Open Network Database
for Application-based Control
in Multi-Layer Networks

F. Paolucci, F. Cugini, G. Cecchetti and P. Castoldi

Abstract—**Traffic Engineering (TE) solutions are enabled by the adoption of the Path Computation Element (PCE) architecture. Cooperative inter-PCE communications, in combination with TE information export mechanisms, such as the Link State extensions to BGP protocol (BGP-LS) enable TE in heterogeneous networks scenarios with different layers, domains and technologies. Multi-layer networks, such as IP/MPLS over Elastic Optical Networks (EON) are a typical example of such scenarios.**

However, these currently adopted export mechanisms are conceived for limited amount of exchanged information in order to preserve scalability, especially if centralized entities are considered, such as PCEs and controllers. Moreover, some information (e.g., router interfaces and transceiver card association) are not always available at the control plane level. Thus, so far limited deployment of inter-layer TE procedures has been experienced.

In this work we propose a novel network element called Open Network Database. Exploiting the state-of-the-art of database platforms while assuring high scalability and reliability performance, dedicated control and management applications run complex re-optimization procedures with a set of ad-hoc queries without relying on complex control infrastructure or without overloading the orchestrators, thus simplifying the implementation of network applications and achieving advanced TE. A YANG-based database for interconnected TE networks is proposed and implemented to enable dedicated applications to perform complex multi-domain/layer operations. The database is successfully utilized for effective and extremely scalable maintenance applications in multi-layer networks.

Index Terms—**Elastic Optical Networks, Multi-Layer Networks, Traffic Engineering Database, open database, IP-over-WDM, network maintenance, LSP, PCE, YANG.**

I. INTRODUCTION

THE Path Computation Element (PCE) architecture [2] has been introduced to provide effective Traffic Engineering (TE) solutions by exploiting a dedicated network entity devoted to path computation process. Communications between path computation clients (PCC) and PCEs, realized through the PCE Protocol (PCEP), also enable inter-PCE communications offering an attractive way to perform TE-based path computations among cooperating PCEs having different domains of visibility. The multi-layer scenario is a typical example of different domains of visibility, mainly due to different technologies (e.g., optical or packet switching), different equipment vendors, different administrative management and scalability issues.

In multi-layer networks, correlation and exchange of TE information may be achieved by exploiting PCEP-based procedures. For example, a PCE controlling the Packet Switching Capability (PSC) domain may request a PCE dedicated to the Lambda Switching Capability (LSC) domain to expand loose Label Switched Paths (LSPs) in the optical domain by means of optical bypasses, forwarding adjacencies and virtual links. More recently, the Link State extensions to Border Gateway Protocol (BGP-LS) have been also introduced to provide the cooperating PCE with TE information [3].

However, PCEP and BGP-LS protocols are typically deployed to limit the amount of exchanged information (e.g., limited amount of segment expansions, exchange of address reachability information only), such that scalability issues are avoided on the remote elements. In addition, and most important, some pieces of information are often not dynamically retrieved by the control plane, such as the association between a card in the PSC router and the physically attached LSC transponder in a Reconfigurable Add-Drop Multiplexer (ROADM). For these reasons, several multi-layer operations and network application (NetAPP), e.g. for maintenance and re-optimization, are still left to manual interventions, without relying on fully automatic mechanisms.

In this work, we take advantage of the impressive scalability performance of the latest generation of database technologies [4] to provide an effective and standard way to correlate TE information originated by different domains/layers. An Open Network Database (ONDB) is proposed as a novel and independent network element storing YANG-defined TE information [5] of heterogeneous network segments having limited visibility of each other (e.g., areas, layers, domains). That is, the database storing each network domain/layer resources is no more considered as a close element within its PCE/controller, instead it is an additional open element directly available to NetAPP (subject to policy control). The proposed ONDB is the key element of an enhanced control and management architecture enabling effective TE while preserving controller scalability. The scalability of this proposal resides in the enormous performance improvements achieved by latest database technologies [4].

The ONDB is implemented to enable dedicated management/control NetAPP to perform complex TE operations, including automated maintenance and re-optimization in multi-

layer networks. Automated maintenance application is evaluated in terms of scalability and experimental validation in a multi-layer PCE-based control plane testbed including Elastic Optical Network (EON) and Segment Routing MPLS domain.

With respect to its preliminary version [1], this work, besides motivations and a related work analysis, proposes a general and innovative APP-based control architecture encompassing the ONDB and extensive implementation results evaluating the ONDB scalability in terms of query time and required storage.

II. RELATED WORK AND MOTIVATIONS

Multi-layer control architectures and procedures exploiting TE export mechanisms have been mainly focused on planning/provisioning and restoration operation. The former targets efficient mapping of PSC-LSP onto either existing or new lightpaths with the aim of improving combined network utilization. The latter addresses the efficient rerouting of a potentially high number of PSC LSPs due to node/link failure events affecting the optical layer. The most considered architectures rely on the adoption of Path Computation Element (PCE) [6] queried by the Virtual Network Topology Manager (VNTM) in order to enable inter-layer computation, instantiation and encapsulation.

Multi-layer planning relying on traditional TE export mechanisms has been investigated deeply with focus on the benefits provided by optical bypass [7]. Efficient traffic grooming strategies encapsulating PSC LSPs over EON lightpaths have been studied also in the context of software defined networking (SDN) framework [8]. Inter-layer applicability has been assessed in different testbeds, mainly based on PCE and the Application Based Network Operation (ABNO) architecture [9], [10]. Work in [11] proposes multi-layer restoration using global concurrent optimization for bulk restoration of LSPs.

