Experimental Validation of the ACTN architecture for flexi-grid optical networks using Active Stateful Hierarchical PCEs

Ramon Casellas, Ricard Vilalta, Ricardo Martínez, Raül Muñoz, Haomian Zheng +, Young Lee *

Centre Tecnològic de Telecomunicacions de Catalunya (CTTC/CERCA), Castelldefels, Spain
*+ Huawei Technologies Co., Ltd. Shenzhen, P.R.China * Huawei Technologies, Plano, TX, USA*
e-mail: ramon.casellas@cttc.es

ABSTRACT

The Abstraction and Control of Traffic Engineered Networks (ACTN) defines the requirements, use cases, and an SDN-based architecture, relying on the concepts of network and service abstraction, detaching the network and service control from the underlying data plane. The architecture encompasses Physical Network Controllers (PNCs), responsible for specific technology and administrative domains, orchestrated by Multi-Domain Service Coordinator (MDSC) which, in turn, enables underlay transport resources to be abstracted and virtual network instances to be allocated to customers and applications, under the control of a Customer Network Controller (CNC). In this paper, we present the application of the ACTN architecture to the control of a multi-domain flexi-grid optical network, by proposing, adopting and extending i) the Hierarchical active stateful Path Computation Element (PCE) architectures and protocols ii) the PCEP protocol to support efficient and incremental link state topological reporting, known as PCEP-LS (PCEP link state), iii) the per link partitioning of the optical spectrum based on variable-sized allocated frequency slots enabling network sharing and virtualization, and iv) the use of a model-based interface to dynamically request the instantiation of virtual networks for specific clients / tenants. We report the design and the implementation of the testbed in order to validate the approach.

1. INTRODUCTION AND MOTIVATION

As transport networks evolve, the need to provide network abstraction and virtualization in a multi-domain setting has emerged as a key requirement for operators. Network virtualization refers to allowing customers and applications (tenants) to utilize and independently control allocated virtual network resources as if resources were real, thus supporting multitenancy. The network is "sliced", with tenants being given a different partial and abstracted topology view of the physical underlying network.

2. ABSTRACTION AND CONTROL OF TRAFFIC ENGINEERED NETWOKS ARCHITECTURE

The ACTN [1][2] Software Defined Network (SDN) architecture (Fig. 1), initially intended for optical transport and extended to other TE technologies, specifies the functional entities and methods for the coordination of resources across multiple domains, to provide end-to-end services. It comprises of Physical Network Controllers (PNCs), responsible for specific technology and administrative domains, connected to a Multi-Domain Service Coordinator (MDSC). The MDSC facilitates underlay transport resources to be abstracted and virtual network instances to be allocated, while a corresponding Customer Network Controller (CNC) manages virtual network instances. ACTN facilitates heterogeneous domain transport networking and control/management, while enabling a logically centralized multidomain orchestration, using a hierarchical architecture to scale.
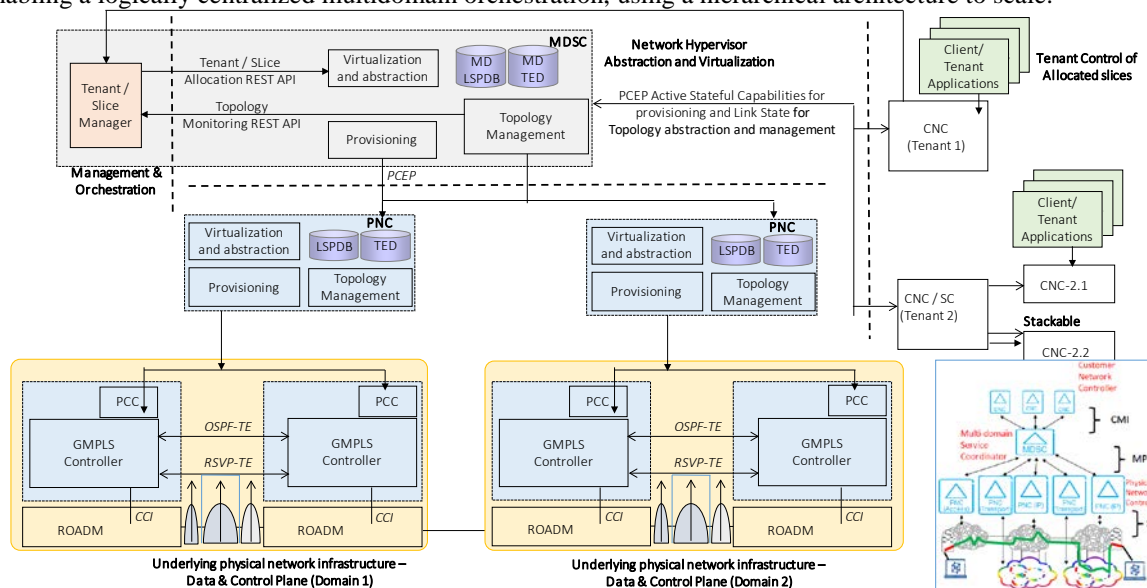

