Towards a Transport SDN for Carriers Networks: An Evolutionary Perspective

Victor Lopez, Luis Miguel Contreras, Oscar Gonzalez de Dios, Juan Pedro Fernandez Palacios
Telefonica I+D/GCTO
Ronda de la Comunicación, s/n
Madrid, Spain

Abstract— **Network operators have worked in interoperable scenarios for transport network from several years. The main motivation is to have a rich ecosystem, which encourages the competition to have more efficient network solutions. The bandwidth increment in the transport network puts a lot of pressure to have revenues on an environment where the end-user is willing to pay less and less for the service.**

Software Define Networks presents a new hope to achieve such desire multi-vendor interoperability. The aim of this paper is to present some architecture to enable interoperability in transport networks. Some of these architectures are market ready and they have been tried in the field, while there are some approaches which are under standardization.

Keywords—SDN; control plane; optical networks; interoperability

I. INTRODUCTION

Multi-vendor interoperability is an old aim from the network operators for transport networks. From the operator's perspective, the interoperability encourages the competition between the vendors, which leads on a rich environment and more efficient network solutions. Various bodies such as the International Telecommunication Union (ITU-T) [1] and the Internet Engineering Task Force (IETF) [2] are working toward developing standard specifications for control and data plane functionalities which facilitate interoperability between different vendors. However, in practice vendors compete with each other trying to gain higher market share by implementing specific non-standard functionalities in their equipment, a fact that unquestionably leads to customer lock-in. This in turn creates a buyers dependency on the seller.

The explosion of broadband connections imposes an unprecedented traffic growth in telecommunication networks with very high cumulative annual growth rates. Such bandwidth increment puts a lot of pressure to have revenues on an environment where the end-user is willing to pay less and less for the service. The operators have to efficiently deploy architectures which are flexible and adapts to the traffic demands in this scenario dynamic interoperability is a must. So far, residential Internet access and business services have driven the connectivity demands in the network. For these customers, the end-to-end service provisioning can be around several weeks for network construction and days to deliver an equipment to the customer, which is acceptable for most end-users. Nevertheless, data center services are becoming an essential component in the traffic sources for network operators [3]. Besides, the bandwidth requirements for cloud services are much more variable [4] than traditional services, and their network usage highly depends on the kind of service installed by the user in the cloud. Measurements used in [4] are for intra-data center traffic, but businesses are migrating from a private cloud paradigm to a hybrid cloud [5]. The internal infrastructure of a company must coordinate with external resources (public or private) to enable a hybrid cloud scenario. This forecast situation hints an increment in the number of service provisioning demands due to the inter-data center variable traffic. Consequently, network operators require adapting their current static networks in a dynamic end-to-end scenario.

The data plane technologies for optical networks are evolving towards an adaptive scenario. Existing DWDM optical communication systems divide the C-band optical spectrum into discrete bands, spaced usually by 50 or 100 GHz, which are normalized in the ITU G.694.1 Grid [6]. However, Elastic Optical Networks (EON) [7] enables the optical spectrum to be used in a more flexible way, where chunks of spectrum can be defined more arbitrarily than currently. The ITU has extended [6] and [7] to include the concept of flexible grid. Moreover, Optical Transport Network (OTN) has evolved to define ODUFlex [8] as a solution to adapt the bandwidth of the tributary ports and carry out a more effective wavelength occupation. Even though the data plane technologies are evolving, there is a need to have dynamic solutions, at least from the control plane to enable multi-vendor support.

This article presents alternatives to control plane interoperability. Moreover, it justifies why SDN can be the solution to enable such multi-vendor scenario. The remaining of this paper is as follows: Section II presents which is the scope of interoperability and which are the topics cover in this paper. Section III presents solutions based on UNI technologies. Section IV shows the Path Computation Element (PCE) as an entity to enable interoperability in some scenarios. Section V describes the Transport API as a key interface for SDN scenarios. Finally, Section VI compares the different alternatives and concludes this work.

II. INTER-OPERABILITY IN TRANSPORT NETWORKS

There are two well-defined planes in transport networks: the data and the control plane [9]. The data plane is used for the transmission of the information, while the control plane is in charge of the decentralized tasks issues such as the exchange of routing information, monitoring of link state and the set up and tear down of connections.