Network orchestrators have been also deployed on top of the control hierarchy to coordinate network domains and layers of different technologies, switching capabilities or vendors. Multi-layer Optical Packet Switching (OPS) / Optical Circuit Switching (OCS) [12] and multi-domain heterogeneous network scenarios [13] have been proposed. Multi-domain multi-vendor scenarios with heterogeneous control platforms are considered particularly suitable for orchestrators [14]. The Hierarchical Path Computation Element (HPCE) architecture has been also proposed for the multi-layer scenario [15],however its deployment has been considered mainly for the multi-domain optical-layer scenario [3].

In general, inter-layer procedures define complex work-flows involving different protocols, sessions and control elements (i.e., PCEs, VNTM, ABNO controller). Network orchestration, including the HPCE architecture represents a promising approach, however the scalability of such proposal requires careful evaluation, as long as orchestrators and controllers are supposed to incorporate additional noticeable amount of functions, procedures, protocols in a centralized architecture. The actual results in terms of industrial exploitation are still limited and, even nowadays, telco operators adopt

basic multi-layer procedures, just replicating functions at each layer (e.g., fast-reroute MPLS restoration over 1+1 protected lightpath). Therefore, without strict and effective coordination between the IP and the optical layer, which are historically kept independent and separated, multi-layer traffic engineering is not enforced yet thus implying waste of network resources.

The main motivation of this work relies on the challenge of simplifying, and thus exploiting, inter-layer work-flow without shifting the complexity to a centralized orchestrator and, most important, preserving acceptable degree of independence of IP and optical layers. The adoption of a multi-purpose network database has the potential of drastically simplifying a number of complex procedures such as re-optimization, maintenance, equipment upgrade in an automatic way and thanks to the extremely advanced capabilities of recent database platforms to process complex multi-match queries.

Open source platforms are nowadays hot topic in the context of network controllers (e.g. OpenDaylight, ONOS). Focusing on databases, an open source database has been proposed recently [16] based on ITU-T G.805 data modeling, conceived as open standard internal TED module for multi-layer PCE. Differently from this work, where the Open Network Database is defined as a novel independent network element, the implementation in [16] aims at becoming a standard internal module of existing PCE platforms. The use of dedicated database infrastructures within control and management plane, and in particular within multi-layer networks, have been also proposed recently with different purposes, e.g. in the context of big data analysis for flow classification [17], [18].

III. OPEN NETWORK DATABASE ARCHITECTURE AND USE CASE

The Open Network Database (ONDB) is proposed as an additional database with respect to existing databases in the network (e.g., PCE TED, PCE LSP-DB). It stores information retrieved from a set of existing controller databases and additional information directly stored by the management plane (e.g., inventory, static associations). The main purpose of the ONDB is to implement efficient TE information export mechanism for supporting effective NetAPP. Therefore, the proposed ONDB is conceived as an efficient enabler of inter-domain/layer orchestration, avoiding the involvement of the orchestrator for selected network operations. The definition of the ONDB is general and can be considered for databases of different layers, vendors, domains, carriers. Standard YANG definitions are considered within the ONDB, such that vendor-independent NetAPPS can perform queries and elaborations having policy-enabled access to selected parameters of the whole network [5]. In this work we focus on databases related to different network layers.

In particular, [5] defines a YANG data model for representing TE topologies through node and link attributes in networks operated with MPLS-TE and GMPLS. The objective is to model topology parameters in a protocol independent way (i.e., technology agnostic), such that domains operated with different interior gateway protocols can be modeled in a common way and path computation is enabled across
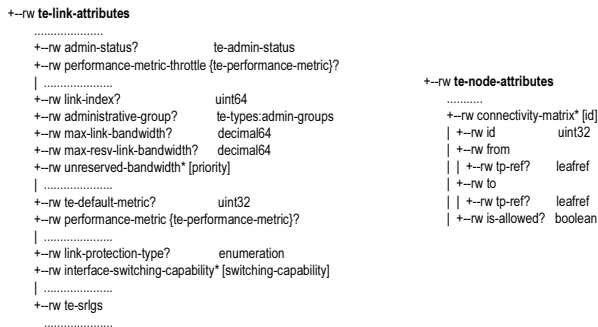
```
+--rw te-link-attributes
     ...............
     +--rw admin-status?                    te-admin-status
     +--rw performance-metric-throttle {te-performance-metric}?
     |    ...............
     +--rw link-index?                      uint64
     +--rw administrative-group?            te-types:admin-groups
     +--rw max-link-bandwidth?              decimal64
     +--rw max-resv-link-bandwidth?         decimal64
     +--rw unreserved-bandwidth* [priority]
     |    ...............
     +--rw te-default-metric?               uint32
     +--rw performance-metric {te-performance-metric}?
     |    ...............
     +--rw link-protection-type?            enumeration
     +--rw interface-switching-capability* [switching-capability]
     |    ...............
     +--rw te-srlgs
          ...............
```

```
+--rw te-node-attributes
     ...........
     +--rw connectivity-matrix* [id]
     |   +--rw id              uint32
     |   +--rw from
     |   |  +--rw tp-ref?      leafref
     |   +--rw to
     |   |  +--rw tp-ref?      leafref
     |   +--rw is-allowed?     boolean
```

Figure 1.   YANG models of generic link and node attributes [5].

