End-to-end service provisioning across MPLS and IP/WDM domains

F. Muñoz[1], R. Muñoz[1], J. Rodríguez[1], V. Lopez[1], O. González de Dios[1] and J.P. Fernández-Palacios[1]

[1]Telefónica I+D

c/ Don Ramón de la Cruz 82-84, Madrid, Spain

***Abstract*— Metro and core networks have been traditionally designed and managed by different departments within network operators. Moreover, even in the cases where both metro and core networks are based on the same technology, e.g. MPLS, services need to be provided separately per domain. This historic separation is now being questioned in light of the potential cost efficiency and lower time-to-market that an end-to-end service provisioning could bring, especially in a cloud ready context, where the traffic demands are less predictable and fast provisioning is needed.**

This paper presents the drivers, enablers and challenges for an end-to-end service provisioning, in which a multi-layer IP/MPLS over WDM core and MPLS based metro networks are coordinated to support dynamic service provisioning. The feasibility of the end-to-end service provisioning is demonstrated in an experimental activity using commercial IP/MPLS routers, optical nodes and a network controller triggering control plane mechanisms.

Keywords: E2E Network Management; Seamless MPLS; Multi-layer traffic engineering; optical networks.

I. INTRODUCTION

EXISTING transport networks have been designed to support a relatively small number of connections with fairly static demands. However, the explosion of broadband connections imposes an unprecedented traffic growth in telecommunication networks with very high cumulative annual growth rates. An example of this huge traffic growth is the forecast from Cisco [1] that predicts an annual IP traffic growth of 29% from 2011 to 2016.

So far, the main connectivity demands have been driven by residential Internet access and business services. In this context, end-to-end service provisioning can take up to several weeks, which is acceptable for most end-users. However, data center services are becoming an essential component in the traffic sources for network operators [2]. The bandwidth requirements for cloud services are much more variable [3] than traditional services, and their network usage highly depends on the kind of service installed by the user in the cloud. Measurements used in [3] are for intra-data center traffic, but businesses are migrating from a private cloud paradigm to a hybrid cloud [4]. In hybrid cloud scenarios, the internal infrastructure of a company must coordinate with external resources (public or private). Therefore, hybrid cloud leads to a single cloud where some resources are in the company's network and others are in the Internet. For daily tasks, users might utilize the IT resources in the company, but external resources could also be used on-demand. This foreseen scenario leads to an increment in the service provisioning demands due to the inter-data center variable traffic. Thus, network operators will need to manage this context of dynamic end-to-end service provisioning efficiently.

During the last decade, operators have started introducing MPLS in new areas within their networks (for example, to implement transport services in their metros). The resulting MPLS domains are typically managed independently and have little interactions with the rest of the network segments. In such context, as MPLS is a common underlying technology, there is an effort in the industry to find solutions towards a single, yet scalable, MPLS domain, able to offer end-to-end services. This unique signaling domain in the whole network would be the main outcome of MPLS E2E architectures.

In large operators' networks typically more than one IGP area exists (e.g. metros and core). Currently, services, which traverse these areas, require of specific manual configurations in the areas' borders, which makes the service provisioning more complex, time consuming and expensive. Also, with time, the allocation of network resources would not be optimized, and there would be a higher possibility of failure. In this context, it would be desirable that an automatic mechanism to provision services through different areas existed. Moreover, if certain network parameters, such as bandwidth, delay, etc. could also be provisioned, that would facilitate the provisioning of certain services.

While MPLS E2E solution solves the multi-area problem, core networks are typically based on an IP/MPLS network over reconfigurable wavelength switched optical network (WSON). Again, both layers are operated isolated, and the configuration of the IP/MPLS routers is done by a different department than the optical equipment, which leads to inefficiencies. To optimize them there is an increasing interest in the coordination in this multi-layer network [5]. The interface between routers and optical equipment was traditionally UNI. However, as new parameters are required to increase the information between both layers, an extended UNI similar to an E-NNI interface is under discussion [6].

Two additional key elements are proposed to help as well in the coordination of such multi-layer architectures: the Path Computation Element (PCE) and the Virtual Network Topology Manager (VNTM) [7]. The aim of the PCE is to calculate the route between two endpoints, especially in complex scenarios (e.g. WSON with physical impairments, multilayer or multi-domain). On the other hand, the VNTM is in charge of maintaining the topology of the upper layer by connections management in the lower layer. The final entity required for this process is a network controller, which orchestrates the network creation process based on network measurements or failure detection [8].