From the data plane perspective, four main scenarios can be defined (Fig. 1). The monovendor scenario is the most deployed one, where all network elements are bought to the same vendor. Additionally, the Network Management System (NMS) and the control plane are proprietary from this vendor. The second scenario is based on the alien wavelength or black link concept [10]. An alien wavelength is a lightpath, but the transponders do not belong to the same vendor than the Optical Cross Connects (OXCs). On this scenario there are two vendors, which do not interoperate, but they can work together. The third scenario is the multi-ROADM scenario, where the OXC or ROADMs belong to multiple suppliers. The fourth one is the interoperable scenario, where there can be transponders from multiple vendors. This scenario is an old aim of operators, but the Forward Error Correction (FEC) is based on proprietary developments and there are not standard FECs with enough performance to satisfy the needs of current networks.

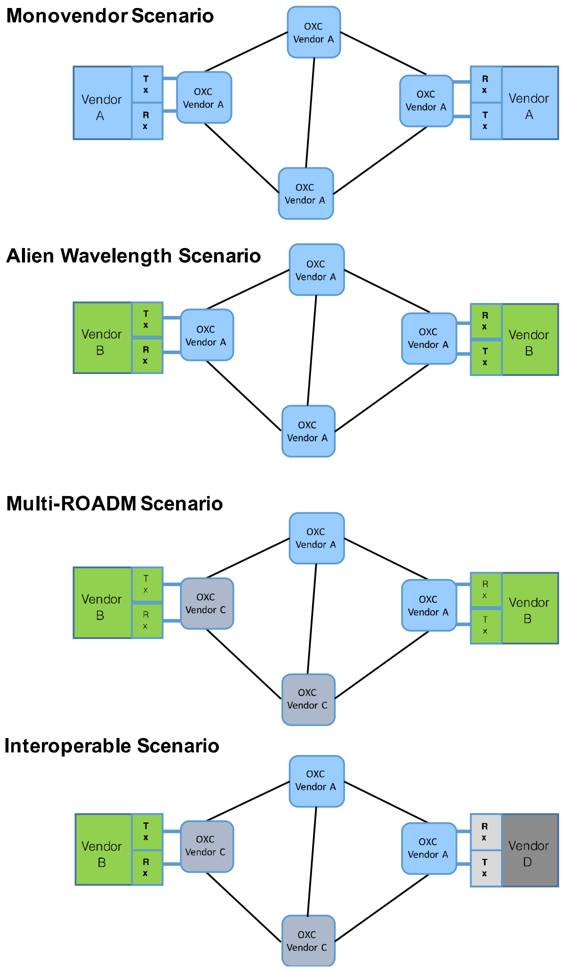

Fig. 1. Data plane interoperability scenarios

As the cost of the optical networks is mainly on the transponders [11], the two scenarios, which are more interesting for operators, are the alien wavelength and the interoperable. The alien wavelength scenario is a reality. It is deployed in production networks without the control plane support. From our perspective, the FEC is still an important barrier to deploy these scenarios.

From the control plane perspective, there are two basic scenarios provisioning for optical networks. The initial one is the L0/L1 dynamic service provisioning on a single domain. For this scenario, the customer requires some bandwidth between two end points in the network. Provisioning capability enables the creation, deletion and update of connections in the network. However, to cope with the service requirements, the capability must support explicit routes, route restrictions, service resilience and traffic engineering parameters such as bandwidth and latency are desirable. This case study is covered by the NMS from the vendors, but it lacks of standard interfaces to easily integrate with the systems of the operators.

The second scenario and more generic is the multi-vendor E2E L0/L1 scenario provisioning, see Fig. 2. The case study is the same, but it requires the interoperability between the vendors.

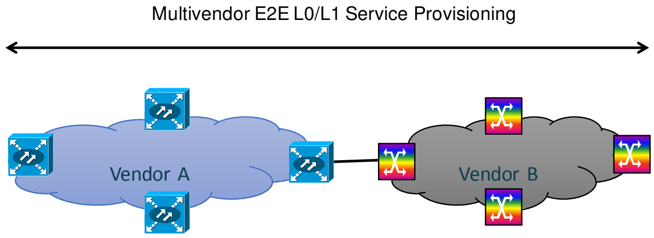

Fig. 2. Multi-vendor provisioning

Once the most basic scenarios are covered, the network operators can create on top of it another use cases. We think they are important, as they are atomic functions to support other more complex scenarios. Some examples of these use cases are elastic bandwidth provisioning, datacenter interconnection, network as a service, optical VPNs, etc. The rest of the paper covers the control plane interoperability aspects. It will present some alternatives to have dynamic operation of the optical network.