Figure 2.   Proposed multi-layer NetAPP-based management architecture including the Open Network Database.

multiple domains. In case of a multi-layer network, nodes and links only allow representation of each network layer as a separate topology, maintaining a per-layer view. Merging of complete or abstract topologies is then supported to enable multi-domain/layer path computation, possibly implemented through the introduction of additional node or link identifiers if the native topologies adopt overlapping naming spaces. Technology agnostic parameters can then be augmented by technology specific models, enabling the control of the full set of network attributes. Node attributes include information related to data plane aspects (e.g. the intra-connectivity matrix in an optical node) and configuration information (e.g. IP address or Router ID). Similarly, link attributes include parameters related to the data plane (e.g. unreserved bandwidth) as well as configuration information (e.g., remote node/link IDs, Shared Risk Link Groups - SRLGs, administrative colors, etc.). A TE Path is then defined as an ordered list of links and/or nodes on the topology graph, obtained as a product of successful path computation. Fig. 1 shows an example of the generic node and link attributes defined in [5]. The figure also shows for each attribute the related data type (e.g., unit32, boolean, etc). In the ONDB implementation proposed in this work, the attribute names and data types defined in [5] are adopted.

The general architecture of a multi-layer network encompassing the proposed ONDB is shown in Fig. 2. The multi-layer management plane is supposed application-based, where each procedure is conceived as a different NetAPP. Link/node maintenance, constrained re-optimization, greenfield full-mesh provisioning are examples of possible NetAPPs. Each layer is assumed to be handled by a dedicated controller, e.g. the PSC Controller and the LSC Controller. The proposed ONDB is a multi-table database which may be populated by different entities. In particular, the ONDB stores the Traffic Engineering information of each layer, populated by the related controller. Write accesses to the ONDB are granted to layer controllers and to the management plane to enable automatic TE replica. In the case of stateless controllers, the Traffic Engineering Database (TED) of each layer is stored, whereas in the case of stateful controllers also the additional LSP database (LSP-DB) is stored.

Different interfaces and protocols may be utilized for TE replica and database population. For example, if PCEP is used, Path Computation Report Messages (`PCRpt`) reporting the LSP establishment/modification/release event may be for-
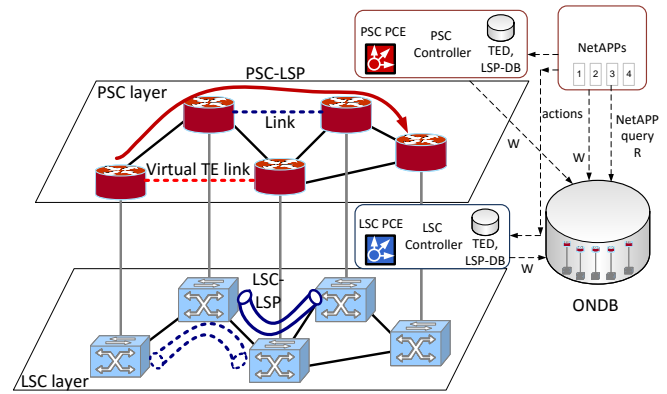
warded by the controller towards the database, thus storing the new or updated LSP entry. It is worthwhile to note that `PCRpt` messages encloses Explicit Route Object (ERO) or Record Route Object (RRO) detailing the entire end-to-end LSP path, including nodes, links and attributes (e.g., bandwidth, spectrum). Similarly, BGP-LS update messages may be used to export Link State TE information from controllers to the database. Additional interfaces may be used, for example database client protocols (e.g., SQL) or Representational State Transfer (REST) based on HTTP. All the mentioned interfaces may enable automatic TE replica towards the ONDB with limited impact on the complexity of the controller. Management plane has write access rights in order to store static management information. For example, the association between router interfaces and transceiver cards is static and typically handled by manual interventions and stored in the management inventory.

NetAPPs have the rights to perform read access to the ONDB. In the case of read access, the use of native database protocol, e.g., Structured Query Language (SQL) protocol, to perform queries is assumed. In particular, each NetAPP runs dedicated set of queries in order to correlate specific inter-layer information. Once retrieved correlated results, NetAPPs trigger the related multi-layer procedures by sending proper list of actions towards the layer controllers. In the case of sequential multi-step actions, the acknowledgement of the last action may be required in order to the enforcement of the next action. Actions are intended as correlated results of the network resource set controllers will need to operate over. This way, NetAPPs do not need to take any decision and their implementation is kept extremely simple (i.e., few lines of code made of a set of queries towards the database).

A. Use case: Optical Link Maintenance

The use case of optical link maintenance in a single-carrier multi-layer network is considered to assess the ONDB-based architecture. The sample topology composed of five sites is assumed, as shown in Fig. 3a. Each site includes a router (equipped with PSC interfaces) and a ROADM (equipped with

LSC interfaces). Each router interface is physically attached to a ROADM interface (i.e., transponder). In this example, lightpaths are established between adjacent ROADMs. Moreover, one additional lightpath is computed by the LSC PCE and established between ROADM A and C, having optical bypass in D. The obtained virtual topology at the PSC layer is shown in Fig. 3b. Note that a direct link is now introduced between Router A and C. Several MPLS LSPs are then computed by the PSC PCE and established, including an LSP from A to F passing through router C (LSP PSC 3 in Fig. 3b).

Maintenance operation of an optical link is then considered. To avoid traffic disruption due to rerouting operations at the LSC layer, rerouting of involved PSC LSPs is typically performed in advance at PSC layer, subject that adequate bandwidth resources are available in the rest of the PSC network. That is, although protection and restoration functions are available at the optical layer, it is desirable to minimize the possibility that SLA is affected, or that traffic remains unprotected during the maintenance operations. Thus, maintenance operations at the optical layer are carefully performed in off peak hours trying to free the link in advance from its data traffic. First, affected lightpaths have to be identified and associated to the related virtual links at the PSC level. Then, the PSC LSPs traversing these virtual links have to be rerouted.