The remaining is organized as follows: Section II summarizes current approaches towards a higher interaction on the one hand among MPLS segments, with focus on Seamless MPLS, and on the other hand between IP/MPLS and WDM networks. Next, Section III presents how both approaches can be combined to support end-to-end services in metro and core networks. Section IV presents the details of an experimental demonstration of end-to-end service provisioning, showing the feasibility of the multi-domain multi-layer end-to-end approach. To the best of the authors' knowledge this is the first experiment coordination in metro and multi-layer core networks. Finally, Section V concludes the paper.

II. Towards Network Automation To Support Cloud-Services

Current network architectures implemented by most operators do not enable dynamic provisioning of connections in multi-area scenarios, e.g. interconnecting metro and core networks. Moreover, these networks seldom present multi-layer capabilities. Latest industry efforts towards these goals are described hereafter, as a first requirement for elastic on-demand bandwidth provisioning for cloud services [2].

The cloud computing paradigm provides a new model for service delivery where the Information Technology (IT) resources form a pool able to attend multiple service demands by means of a dynamic assignment of resources, like CPU or storage capacity, either physical or virtual (by using some abstraction mechanisms). The variability of the cloud computing resides on the capability of virtualizing host instances where some services or processes are deployed. The virtualization technology allows a flexible management of IT resources, distributing them as needed either among distinct servers into a data center, or even spreading them across several data centers connected to the network. That distribution can even cross administrative boundaries, and involve public and private data center.

A highly scalable and flexible infrastructure providing on-demand capabilities easily adapts to the business requirements of cloud users, as well as allows the efficient utilization of the cloud provider resources. In this multi-tenant model, the sharing of resources among users reduces costs and maximizes utilization, leveraging the economy of scale. For instance, enterprise oriented cloud usage scenarios already demand combined computing and network resource provisioning. This implies requirements to address issues such as low latency, guaranteed bandwidth, application-centric management, security service consistency and energy efficiency. The combined cloud and network resource provisioning requires that a number of services and control systems interoperate at different stages of the whole provisioning process.

As stated in [2], three fundamental pillars are required in the migration path for cloud-aware transport networks:

1. **Flexible data plane technologies.** It is required to introduce new technologies providing transport flexibility reconfiguration and adaptability to change the assigned capacity to the actual demand.
2. **New control mechanisms**. Network control mechanisms have to be deployed allowing intelligent and automatic network processes.
3. **Coordination**. There are two main topics regarding coordination. On the one hand, it is needed to coordinate the context of both the application and the network, jointly considering the needs of both strata in a coherent manner. On the other hand, a network controller is required to orchestrate such process.

For this work, flexible data plane is possible thanks to the utilization of Seamless MPLS. Secondly, network control is carried out using a multi-layer control plane. Both technologies are explained in the next section. Regarding coordination, this work proposes a neutral term we have called network controller. According to the state of the art, there are three approaches which can play such role of coordination, namely: (1) Active Stateful PCE [9], (2) Software Defined Networking (SDN) controller [10], and (3) ABNO architecture [11].

Active Stateful PCE is a path computation entity, which can maintain the sessions for the LSPs or even create LSPs in the network. To cope with such coordination role, the main problem Active PCE has is that its interface in PCEP. To use Active PCE from a cloud service, cloud system should know which are the routing information of the network elements and PCEP is not a protocol included usually in cloud frameworks like OpenStack [12] or OpenNebula [13].

SDN strictly decouples control and data plane in the network equipment. Almost bare-bones hardware network elements switch data packets, while there is a SDN controller, which configures network elements based on the network information it has. This SDN controller has to fulfill three functions: (1) discovery, (2) provisioning and (3) monitoring. These functionalities are not defined yet for a SDN controller. OpenFlow protocol [14] is gaining momentum as the interface for provisioning, but there is not a clear candidate for the other functions. However, OpenFlow does not support all configuration capabilities so other interfaces should be included like NetConf or CLI.

Finally, ABNO architecture is a toolbox with multiple standard components, which can cope with orchestration use cases as it is defined in [11]. Although ABNO defines multiple components in the architecture, depending on the scenario or the use case, they can be used or not. The advantage of ABNO architecture is that interfaces and modules are defined based on standard protocols and entities and functionalities are clearly split, which facilitates its deployment in realistic scenarios.