Figure 1. Implemented ACTN with AS-PCE as CNC, MDSC and PNC over a GMPLS flexi-grid network.

3. PCE/PCEP BASED ACTN IMPLEMENTATION

While the ACTN architecture remains generic, it can be mapped to existing entities, interfaces and protocols in view of leveraging on and supporting existing deployments and mature solutions. In particular, the Path Computation Element (PCE) [3] architecture can be extended to support ACTN requirements, and the PCEP protocol can be used to support ACTN interfaces and, in particular, the MDSC-PNC Interface (MPI) between an MDSC and a PNC for the requests for new connectivity or for bandwidth changes adapting the Active Stateful Hierarchical PCE framework [4] (in addition to the provisioning interface for creating forwarding state in the physical network, requested via the PNC).

Child domains: an Active Stateful Path Computation Element (AS-PCE) is an entity that is able to, on the one hand, perform constrained path computation on a network topology (Traffic Engineering Database, TED) and, on the other hand, provision and reroute connections either directly [5] or driving a GMPLS control plane [6].

Topology Management: the PCE architecture does not specify how a PCE obtains its Traffic Engineering Database (TED). That said, it is assumed that each PNC has a detailed view of the topology (e.g. by participating or sniffing IGP such as OSPF-TE or retrieving topological information directly from the PCC at the nodes). Similarly, the architecture does not mandate what topological information the MDSC has beyond basic domain-connectivity and endpoint reachability. In our architecture, each PNC/cPCE is responsible for abstracting and aggregating domain topological information and disseminate it to the MDSC using PCEP for Link State [7].

Service Path Computation and Instantiation: the end-to-end path computation is based on a sequence of interactions between the MDSCs and the PNCs using abstracted multi-domain TED involving optimal domain selection and segment expansion. After the path computation is complete, the MDSC coordinates the provisioning across the domains by requesting the segment provisioning using PCInitiate messages.