For example, in case of maintenance of link A-D, lightpaths A-D and A-C have to be identified, corresponding to virtual links A-D and A-C at the PSC level. Then, three LSPs have to be selected for rerouting: A-D, A-C and also A-F. On the other hand, in case of maintenance of link B-C, LSP A-F is not considered by rerouting operation.

However, due to the separation between IP and optical layers management and control maintained by network operators, the PSC PCE is not typically able to automatically perform the needed correlations between the actual optical links exploited by the PSC LSPs. Indeed, PCEP and BGP-LS may not be adequate, since the interconnection between router interface and ROADM transponder is not typically included in control plane advertisements, although mechanisms have been proposed to advertise such information [19]. Moreover, concerns on scalability issues are considered when providing each PCE with excessive amount of information, particularly in the case of active stateful capabilities.

The ONDB contains layer-specific information (e.g., active LSPs collected by each layer-specific PCE through PCEP `PCRpt` messages) as well as inter-layer information, (e.g. static associations between the router interface and the ROADM transponder). In this considered single-carrier scenario, all relevant network information are exported without applying policy restrictions (i.e., no confidentiality issues have to be considered in this use case).

For the use case of link maintenance, a portion of the most relevant information included in the ONDB is reported in Tables I, II and III of Fig. 3. As a first step, the NetAPP devoted to maintenance procedures receives as input the optical link under expected maintenance. Then, it performs a set of queries to the ONDB. As outcome, the ONDB provides the NetAPP with (1) the list of links at the PSC level to be excluded by the rerouting process and (2) the list of involved PSC LSPs.

The set of queries performed in the use case is the following. A first query, involving Table I, selects the association between lightpaths, identified by the LSP Lambda id ($LLid$) and traversed optical interfaces, identified by unnumbered interfaces (i.e., $IPv4$ LSC node address and $port$). A second query, involving Table II, finds the association between lightpath endpoint optical interfaces, i.e., transponder source (i.e., $Tsrc$ and $Tsrcport$) and destination (i.e., $Tdst$ and $Tdstport$) addresses, and attached router interfaces, i.e., router source ($Rsrc$) and destination ($Rdst$) addresses. A further query, based on the results of the previous queries, is performed involving Table III, identifies the involved PSC LSPs, identified by the LSP Packet ID ($LPid$) and by their traversed router interfaces and virtual links ($IPv4$ PSC node).

For example, in the case of maintenance applied to link A-D, the NetAPP queries the ONDB as in the following. The A-D link described with its unnumbered interface, i.e. composed by the node IPv4 address (10.0.0.1) and the local identifier (value 2), is considered as input parameter. According to Table I, the ONDB first returns two LSC LSPs (i.e., $LLid$ 3 and 5) traversing A-D. Their LSC interfaces (i.e., 101 and 102 at 10.0.0.1) are used within Table II to retrieve the IP addresses of the attached interfaces at the packet layer (i.e., 172.16.101.201 and 172.16.101.202 respectively). Such IP addresses are used within Table III to retrieve the PSC LSP Identifiers utilizing these interfaces (i.e., $LPid$ 13 and 11) and the related LSP Explicit Route Objects (ERO). Note that all these correlations are performed without involving the PCEs. Only after this stage, the NetAPP triggers a PCEP PCReq to request the PSC PCE to perform a Global Concurrent Optimization (GCO) of the identified involved LSPs (the three PSC LSP A-D, A-C and A-F), including the specific PSC links (A-D and A-C) within the eXclude Route Object (XRO). The PSC PCE then performs joint path computation and, in case of active stateful condition, it is also able to apply the computed changes, releasing the resources actually traversing the optical link A-D and redistributing them in the LSC layer (typically over-provisioned) according to the indicated GCO parameters.

IV. OPEN NETWORK DB IMPLEMENTATION AND EXPERIMENTAL DEMONSTRATION

In this section the ONDB architecture has been assessed through experimental demonstration, in order to verify the feasibility of the proposal in terms of scalability and effectiveness. In Sec. IV-A, a MySQL implementation of the ONDB employing link maintenance as a sequence of queries is detailed and its performance is evaluated as a function of the database size. In Sec. IV-B the ONDB architecture implementation, including the NetAPP of link maintenance, is validated in a multi-layer network testbed.