III. Technologies in Metro and Core Networks

To allow horizontal and vertical orchestration for end-to-end services across MPLS and IP/WDM domains, it is required two technologies seamless MPLS and Multi-layer Management.

A. Seamless MPLS

Seamless MPLS [15] is not a new protocol or set of protocols, but a new network architecture, based on existing standards, for the resolution of scalable E2E MPLS networks.

An important feature in Seamless MPLS is that the network boundaries disappear in the service plane: services can be created provisioning only at the source and destination nodes. The resulting plane decoupling (service vs. transport) provides the following nodes classification: access (AN), service (SN) and transport (TN). Paths are established among access and service nodes, while transport nodes have no knowledge of the service itself.

Seamless MPLS solves the scalability issues by the introduction of two concepts: (1) division in routing areas and (2) MPLS hierarchy. In the most typical case, there would be three label levels: (1) intra-area, (2) inter-area and (3) E2E service. LDP or RSVP can be used for intra-area signaling. Inter-area signaling is solved by Labeled-BGP, both in the edge nodes where ANs and SNs are connected, and at the Area Border Routers (ABR), which are the frontier nodes between different routing areas. Finally, E2E services are typically instantiated over pseudowires, establishing T-LDP sessions between ANs and SNs only.

Fig. 1 presents the previously defined E2E architecture and its service provisioning. As depicted, there exist different areas where LSPs must be created. These intra-area LSPs can be established by using either RSVP or LDP allowing, together with the inter-area L-BGP signaling, the E2E services establishment over pseudowires. Connection mechanisms proposed in Seamless MPLS for access and service nodes are out of the scope of this paper, but are also MPLS-based.

Thus, the described Seamless MPLS architecture allows for multi-area service creation with management operations only at the service ends, avoiding intermediate equipment configuration steps that would increase Network Management Systems (NMS) complexity and Operational Expenditures (OpEx).

B. *Multilayer Management*

On most occasions, multi-layer core networks are currently managed separately in each layer. This means a different NMS per layer, and therefore an increment in the cost and in the network management tasks that imply human interaction.

The control plane appears as a set of functions (routing, resource reservation and link management) which are done automatically and distributed in the network. This automatic

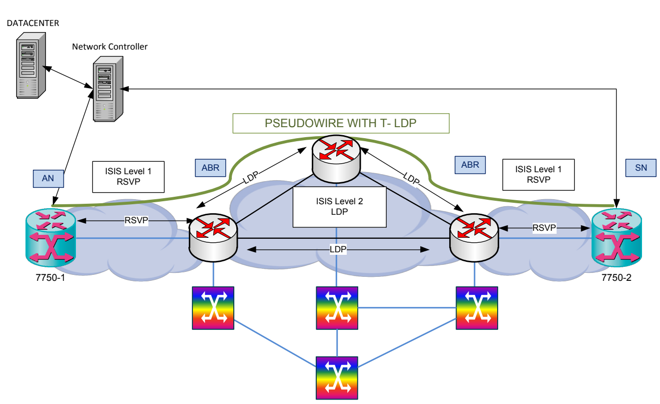

Fig. 1. Seamless MPLS E2E provisioning

process allows reducing the management complexity for a network operator: the operator must trigger the path establishment, but routing and resources reservation are done automatically by the control plane. Nowadays in multi-layer networks, control plane, as happens for management systems, also operates separately in each layer, with different elements and switching technologies. This section presents a multi-layer control plane architecture to solve such lack of interoperability between layers.

First, the key elements of the multi-layer control plane (TE-Links, Hierarchical LSPs and ML-PCE) are presented, together with two network entities (Network Management System and Virtual Network Topology Manager) which help to automate the process while there is not an integrated multi-layer control plane.

The ML-PCE (Multi-Layer PCE) is a standardized entity, evolution from the transport layer PCE, which also considers the IP/MPLS topology, allowing multi-layer path computation.

Hierarchical LSPs is a model definition to encapsulate different switching technologies to allow higher granularity in resource reservation across layers. This hierarchy of LSPs is achieved through the encapsulation of LSPs and helps to give a logic vision of all the connections in the different layers of the network. One of the main goals of hierarchical LSPs is to increase network scalability.

