (PCE): Design, Implementation and Performance," *IEEE J. Lightwave Technology 2011*, 2011, Source Online, available: http://opensourcePCE.org
[13] Atlanta Network Problem Instance, SNDlib — Library of test instances for Survivable fixed telecommunication Network Design, available: http://sndlib.zib.de/
[14] R. Alimi, R. Penno, and Y. Yang, "ALTO Protocol," IETF Internet Draft, May 2013, http://datatracker.ietf.org/doc/draft-ietf-alto-protocol/.
[15] N. McKeown *et al.*, "OpenFlow: Enabling Innovation in College Networks," Whitepaper, Mar. 2008, http://www.openflow.org//documents/openflow-wp-latest.pdf.

BIOGRAPHIES

MOHIT CHAMANIA (m.chamania@tu-bs.de) is currently pursuing his Ph.D. at the Technical University of Braunschweig, Germany. Prior to his Ph.D. studies, he obtained his B.Tech and M.Tech in Mechanical Engineering at the Indian Institute of Technology, Bombay (2002-2007). His research interests include hybrid IP-over-optical networks, interdomain routing, Ethernet and Optical technologies. Mohit is the main software architect in the open source PCE project (opensourcePCE.org).

ADMELA JUKAN (a.jukan@tu-bs.de) is Chair Professor of Communication Networks in Electrical and Computer Engineering Department at the Technical University of Braunschweig, Germany. She currently coordinates the EU-Project ONE (ict-one.eu), with focus on next generation network management systems in combined IP and optical networks, and her group has developed the first open source PCE emulator (opensourcePCE.org).

FERNANDO MUÑOZ, B.S. Telecommunications Engineering in 2013. He joined Telefonica R&D in 2009 and his work is focused in convergence of multi-layer networks especially in IP/MPLS network over GMPLS photonic transport networks. Control plane integration and automated core network operation are also of Fernando's interests.

MARTA CUARESMA received her degree in Computer Engineering and Mathematics from Universidad Autónoma de Madrid (UAM), Spain in 2012. From October 2011 to March 2013, she did an intership in Telefonica I+D. Currently she is working at Meteologica and studying a M.Sc. degree in Research and Development in Telecommunications.

OSCAR GONZALEZ DE DIOS (ogondio@tid.es) received his MS in Telecommunications Engineering in 2000 and his Ph.D. from University of Valladolid in 2012. He is a researcher in Telefonica I+D working in the Core Network Evolution Initiative, part of the Global CTO Unit. His research interests include network planning and design, multi-domain and multi-layer networks and control plane. He is currently active in IET CCAMP and PCE working groups, where he is co-author of several drafts.

VÍCTOR LÓPEZ received his M.Sc. degree in telecommunications engineering from Universidad de Alcalá de Henares, Spain, in 2005 and his Ph.D. degree in computer science and telecommunications engineering from Universidad Autónoma de Madrid (UAM), Spain, in 2009. In 2004, he joined Telefónica I+D as a researcher and he joined UAM in 2006, where he became assistant professor. In 2011 he joined Telefonica I+D as a technology specialist. His research interests include metro and backbone architectures.