III. CONTROL PLANE PROVISIONING VIA UNI

The features of the control plane are achieved through the use of the GMPLS architecture, which extends MPLS to other switching types, which are: Packet Switch Capable (PSC), Layer-2 Switch Capable (L2SC), Time-Division Multiplex Capable (TDM) and Lambda Switch Capable (LSC). GMPLS also enables the organization of such switching types hierarchically, which will be very useful for multi-layer networks.

GMPLS architecture has the following features: (1) auto-discovery of the network topology, (2) reporting all network resources and management of the available links, (3) routing and (4) configuration of the chosen paths. However, GMPLS is not a protocol but a set of protocols that handle the functions previously named. The main protocols of GMPLS are:

- RSVP-TE (Resource reSerVation Protocol): is responsible of signalling the links in the computed paths for resource reservation of data flows.
- OSPF-TE (Open Short Path First): is responsible of the dissemination of the information of the topology and the traffic engineering, and construct a Traffic engineering Database (TED). These functions enable the routing at each node in the network.
- LMP (Link Management Protocol): is responsible of links management. Monitors the proper functioning of the links and checks the connectivity between adjacent nodes.

The control plane has three main interfaces the Internal-Network to Network Interface (I-NNI), the External NNI (E-NNI) and the User to Network Interfaces (UNI). The UNI is the control plane interface from the routers to the optical equipment, as shown in Fig. 3.

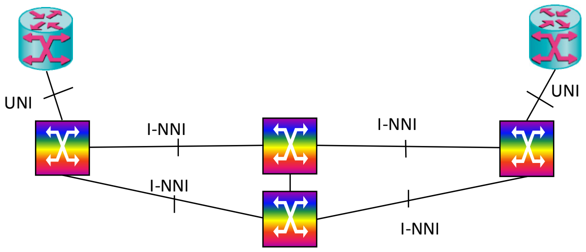

Fig. 3. UNI interoperability scenario

There are different models for the control plane to be used in this environment with UNI. The "peer" model and "overlay" model are the two main approaches, but current operator's networks typically use "overlay" because of the lack of multi-vendor interoperability between layers. The "overlay" model for multi-layer networks works as a client/server model. The IP/MPLS upper layer can be considered as the client layer, while the Transport layer works as the server layer. In this context, the client layer requests a connection to the transport layer through the UNI, which is an interface standardized by the IETF.

UNI works with RSVP-TE (and the extensions to GMPLS) for resource reservation, and OSPF-TE to notify the new adjacencies in the client layer after the resource reservation in the transport layer. To set-up a connection via UNI, the router can signal the following parameters in the request:

- *The switching capability for the LSP.* When a new point-to-point connection between two routers is created, this connection needs to use an interface in both routers. Each interface recognizes different types of switching capabilities, which means different technologies that characterizes the established connection.
- *The bandwidth required.*
- *The path to reserve given by the ERO (Explicit Route Object).* This object defines the hops in the path. In the actual model of separated layers, the router that makes the path request does not have all information of the network so it can not specify the path in the lower layer. The ERO in the multi-layer case needs to include the identifiers of the TE-Links of the router making the request, the identifier of the connected terminal in the lower layer and the identifier of the target terminals (both the terminal in the transport layer and the one in the IP/MPLS layer).

In the case of having colored interfaces in the routers, the control plane may change. As there is not a client/server model, the terminals work in both ways interpreting requests as server and sending requests as client. This type of working model figures in the standards as the interface I-NNI (Internal-Network to Network Interface), instead of UNI. The use of I-NNI is reduced to one domain, so there is an integrated single control plane in each terminal. This model of control plane is known as peer model and all the terminals in the network work with the same instance of the control plane. On this scenario the router has to signal the lambda to the optical layer to provision the connection. The authors in [12] demonstrate the interoperability between vendors using colored interfaces in the routers. Based on this work, we can assume that the L0/L1 dynamic service provisioning on a single domain is supported by this architecture.