Finally, TE-Links concept is applied to two types of network adjacencies: the connections between different layers equipment (known as inter-layer links) and the logical connections between two nodes on the same layer, but which need to cross other layers to establish the adjacency. This concept allows operating in a simple way keeping in mind at any moment the information relevant for each procedure to be performed in the network.

There are different models for the control plane to be used in multi-layer environments. "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 (User to Network Interface), 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. 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.

To do this in an automated fashion, VNTM provides control plane information and functions to decrease the network controller tasks to be done in terms of multi-layer topology computation and configuration. The VNTM's main functions are the following:

- Gather and update inter-layer connectivity information between IP/MPLS and transport equipment.
- Compute and create a virtual topology table with current and potential forwarding adjacencies between IP/MPLS nodes. This table will include specific information about links such as metrics and Shared Risk Link Group (SRLG) traversed in the virtual topology.
- Automatically trigger multi-layer restoration mechanisms when critical unrecoverable failures take place in the network.

IV. Seamless MPLS and Multi-Layer Architecture Compatibility

Seamless MPLS consists of an architecture where MPLS can be used to provide connectivity between each node of the MPLS network. It is especially useful for network scenarios where it is needed to create services between two metropolitan areas crossing a core network with a different MPLS protocol stack. On the other hand, core network may be based in a multi-layer approach with some degree of coordination.

For some services, like Cloud-based applications, compatibility between Seamless MPLS and multi-layer core network may imply that the core network must be able to provide network connectivity with some degree of QoS for the services demanded by Seamless MPLS pseudowires (PW).

To achieve the required compatibility degree, core network operation must be automated to provide bandwidth for PW requests. To do so, core network must have an extended UNI, automated IP equipment configuration, and VNTM and ML-PCE entities.

A use case of Seamless MPLS and multi-layer architecture coordination is a data center request to connect to a public cloud in a different MAN. The idea is that the network operator just configures both ends (terminal equipment in both MANs) and the full procedure of the service establishment is done automatically. If all core network links are already established and the free bandwidth is enough, the multi-layer architecture is not involved in the request handling. However, if we take into count other parameters as for example, end to end delay, jitter or other QoS parameters, it should be checked if the PW copes with them.

If the existent established paths cannot cope with the parameters requested, then the multi-layer core network has to create new paths to provide the desired QoS for the requested PW. This is the use case demonstrated in Section IV. For this use case, we assume there is not enough bandwidth in the LSPs already established in the core network. As result of a new service request an E2E PW is tried to establish, but the network controller detects there is not enough bandwidth for the service. As a consequence, the automated multi-layer core network reconfigures itself creating a new path increasing the bandwidth to allow the new service reservation.

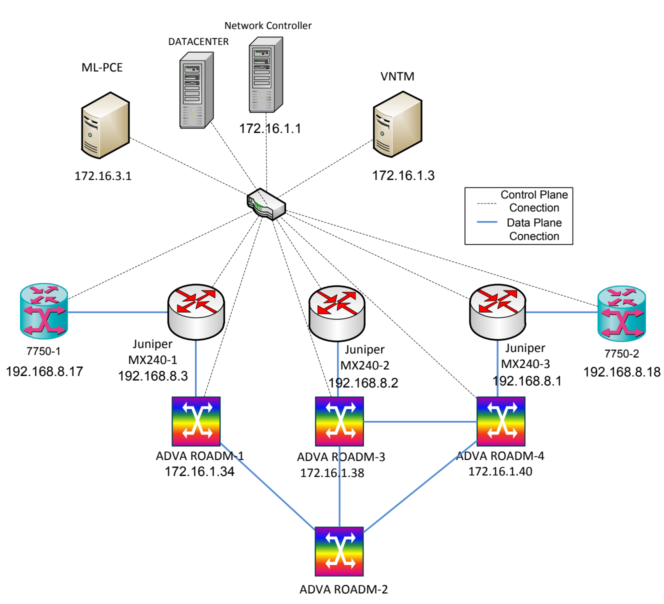

Fig. 2. Seamless MPLS E2E provisioning

V. Experimental Demonstration of E2E Service Provisioning

A. *Experimental Set-Up*

The experimental set-up consists of three MPLS network domains, emulating two MANs connected by a core network. Each MAN also has an ALU 7750, which is used as end point for the E2E services (the pseudowires are established between both ALU routers). The multi-layer core network has four ADVA FSP3000 optical nodes managed with GMPLS control plane and three Juniper MX routers. Two of them are acting as Seamless ABR nodes separating core and metro areas.