A. Open Network Database implementation results

The considered ONDB implementation has been configured on a MySQL platform (version 5.6) running on a remote FreeBSD server (AMD Opteron, 12-core CPU at 2.4GHz, RAM 16 GB). The ONDB internal structure, in accordance
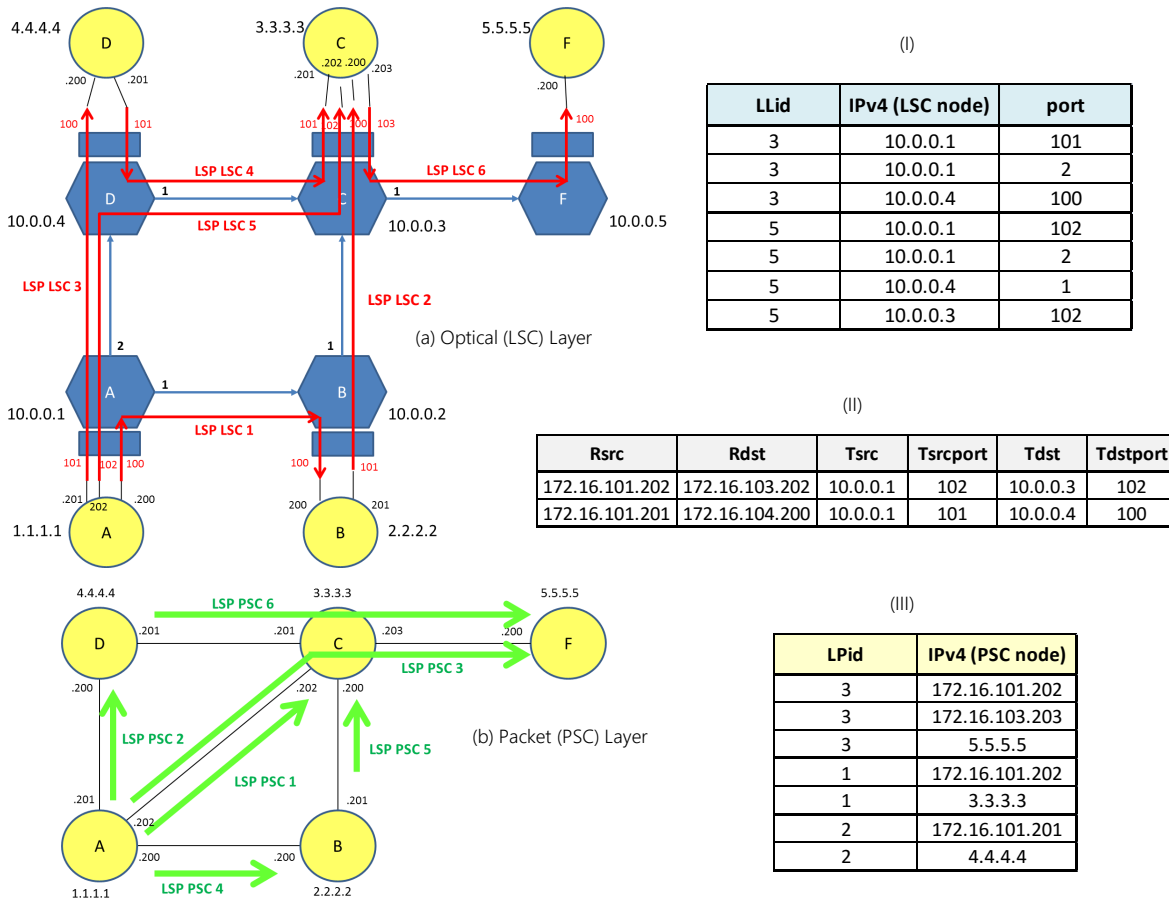
Figure 3. (a) Optical (LSC) network: ROADMs, active LSC LSPs and interconnected Routers; (b) Virtual network: PSC links and PSC LSPs; Table I: active LSC LSPs and traversed LSC interfaces; Table II: interconnection between LSC and PSC interfaces; Table III: active PSC LSPs and traversed PSC interfaces.
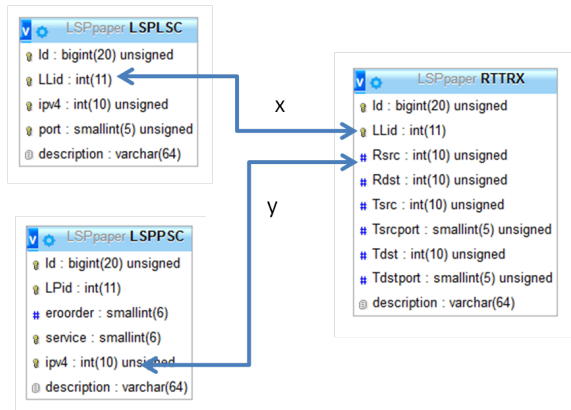
(I)

LLid	IPv4 (LSC node)	port
3	10.0.0.1	101
3	10.0.0.1	2
3	10.0.0.4	100
5	10.0.0.1	102
5	10.0.0.1	2
5	10.0.0.4	1
5	10.0.0.3	102

(II)

Rsrc	Rdst	Tsrc	Tsrcport	Tdst	Tdstport
172.16.101.202	172.16.103.202	10.0.0.1	102	10.0.0.3	102
172.16.101.201	172.16.104.200	10.0.0.1	101	10.0.0.4	100

(III)

LPid	IPv4 (PSC node)
3	172.16.101.202
3	172.16.103.203
3	5.5.5.5
1	172.16.101.202
1	3.3.3.3
2	172.16.101.201
2	4.4.4.4

Figure 4.   Implementation of the ONDB tables and correlation relationships.

Figure 5.   Standard and optimized query execution time to get lightpath list from the database.

with the use case description of Sec. III, has been organized into three tables: $LSPLSC$, $LSPPSC$, $RTTRX$.

The $LSPLSC$ table stores the description of the active LSC LSPs (i.e., lightpaths). Each table entry includes two identifiers: the unique identifier within the considered database table and the unique $LLid$ lightpath identifier within the network. The two identifiers univocally define the lightpath in the database. Moreover, each entry include the ERO subobject of the lightpath, expressed as unnumbered interface with router
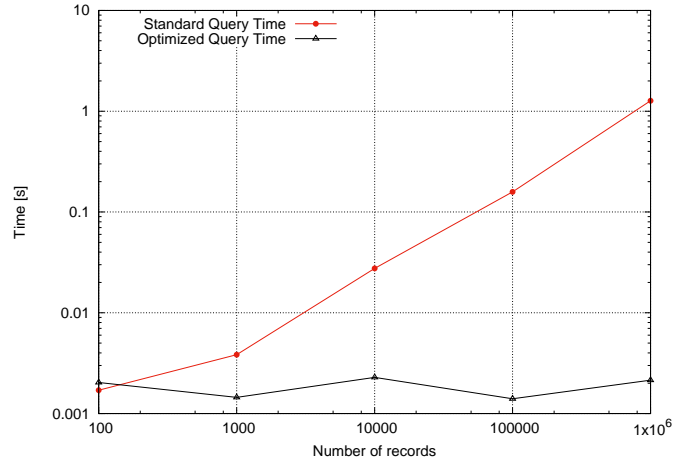
id ($IPv4$ address) and $port$ number attributes.