Virtual Network (Network Slicing) and Client support: the MDSC is extended to support the allocation of slices (partitions of the optical infrastructure), by using a dedicated REST interface. It exports a tenant REST API that allows the client (through the CNC and via the NMS/tenant manager) to request the allocation of a slice/tenant. The request includes the sub-TED (defined as a partitioning or sub-graph with the nodes of the sub-graph and link objects that include a bitmap mask that defines the usable optical spectrum ranges) as well as the address of the corresponding CNC. After the initial synchronization, PCEP-LS updates for the slice links are forwarded (see Fig. 2). Subsequent topology updates are only sent for the network elements and links within the CNC allocated slice. A CNC can thus provision LSPs choosing either PCEP or a REST API (the latter expected to be used by end applications).
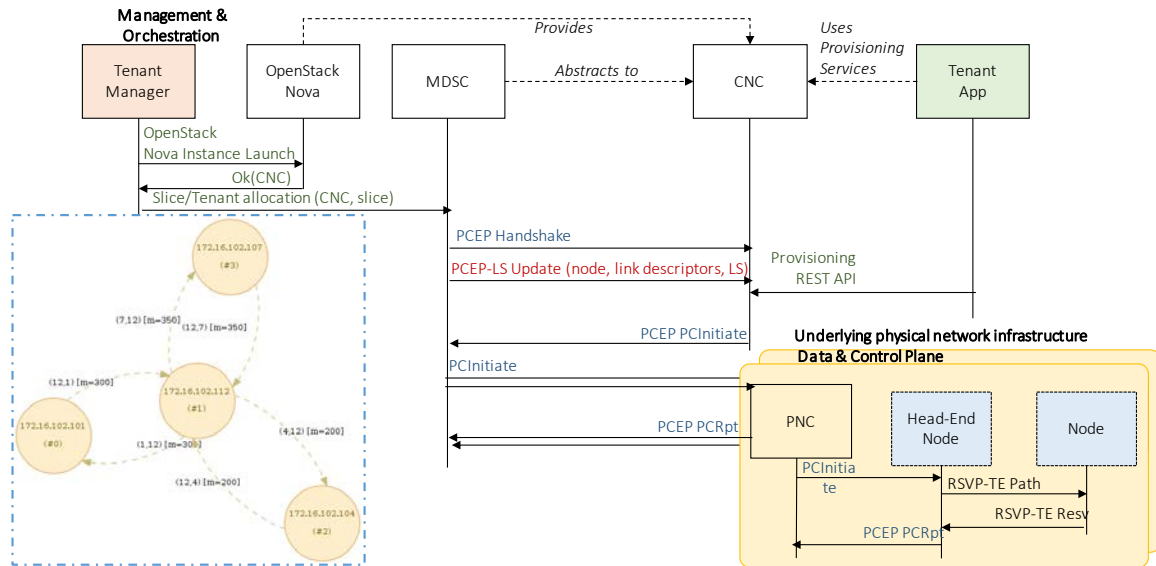

Figure 2. Message flow for the dynamic allocation of CNCs/ flexi-grid slice creation and per client LSP setup.

4. EXPERIMENTAL EVALUATION

The implemented architecture has been evaluated in a control plane testbed, using 4 child domains (10.0.50.0/24, 11.0.50.0/24, etc.) each composed of 14 nodes, with 44 links and 10 Nominal Central Frequencies (NCF) per link, each with its own GMPLS control plane, with LSP control delegated to a single AS-PCE child of a unique parent AS-PCE/PNC. At each domain, the AS-PCE periodically computes (with a configurable timer set to 60s) a full-mesh virtual links between the domain border nodes using a maximum residual bandwidth (shortest widest) path computation ensuring spectrum continuity. Each virtual link is subsequently announced to parent MDSC using PCEP-LS, notably including a bitmap encoded representation of nominal frequencies that represent available frequency slot ranges. Thus, the parent topology encompasses inter-domain links along with the

announced virtual links, resulting in an abstracted and aggregated multidomain network. Figure 3 (center) shows the aggregated topology seen at the MDSC.
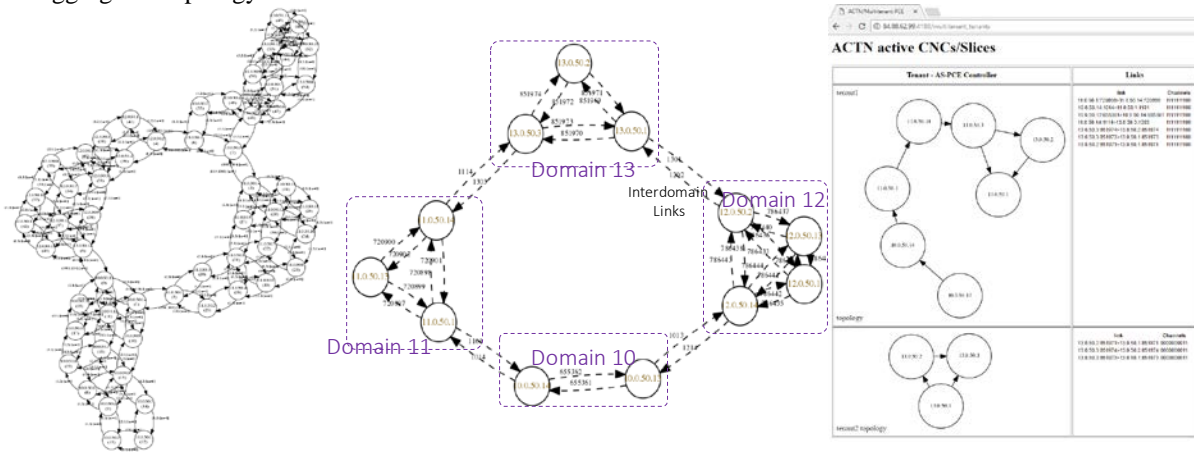