RSVP and LDP are used in the set-up. The former is used in both MANs and the latter for the core network. Target LDP is used for the E2E PW signaling and GMPLS and standard UNI for resource reservation across core network. The addressing scheme is presented in Fig. 2.

In the test bed there have also been deployed in house developments of a network controller, a ML-PCE and a VNTM for the tasks defined in Section II-B. The network controller configures ALU 7750 nodes for PW establishment via CLI and of the Juniper MX routers to establish the new IP/MPLS adjacency across the transport layer. All entities are connected in a LAN to have control plane reachability.

B. *Use Case Scenario*

A bandwidth on-demand use case is demonstrated in this experimental set-up. To show the coordination between the metro and core networks, the initial scenario is configured without enough bandwidth in the core network for the requested service. The associated service provisioning workflow is depicted in Fig. 3.

Let us assume that a data center requests a new connection to the network controller (1). First, network controller checks if the required resources (BW in this case) are or not available. To do so, a PCEP request is sent to the ML-PCE (2). As we have defined the network without enough bandwidth for the new request, the ML-PCE responses with NO PATH based on TED information (3). The network controller launches a request to VNTM (4) for creation of new resources between the end points in the core network. VNTM then asks ML-PCE which is the best alternative for a new link across the core network (5-6), and triggers its creation within the network (7), which is done by means of the standard UNI. After acknowledgment from the network nodes (8), VNTM

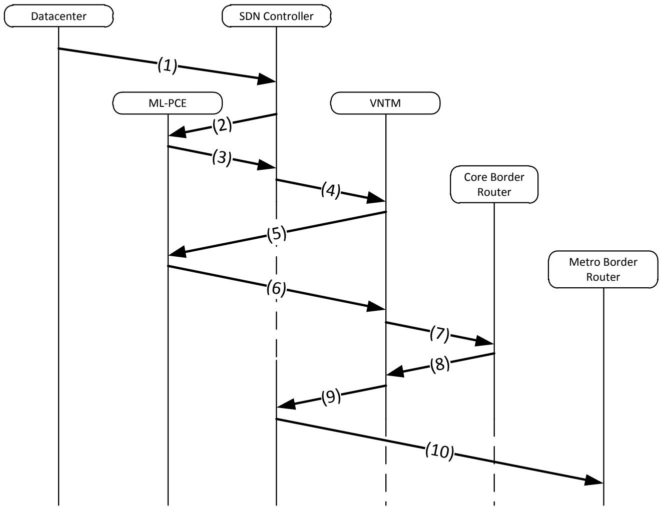

Fig. 3. Full E2E establishment procedure

responses to the network controller, informing that the new path is available (9). At this moment, the network controller instantiates the new service only in the edge nodes (10).

Following some Wireshark captures are shown with some steps in the process. Initial steps in the process can be seen in detail in previous work [8]. Fig. 4 presents the PCEP response to a VNTM request for a new multi-layer path (step 6) from the ML-PCE communication for network creation. UNI requests messages from the IP routers to the optical nodes are shown in [8]. The delay to set-up the new path in the core network was around 1 minute (mainly due to the optical devices).

Once the core network is configured, the E2E PW establishment process is done by both 7750 using T-LDP (Fig. 5). Label distribution procedure starts at PW level as soon as the core connectivity is established.

VI. Summary and Conclusions

This paper proposes an E2E architecture for multi-area and multi-layer scenarios and validates its feasibility in an experimental test bed with commercial equipment. To the best of the authors' knowledge this is the first proof of concept for the coordination in metro and multi-layer core networks.