Even though UNI allows reserving resources across the transport layer, it has not the multi-layer topology knowledge which is needed to plan in an optimal way the network creation process in any network. A central entity is needed to have this global view of the network.

IV. PATH COMPUTATION ELEMENT

The path computation element is "an entity (component, application, or network node) that is capable of computing a network path or route based on a network graph and applying computational constraints" [13].

The PCE requires information about the network state for the path computation. Traditionally, such collection process is done using the link state protocols. GMPLS networks compute the paths based on the state information transmitted in a distributed way via the IGP, OSPF-traffic engineering (TE) or ISIS-TE. IGP-TE protocols exchange two kinds of information which is link state (LS) and TE, so two databases exist namely the link state database (LSDB) and the traffic engineering database (TED). The TED is a subset of the LSDB. The IP protocol only uses LS information, but MPLS and GMPLS can also use TE information next to LS information. However, as PCE extends this computation process with more complex algorithms, it requires new information that may not be present at each network node.

A. *Active Path Computation Element*

The original definition of the PCE was stateless in the sense that a network element queries the PCE to obtain the path for a connection. A stateful PCE knows which are the connections on the network and can make decisions based on this information. An active stateful PCE goes a step further and it is a path computation entity, which can maintain the sessions for the LSPs and can even create LSPs in the network.

With this approach the network operator enters in the PCE and it can set-up a connection on the network via a PCInitate message which is sent to the network elements. The lightpath is signaled using RSVP with the constraints sent by the PCE in the PCInitiate message. Once the message is RSVP Resv message is received, the network element sends a PCReport to the PCE.

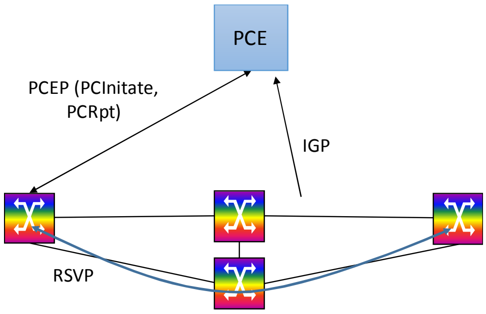

Fig. 4. Active PCE interoperability scenario

Telefonica has developed an open source implementation of this scenario. The Telefonica Netphony suite is a set of java based libraries that enable to create a PCE/GMPLS based control plane. It comprises a set of components, distributed as jar files, which are hosted in publicly available github repositories. The netphony-pce [14] contains the implementation of a Java based Path Computation Element. The repository also contains two Path Computation Clients, QuickClient, which by means of command line options can generate and receive PCEP messages. The GMPLS network emulator is part of the netphony-gmpls-emulator [15].

B. Hierarchical Path Computation Element

There are scenarios where the network operator wants to connect two different domains. For this multi-domain scenarios, the Hierarchical PCE (H-PCE) architecture achieves a lower blocking probability and increases the network utilization [16].

In this architecture there is a parent PCE and some child PCEs, and they are organized in multiple levels [17]. The parent PCE does not have information of the whole network, but is only aware of the connectivity among the domains and provide coordination to the child PCEs. The path request is sent to the parent PCE, which selects a set of candidate domain paths and sends requests to the child PCEs responsible for these domains. Then the parent PCE selects the best solution and it is transmitted to the source PCE. This hierarchical model fits with the model for the Automatic Switched Optical Network (ASON), since the networks are composed by sub-networks and the routing areas have relationship between peers.

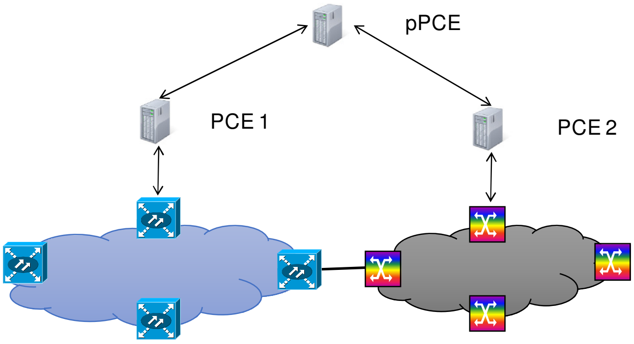

Fig. 5. Hierarchical PCE scenarios

From a bottom-top approach based on Fig. 4, each domain deploys an extended GMPLS control plane including, notably, the OSPF-TE protocol for topology dissemination and the RSVP-TE protocol for the signalling of the Label Switched Paths (LSPs). On top of the GMPLS control plane, each domain deploys an active stateful Path Computation Element (AS-PCE), for the purposes of both optimal path computation and service provisioning within its domain. Thus, multi-domain path computation and provisioning is carried out by means of a Hierarchical Path Computation Element (H-PCE), with the parent PCE (pPCE), coordinating the procedures between children PCEs (cPCE): the interface between pPCE and domain cPCEs (based on PCEP protocol) is thus used by the pPCE for path computation and instantiation.