The $LSPPSC$ table stores the description of the PSC LSPs. The table entry identifies an ERO subobject, expressed as $IPv4$ address node, the sorting id inside the ERO, the LSP identifier which the entry belongs to (i.e., $LPid$) and a $service$ field indicating the service priority of the PSC LSP. The sorting
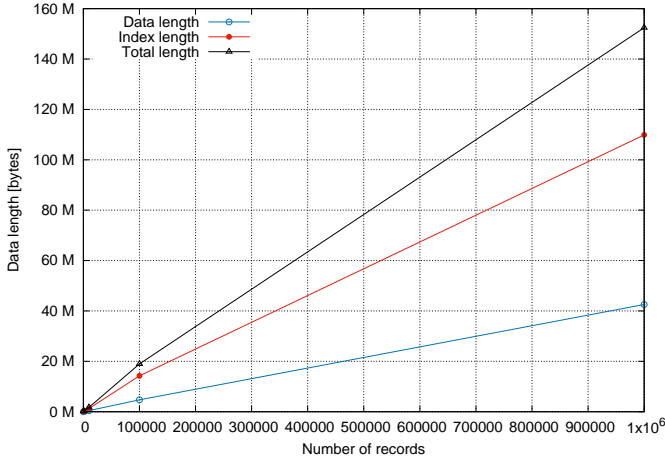
Figure 6. Storage resources used by the ONDB tables as a function of LSPPSC records size.

id, i.e. $eroorder$, is enclosed to allow enhanced inter-layer correlation and sorted output results.

Finally, the $RTTRX$ table stores the inventory of the unidirectional association between optical cards and router interfaces. Each entry is described by the source and destination IPv4 addresses of the router interfaces ($Rsrc$, $Rdst$) and the source and destination unnumbered interfaces addresses of the transponder cards ($Tsrc$, $Tdst$, $Tscrport$, $Tdstport$).

The correlation relationships between the three tables of the database are shown in Fig.4. In particular, the unique lightpath identification is enclosed in both $LSPLSC$ and $RTTRX$ tables (relationship X in the figure), thus allowing the match of a lightpath and its endpoint router card interface addresses. Then, the source router interface address is enclosed in both $RTTRX$ and in the LSPPSC entries (with the same IP address, in the match case) identifying the ERO subobject of a PSC LSP (relationship Y in the figure), thus allowing the match between the considered interface and the PSC LSPs that use such interface as ERO. Based on the value of $eroorder$, it is possible to detect whether the PSC LSP utilizes the selected lightpath and in which segment of the PSC LSP.

Following the link maintenance procedure, in order to get the list of PSC LSPs affected the maintenance of an optical link (for example, the link identified by the 10.0.0.1:2 unnumbered interface address) the following query executes a join between the three database tables.

```
1  SELECT LPid FROM LSPLSC, LSPPSC, RTTRX WHERE
   LSPLSC.ipv4=inet_aton("10.0.0.1") AND port=2 AND
       LSPLSC.LLid=RTTRX.LLid AND
3  LSPPSC.ipv4=Rsrc;
```

Then for each element of the list the following script queries the LSPPSC table to find the ordered list of the tuple composed by $eroorder$, $service$ and $IPv4$.

```
1  for i in $LPid_list do
     mysql —login—path=local LSPpaper −t −e "SELECT
        eroorder, service, inet_ntoa(ipv4) AS IPv4 FROM
        LSPPSC WHERE LPid=$i ORDER BY eroorder"
3  done
```

Therefore, by just relying on these two simple queries, the list of affected PSC LSPs including their ERO is provided.

In Fig. 5 the query set time is depicted as a function of the database size (plots are averaged over five repetitions per experiment). Two different database configurations have been considered: one standard containing the tables without any optimization, and the other with tables optimized through indexes. Executing the same set of query to get the LSP list, the former increases the query execution time while the latter keeps such time at a constant level (around 1ms).

Therefore, in order to achieve remarkable scalability, it is necessary to optimize the tables creating the appropriate indexes of the field used to select the desired records. The indexes used are *b-tree* type [20]. B-tree is a self-balancing tree data structure that keeps data sorted and allows searches, sequential access, insertions, and deletions in logarithmic time. In fact, accessing an entry in a b-tree of n elements is in logarithmic amortized time O($\log(n)$): this results in extremely reduced and stable query time, even when the ONDB tables have million of records.

Along with the query time, also the storage used by the database is evaluated to figure out the resource needed on the selected hardware. The LSPLSC and RTTRX tables gross size is set to 16 KB, while the LSPPSC table size grows with the number of record inserted up to a value of one million (i.e., representing a network with one million of active PSC LSPs). Fig. 6 shows the storage resources used by the ONDB tables for the two different database configurations. Both configurations require limited storage capabilities, in particular the amount of required storage grows linearly with the number of the entries of the database, even in the case of the optimized configuration, which requires additional resources to store the indexes (with an increasing factor of around 4 times with respect to standard configuration). However, such values do not represent a critical parameter (e.g., for 1M entries representing a very large LSP database, only 160 MB are required for the optimized configuration, which can be handled directly in RAM) also considering the huge amount of included network information and the impressive query time performance achieved by the optimized configuration, as shown in Fig. 5.