Figure 3. Topological testbed setup: 4-domain full topology when PNCs disclose all topologies (left). MDSC aggregated topology for the virtual link aggregation policy (center) and MDSC management of two active CNC/Slices showing each slice topology (right).

As explained, a client may request a slice of the end-to-end abstracted flexi-grid network. For this, currently a simple information model is used to request the slice, although work is ongoing to use ACTN specific model [8]. Additionally, CNCs are allocated on demand using an ETSI/NFV Network Service with a single VNF and, once instantiated, an initial synchronization is done using PCEP-LS. Figure 3 (right) shows two active slices that have been requested, as managed by the MDSC. Figure 4 shows the details of the REST interface for the creation of a tenant and the link list seen by the CNC.
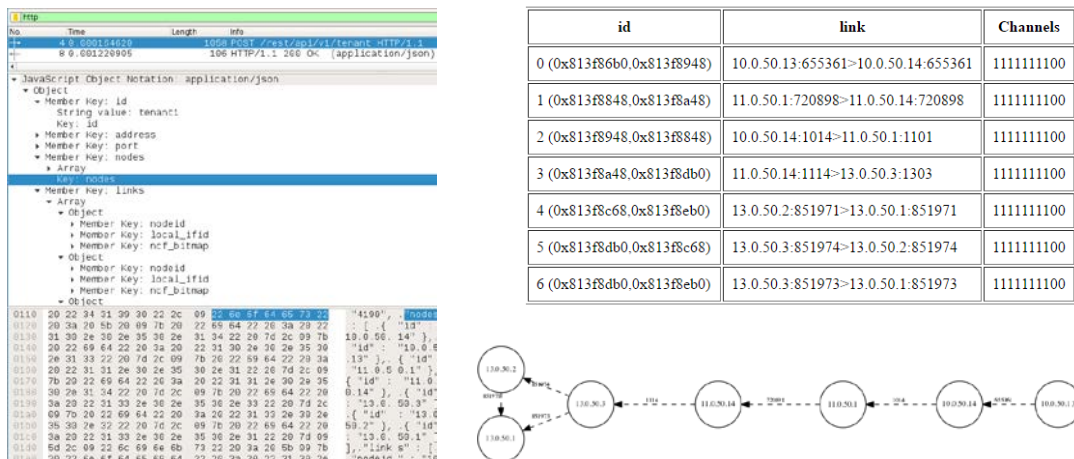

Figure 4. Wireshark capture of the REST interface to request a partition of the topology seen by the MDSC.