The netphony-pce [14] contains the implementation of a Java based parent Path Computation Element. Recently this implementation has been demonstrated on a flexigrid environment with implementations from multiple institutions in each network domain [18].

V. TRANSPORT API FOR SDN ENVIRONMENTS

Software Define Network (SDN) concept is based on the idea of decoupling the control and data plane. This concept is inherent to the optical networks as the signaling is done always via an out of band channel. The NMS was the controller, which configures the optical equipment and there was not a standard interface from the NMS to the devices. The utilization of open and standard interfaces to enable interoperability is the first advantage of this architecture.

Most of the solutions in the market for SDN are based on single domain and mono vendor solutions. However, network operators usually have in place multiple technologies (provided by different vendors) in their networks and multiple domains to cope with administrative and regional organizations. A single SDN controller cannot configure the whole network of an operator for scalability and reliability issues. This is even more complicated when considering an architecture that should deal with multiple South Bound Interfaces (SBI) like OpenFlow and GMPLS. The ONF proposes a hierarchical architecture that fits with the multi-vendor/multi-domain scenario. In this approach, there are multiple SDN controllers interacting with a SDN orchestrator hierarchically placed on top of them.

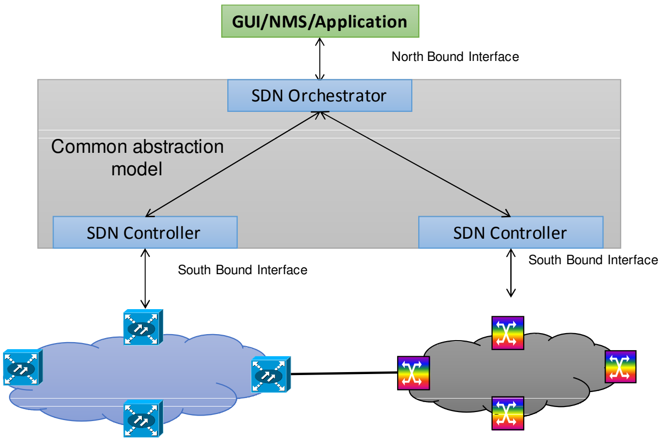

Fig. 6. Hierarchical SDN architecture

The implementation of the controller depends on the vendor, but there are open source approaches like ODL, ONOS or netphony-abno [19]. Even though there is a debate in the operators to use or not open source implementations, there is a wide agreement that the interfaces to the controllers and orchestrator must be standard.

A. *Introduction to the T-API*

The Transport API (T-API) is under definition within the optical transport group in the ONF [20]. The T-API abstracts four main functionalities: (1) Network Topology, (2) Connectivity Requests, (3) Path Computation and (4) Network Virtualization to a set of Service interfaces.

Network Topology functionality requires, at a minimum, that the interface exports network topology information with unique identifiers. However, network identifiers (such as IPv4 or datapath-IDs) help to carry out path computation and to integrate the nodes for an end-to-end scenario. Further, the controllers can provide information about the links in the domain (physical or virtual), their utilization or even information about physical impairments, which the orchestrator may apply to a physical impairments computation model. It is clear that the more information is shared, the less abstracted the network appears. The Connectivity Requests functionality enables the set-up, tear down and modification of connections in the network. Its most basic feature is to set up a point-to-point connection between two locations. However, there are other characteristics that a client interface can have like (a) excluding or including nodes/links for traffic engineering, (b) defining the protection level, (c) defining its bandwidth or (d) defining its disjointness from another connection. The Path Computation function is a critical and fundamental feature because individual controllers in each domain are only able to share abstracted information that is local to their domain. An orchestrator with its global end-to-end view can optimize end-to-end connections that individual controllers cannot configure. Without a path computation interface, the orchestrator is limited to carrying out a crank-back process. Finally, Network Virtualization services enables to expose a subset of the network resources to different tenants.