The architecture is based on current standard solutions, although there are three in house developments: ML network controller, ML-PCE and VNTM. The definition of standard interfaces for the configuration of IP/MPLS and optical devices is pushed in FP7 ONE project [5]. When VNTM and the interface to configure network elements become standard

```
⊟ Path Computation Element communication Protocol
  ⊞ PATH COMPUTATION REPLY MESSAGE Header
  ⊞ RP object
  ⊟ EXPLICIT ROUTE object (ERO)
      Object Class: EXPLICIT ROUTE OBJECT (ERO) (7)
      Object Type: 1
    ⊞ Flags
      Object Length: 96
    ⊞ SUBOBJECT: IPv4 Prefix: 192.168.8.3/32
    ⊞ SUBOBJECT: SERVER LAYER INFO: Switching cap Lambda Switch Capable (SC=150)
    ⊞ SUBOBJECT: Unnumbered Interface ID: 172.16.1.34:0
    ⊞ SUBOBJECT: Label Control
    ⊞ SUBOBJECT: Unnumbered Interface ID: 172.16.1.36:0
    ⊞ SUBOBJECT: Label Control
    ⊞ SUBOBJECT: IPv4 Prefix: 172.16.1.40/32
    ⊞ SUBOBJECT: SERVER LAYER INFO: Switching cap Lambda Switch Capable (SC=150)
    ⊟ SUBOBJECT: IPv4 Prefix: 192.168.8.1/32
        L=0 Strict Hop
        Type: SUBOBJECT IPv4 (1)
        Length: 8
        IPv4 Address: 192.168.8.1
        Prefix Length: 32
        Padding: 0x00
```

Fig. 4. ML-PCE PCEP Response to VNTM Request

23.2€192.168.8.17	192.168.8.18	LDP	92 Hello Message
23.2€192.168.8.18	192.168.8.17	LDP	96 Hello Message
23.2€192.168.8.18	192.168.8.17	LDP	114 Initialization Message
23.2€192.168.8.17	192.168.8.18	LDP	120 Initialization Message
23.2€192.168.8.18	192.168.8.17	LDP	96 Keep Alive Message
23.2€192.168.8.17	192.168.8.18	LDP	92 Keep Alive Message
23.2€192.168.8.18	192.168.8.17	LDP	110 Address Message
23.2€192.168.8.17	192.168.8.18	LDP	164 Label Mapping Message
23.2€192.168.8.18	192.168.8.17	LDP	136 Label Mapping Message

Fig. 5. T-LDP session establishment

(Openflow or NetConf are proposals for this), the rest of the operation would be automated, supported by standard control plane functionalities. In the current provisioning and management systems, each operation would probably demand the human operator's acceptance, after a detailed testing procedure, which might also need to be acknowledged by human intervention.

Acknowledgments

This work was partially supported by the ONE project in the FP7 Program, contract number INFSO-ICT-258300.

References

[1] Cisco, "Visual Networking Index (VNI)," Cisco, Tech. Rep., 2012.

[2] 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.

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

[4] Gartner Report, "Five Cloud Computing Trends That Will Affect Your Cloud Strategy Through 2015.", http://www.gartner.com/resId=1920517

[5] M. Yannuzzi, A. Jukan, X. Masip-Bruin, M. Chamania, R. Serral-Gracia, V. López, O. Gonzalez de Dios, A. Azañon, M. Maciejewski, C. Brunn, M. Roth and J. Altmann, *The Internet and Transport Network Management Ecosystems: A Roadmap Toward Convergence*, in Optical Networking Design and Modeling (ONDM), Apr 2012.

[6] V. Beeram, et al, "Generalized Multiprotocol Label Switching (GMPLS) External Network Network Interface (E-NNI): Virtual Link Enhancements for the Overlay Model", IETF draft, Online Sept 2012.

[7] E. Oki, et al, *"Framework for PCE- based inter-layer MPLS and GMPLS traffic engineering,"* IETF RFC 5623, September 2009.

[8] O. Gonzalez de Dios, M. Cuaresma, S. Martinez, F. Munoz, V. López, J.P. Fernandez-Palacios: *Functional validation of the cooperation between Virtual Network Topology Manager (VNTM) and Path Computation Element (PCE)*, in International Conference on IP + Optical Network (iPOP), Jun 2012.

[9] E. Crabbe et al. "PCEP Extensions for Stateful PCE", draft-ietf-pce-stateful-pce-03, March 2013.

[10] Ping Pan, "Efficient Inter-Data Center Transport within SDN Framework", iPOP 2012, Tokyo, May, 2012.

[11] D. King and A. Farrel, "A PCE-based Architecture for Application-based Network Operations", draft-farrkingel-pce-abno-architecture-03, Feb 2013.

[12] OpenStack Networking Quantum, http://wiki.openstack.org/Quantum

[13] OpenNebula http://opennebula.org/

[14] Open Networking Foundation. Openflow Switch Specification. Version 1.1.0. http://www.openflow.org/documents/openflow-spec- v1.1.0.pdf

[15] N. Leymann et all, "Seamless MPLS Architecture", IEFT draft, Online Sept. 2012.