B. Experimental validation

The effectiveness of the ONDB-based architecture has been evaluated in a network testbed implementing the key portions of the sample network shown in Fig. 3 and following the architectural proposal as described in Sec. III. The testbed includes two PCE-based controllers (i.e., the LSC controller and the PSC controller), the maintenance NetAPP and the ONDB implementation of Sec. IV-A. The C++-based LSC PCE implementation derives from [10] and supports stateful provisioning and update of flex-grid Elastic Optical Networks (EON) lightpaths with transponder extensions as in [21]. The considered C++ PSC PCE operates on IP/MPLS networks and is derived from the implementation described in [22], also
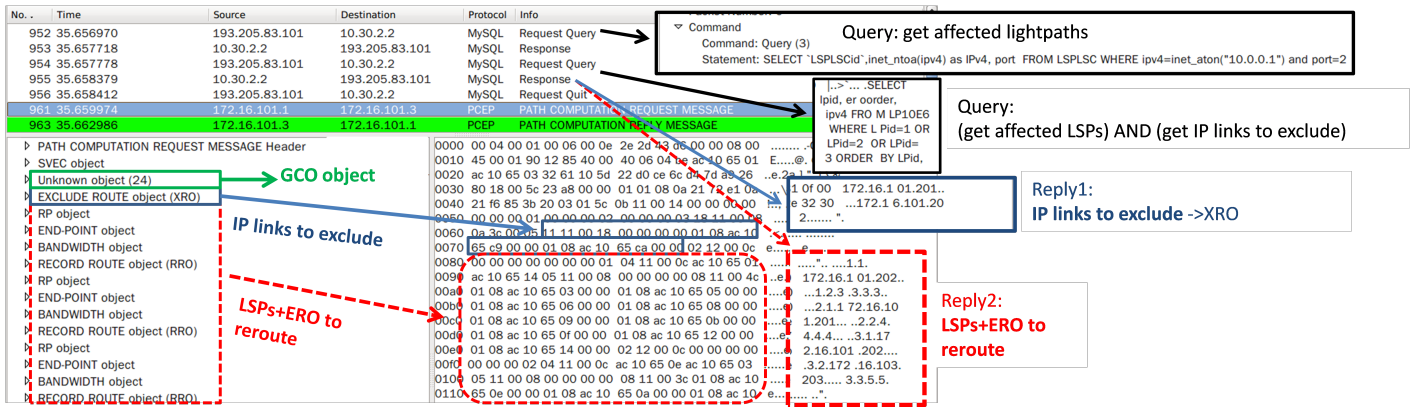
Figure 7. Message exchange by the Maintenance NetAPP with the ONDB and PSC PCE.

supporting Segment Routing extensions. The PSC PCE implements a northbound PCEP interface with standard MPLS-TE capabilities for communications with NetAPPs and the ONDB and a southbound interface towards network nodes supporting ERO Segment Routing ERO extensions, in order to enable segment list configuration at source node. The maintenance NetAPP has been implemented in C++. It includes a MySQL client for the communication to the ONDB and a Path Computation Client (PCC) for the communication to the PSC PCE for triggering the rerouting of selected PSC LSPs.

Fig. 7 shows the Wireshark capture of the full maintenance procedure of link A-D, as detailed in Sec. III-A. Packets 952 to 956 report the queries and responses between the NetAPP (IP addresses 193.205.83.101 and 172.16.101.1) and the ONDB (IP address 10.30.2.2) through the SQL protocol. In particular, the first query (packets 952-953) retrieves the affected LSC LSPs, the second query (954-955) matches the related IP endpoint interface and the IP links to be excluded, thus retrieving the PSC LSPs subject to rerouting and their ERO. Then, packets 961-963 report the subsequent PCEP session between the NetAPP and the PSC LSP (IP address 172.16.101.103), applying the rerouting of the retrieved PSC LSP A-D, A-C and A-F, subject to GCO, and the XRO including addresses 172.16.101.201 and 172.16.101.202 (i.e., router links A-D and A-C to be excluded from rerouting computation). The overall control plane procedure is completed in less than 10ms, including around 4 ms to complete the considered queries.

To assess the scalability performance also in the network scenario, 100k PSC LSP over a 100nodes mesh network have been also emulated and inserted within table LSPPSC (reaching around 1million entries). In this case, the time required to perform the last query on link to LSP correlation has increased from 4ms to 1271ms if database is not optimized, while the query time is still 4ms if the database is optimized, thus substantially confirming scalability results of Sec. IV-A.

V. CONCLUSIONS

An Open Network Database for interconnected multi-layer TE networks was proposed as novel control and management network element within an application-based architecture, implemented and demonstrated for the automated maintenance use case. The Open Network Database exploits standard YANG definition of TE parameters. Such parameters under different domains of visibility are successfully correlated and may be suitably used at the desired layer to perform complex re-optimization and computation procedures. The ONDB solution significantly simplifies the implementation of network applications. For example, the network application for optical link maintenance was deployed by just relying on two simple queries. The experimental evaluation demonstrated that the proposed architecture and the implementation of the ONDB was able to achieve inter-layer correlation information efficiently and the overall optical link maintenance was successfully performed, at the control plane level, in few milliseconds with limited storage resources and impressive scalability performance.

REFERENCES

[1] F. Paolucci, F. Cugini, G. Cecchetti, and P. Castoldi, "Open database for interconnected traffic engineered multi-layer networks," in *Optical Fiber Communication Conference*. OSA, 2016, p. Th4G.5.

[2] F. Paolucci, F. Cugini, A. Giorgetti, N. Sambo, and P. Castoldi, "A survey on the path computation element (PCE) architecture," *Communications Surveys Tutorials, IEEE*, vol. 15, no. 4, pp. 1819–1841, Fourth 2013.