The T-API is positioned as the NBI of the SDN orchestrator or as the common abstraction model between the SDN orchestrator and the controllers.

B. *Open issues within the T-API*

Event though the T-API is a very good candidate technology, there are some open issues. The first comment is the support for optical parameters is not covered in detail. As an example, the ConnectionEndPoint description is shown in Fig. 7. As it can be seen, there is only an indication of the layer of the EndPoint. However, an important feature would be to know whether there are tunable transponders to validate, if a new service can be provisioned.

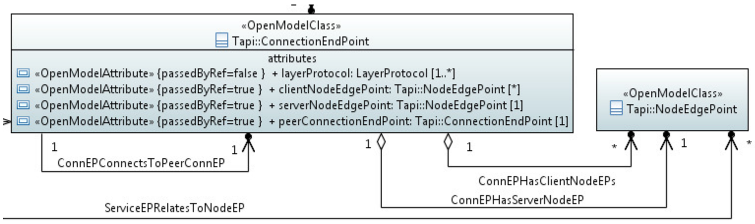

Fig. 7. UML description of the T-API[1]

Let us assume another scenario where the Transport API could be use. The SDN controller can optimize the total power consumption in a wireless transport network. To do so, the controller disables underlying physical ports of wireless L1 LAG (also unknown as a “Link Aggregation Group” or a “Multi Radio Group”, or a “Physical Link Aggregation”, or “Radio Link Bonding”) links when the utilization is below certain thresholds. This case study was proposed within the ONF Wireless Transport Group [21]. This scenario was successfully carried out. The PoC was held in October, 2015, in Madrid, Spain. Telefónica and IMDEA Networks together hosted the PoC. Six vendors participated in the tests, providing a variety of types of equipment and software consistent with the SDN architecture, leveraging ONOS as controller. However T-API was not used, as, right now, the definition of the T-API does not support the technology specific parameters neither for wireless networks, nor for optical networks.

Finally, let us assume two optical use cases: multi-domain failure recovery and congestion detection [22]. Current services require fast recovery from network failures (<50 ms) [17], this is possible in optical network using protection at the optical level. The T-API allows a mono-domain protection, but orchestrator should be able to receive failure alarms and start recovery workflows to create new E2E paths. Similarly, real time monitoring and dynamic network reconfiguration have become increasingly attractive due to the growing bandwidth on demand and the latest standards and protocols that act as enablers to provide more flexibility in the network design. The T-API does not support any monitoring information exchange to fulfill with the use cases previously mentioned.

---

[1] https://raw.githubusercontent.com/OpenNetworkingFoundation/ONFOpenTransport/develop/TransportAPI/UML/figures/Connectivity_Service_Overview.PNG

The orchestrator requires to implement an Operation, Administration and Maintenance (OAM) handler that receives failure alarms, triggers an internal workflow, and allows the orchestrator to obtain information about the affected services and to configure new E2E connections. Our proposal should have an T-API which can respond in both directions. It can have a reactive part, which would be based on websockets subscriptions, and proactive part of the API, which can use a RESTful interface as it is defined right now in the T-API.

VI. SUMMARY AND CONCLUSIONS

The SDN architecture proposed here enables automated and simplified network service provisioning through different vendors, network segments (metro, core, data center…) and technologies (IP/MPLS, optical, OpenFlow…)

Such automation and simplification could be achieved by applying complementary measures:

- Network configuration points minimization by transferring multidomain and multilayer provisioning functionalities from NMS to a control plane.
- PCEP as main SBI. Telefonica has developed an open source H-PCE implementation enabling interoperability. The Telefonica Netphony suite is a set of java based libraries that enable to create a PCE/GMPLS based control plane.
- Transport API as NBI. This interface should export network topology information with unique identifiers.

As future work, we propose to work on two main topics to work related with the T-API interface. The first one consists on extend the T-API to define in more detail the physical details of two main transport technologies optics and microwaves. On the other hand, the T-API does not cover OAM information. The failure management and congestion detection are topics that an SDN framework should solve. Therefore, we propose to use a interface that can support a subscription approach, where the orchestration listen to the events that the controllers can send.