[3] O. G. de Dios, R. Casellas, R. Morro, F. Paolucci, V. López, R. Martínez, R. Munoz, R. Vilalta, and P. Castoldi, "Multipartner demonstration of BGP-LS-enabled multidomain EON control and instantiation with H-PCE [invited]," *J. Opt. Commun. Netw.*, vol. 7, no. 12, pp. B153–B162, Dec 2015.

[4] L. M. Vaquero, L. Rodero-Merino, and R. Buyya, "Dynamically scaling applications in the cloud," *ACM SIGCOMM Computer Communication Review*, vol. 41, no. 1, pp. 45–52, 2011.

[5] X. Liu, I. Bryskin, V. Beeram, T. Saad, H. Shah, and O. Gonzalez de Dios, "YANG data model for TE topologies, draft-ietf-teas-yang-te-topo-05," *IETF, TEAS Working Group*, Jul 2016.

[6] F. Paolucci and F. Cugini, "The path computation element (pce)," *Handbook of Research on Redesigning the Future of Internet Architectures*, p. 237, 2015.

[7] O. Gerstel, C. Filsfils, T. Telkamp, M. Gunkel, M. Horneffer, V. Lopez, and A. Mayoral, "Multi-layer capacity planning for ip-optical networks," *Commun. Mag., IEEE*, vol. 52, no. 1, pp. 44–51, January 2014.

[8] S. Zhang, M. Tornatore, G. Shen, J. Zhang, and B. Mukherjee, "Evolving traffic grooming in multi-layer flexible-grid optical networks with software-defined elasticity," *Lightwave Technology, Journal of*, vol. 32, no. 16, pp. 2905–2914, Aug 2014.

[9] O. de Dios, V. Lopez, M. Cuaresma, F. Munoz, M. Chamania, and A. Jukan, "Coordinated computation and setup of multi-layer paths via inter-layer pce communication: standards, interoperability and deployment," *Commun. Mag., IEEE*, vol. 51, no. 12, pp. 144–154, Dec. 2013.

[10] L. Gifre, F. Paolucci, L. Velasco, A. Aguado, F. Cugini, P. Castoldi, and V. Lopez, "First experimental assessment of abno-driven in-operation flexgrid network re-optimization," *Lightwave Technology, Journal of*, vol. 33, no. 3, pp. 618–624, Feb 2015.

[11] A. Castro, R. Martinez, R. Casellas, L. Velasco, R. Munoz, R. Vilalta, and J. Comellas, "Experimental assessment of bulk path restoration in multi-layer networks using pce-based global concurrent optimization," *Lightwave Technology, Journal of*, vol. 32, no. 1, pp. 81–90, Jan 2014.

[12] R. Munoz, R. Vilalta, R. Casellas, R. Martinez, F. Francois, M. Channegowda, A. Hammad, S. Peng, R. Nejabati, D. Simeonidou, N. Yoshikane, T. Tsuritani, V. Lopez, and A. Autenrieth, "Transport network orchestration for end-to-end multilayer provisioning across heterogeneous sdn/openflow and gmpls/pce control domains," *Lightwave Technology, Journal of*, vol. 33, no. 8, pp. 1540–1548, April 2015.

[13] Y. Yoshida, A. Maruta, K.-I. Kitayama, M. Nishihara, T. Tanaka, T. Taka-hara, J. Rasmussen, N. Yoshikane, T. Tsuritani, I. Morita, S. Yan, Y. Shu, Y. Yan, R. Nejabati, G. Zervas, D. Simeonidou, R. Vilalta, R. Munoz, R. Casellas, R. Martinez, A. Aguado, V. Lopez, and J. Marhuenda, "Sdn-based network orchestration of variable-capacity optical packet switching network over programmable flexi-grid elastic optical path network," *Lightwave Technology, Journal of*, vol. 33, no. 3, pp. 609–617, Feb 2015.

[14] O. Gerstel and V. Lopez, "The need for sdn in orchestration of ip over optical multi-vendor networks," in *Optical Communication (ECOC), 2015 European Conference on*, Sept 2015, pp. 1–3.

[15] D. King and A. Farrel, "The Application of the Path Computation Element Architecture to the Determination of a Sequence of Domains in MPLS and GMPLS," *IETF, RFC 6805*, Nov 2012.

[16] M. Chamania, F. Carpio, and A. Jukan, "Engineering the first open source implementation of the multi-layer network topology databases," in *Optical Communication (ECOC), 2015 European Conference on*, Sept 2015, pp. 1–3.

[17] L. Gifre, L. M. Contreras, V. Lopez, and L. Velasco, "Big data analytics in support of virtual network topology adaptability," in *Optical Fiber Communication Conference*. OSA, 2016, p. W3F.6.

[18] F. Morales, M. Ruiz, and L. Velasco, "Virtual network topology recon-figuration based on big data analytics for traffic prediction," in *Optical Fiber Communication Conference*. OSA, 2016, p. Th3I.5.

[19] M. Chen, R. Zhang, and X. Duan, "OSPF Extensions in Support of Inter-Autonomous System (AS) MPLS and GMPLS Traffic Engineering," *IETF, RFC 5392*, Jan 2009.

[20] D. Comer, "Ubiquitous b-tree," *ACM Comput. Surv.*, vol. 11, no. 2, pp. 121–137, Jun. 1979.

[21] R. Martinez, R. Casellas, R. Vilalta, and R. Munoz, "Experimental assessment of gmpls/pce-controlled multi-flow optical transponders in flexgrid networks," in *OFC Conf.*, March 2015.

[22] A. Sgambelluri, F. Paolucci, A. Giorgetti, F. Cugini, and P. Castoldi, "Experimental demonstration of segment routing," *Lightwave Technology, Journal of*, vol. 34, no. 1, pp. 205–212, 2016.