ACKNOWLEDGMENT

The work on this paper is partially funded by the European Commission within the H2020 Research and Innovation program, ACINO project, Grant Number 645127, www.acino.eu and EU FP7 STRAUSS (FP7-ICT-2013-EU-Japan 608528).

REFERENCES

[1] The Internet Engineering Task Force Official Web Site. http://www.ietf.org/.

[2] The International Telecommunication Union Official Web Site. http://www.itu.int/.

[3] L. M. Contreras, V. López, O. González de Dios, A. Tovar, F. Muñoz, A. Azañón, J.P. Fernández-Palacios, J. Folgueira: *Towards Cloud-Ready Transport Networks*, in IEEE Communications Magazine, September 2012, Vol. 50, Issue. 9.

[4] Hitesh Ballani, et. all, “Towards predictable datacenter networks”, SIGCOMM Comput. Commun. Rev. 41, 4 (August 2011), 242-253.

[5] International Telecommunications Union, "Architecture of Optical Transport Networks," ITU-T Recommendation G.872, October 2012.

[6] M. Jinno, H. Takara, H. Watanabe, and Y. Sone, "Distance-Adaptive Spectrum Resource Allocation in Spectrum-Sliced Elastic Optical Path Network," IEEE Communications Magazine, August 2010.

[7] International Telecommunications Union, "Spectral grids for WDM applications: DWDM frequency grid," ITU-T Recommendation G.694.1, February 2012.

[8] Recommendation ITU-T G.7044, Hitless adjustment of ODUflex (GFP), 2011.

[9] M. N. Ellanti, S. S. Gorshe, L. G. Raman, and W. D. Grover, Next Generation Transport Networks: Data, Management, and Control Planes, 1st ed. Springer, April 2005.

[10] International Telecommunications Union, "Amplified multichannel dense wavelength division multiplexing applications with single channel optical interfaces", ITU-T Recommendation G.698.2, November 2009.

[11] F. Rambach, B. Konrad, L. Dembeck, U. Gebhard, M. Gunkel, M. Quagliotti, L. Serra, V. López, “A Multilayer Cost Model for Metro/Core Networks,” in Journal of Optical Communications and Networking, vol. 5, no. 3, pp. 210-225, March 2013.

[12] V. López, Z. Ali, S. Neidlinger, O. Gonzalez de Dios, D. Patton, T. Remiszewski, P. Drózda, J.P. Fernández-Palacios: Demonstration of control plane interoperability with integrated optical ports in multi-vendor scenarios, in European Conference on Optical Communication (ECOC), Sep 2014.

[13] A. Farrel, J.P. Vasseur, and J. Ash, “A path computation element (PCE)-based architecture,” IETF RFC 4655, pp. 1–40, August 2006. http://tools.ietf.org/html/rfc4655.

[14] Netphony-pce, https://github.com/telefonicaid/netphony-pce

[15] Netphony-gmpls-emulator, https://github.com/telefonicaid/netphony-gmpls-emulator

[16] D. Álvarez, V. López, J.L. Añamuro, J.E. López de Vergara, O. González de Dios and J. Aracil: Utilization of Temporary Reservation of Path Computed Resources for Multi-Domain PCE Protocols in WDM Networks, in the proceeding of Network of the Future conference, Nov, 2011.

[17] RFC6805, “The Application of the Path Computation Element Architecture to the Determination of a Sequence of Domains in MPLS and GMPLS”

[18] O. Gonzalez de Dios, R. Casellas, R. Morro, F. Paolucci, V. López, R. Martínez, R. Muñoz, R. Villalta, P. Castoldi: Multi-partner Demonstration of BGP-LS enabled multi-domain EON, accepted for publication in Journal of Optical Communications and Networking, Dec. 2015, Vol. 7, Issue 12, pp. B153-B162.

[19] Netphony-abno, https://github.com/telefonicaid/netphony-abno/wiki

[20] Open Network Foundation, https://www.opennetworking.org/

[21] Wireless Transport SDN Proof of Concept White Paper, https://www.opennetworking.org/images/stories/downloads/sdn-resources/white-papers/ONF_Microwave_SDN_PoC_White_Paper%20v1.0.pdf

[22] A. Aguado, et al. “Dynamic Virtual Network Reconfiguration Over SDN Orchestrated Multitechnology Optical Transport Domains”, in Journal of Lightwave Technology, April 2015.

}