4.1 Experimental results

In our proof-of-concept, we have evaluated aspects such as PNC topology aggregation and dissemination using PCEP-LS and service provisioning times from different perspectives. A transit PNC/cPCE is able to compute a mesh of virtual links in a time roughly proportional to $O(N^2)$, where N is the number of border nodes (in particular, a measurement yielded ~2ms for N=2 and ~100ms for N=12), for the considered topology and virtual link aggregation algorithm, forwarding the virtual links to the MDSC. The PNC is able to synchronize its abstract topology to the pPCE/MDSC in O(15) ms, using a PCEP persistent connection over a 100Mb LAN (not including TCP/PCEP handshake). When considering the creation of slices, service instantiation time is mainly determined by the deployment of virtualized CNC, and largely depends on the capabilities of hosting nodes and the VM image size. In particular, with Linux images with AS-PCE software ~ 3GB, CNCs are typically operative in 10-60s. The synchronization of slice topology between MDSC and the CNC is similar to the one between PNC and MDSC, i.e., mainly determined by the number of topological elements and the PCEP/TCP throughput between PCEP peers.
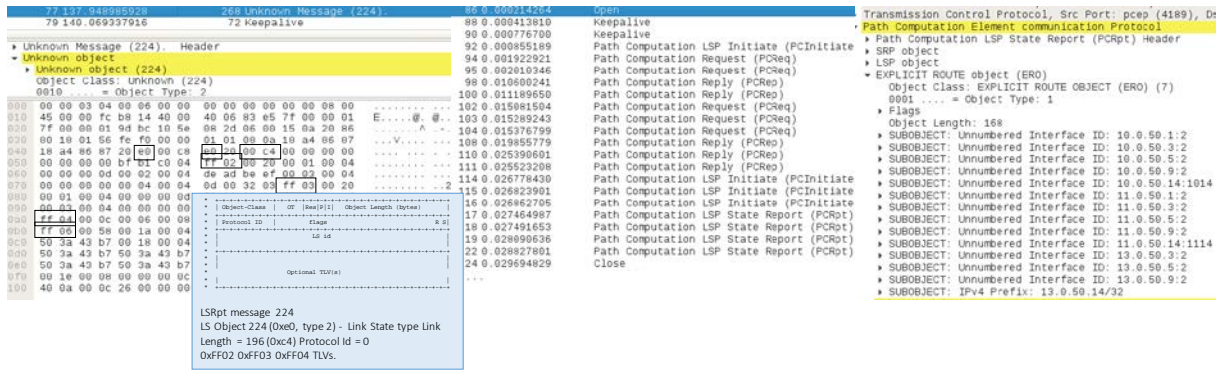
Figure 5. Wireshark captures within the testbed: PCEP-LS message with a Link State advertisement, including PCEP-LS TLVs (left). PCEP messages at the MDSC for the path computation and subsequent segment establishment (center) and computed ERO after segment expansion (right)

Finally, the instantiation of an LSP, measured at the CNC and using recursive PCEP PCInitiate/PCRpt messages (CNC→MDSC→PNC→PCC) is 0(100) ms (e.g. 30 ms without taking into account GMPLS setup delay, Fig 5, center). Let us note that, since the end-to-end flexi-grid connection is segmented, each PNC drives the establishment of LSP segment within its domain (spectrum contiguity being ensured by the MDSC at the frequency slot level) in parallel, setup time does not linearly depend on the number of domains, but on the slowest domain, along with the overhead of intermediate entities. For the considered domains, a segment LSP is stablished by the GMPLS control plane (seen by the ingress node) in O(10-20) ms. The actual LSP (once the domain segments have been expanded) is shown in Fig.5, right.

5. CONCLUSIONS

The ACTN architecture aims at addressing advanced requirements related to the control of multi-domain Traffic Engineered networks while enabling use cases such as multi-tenancy and per client control of virtualized infrastructure. In this paper, we have detailed the implementation of the ACTN architecture in terms of hierarchical active stateful PCEs, using PCEP-LS extensions for (aggregated) topology management, and how per-client controllers are instantiated on-demand, to control allocated slices, defined in terms of partitioning of the flexi-grid underlying optical network, both in terms of links and nodes as well as frequency slot ranges.

ACKNOWLEDGEMENTS

REFERENCES

[1] Y. Lee, editor, "Requirements for Abstraction and Control of TE Networks", IETF draft, work-in-progress https://tools.ietf.org/html/draft-ietf-teas-actn-requirements-04.txt, January 3, 2017.

[2] D. Ceccarelli, Y. Lee, editors, "Framework for Abstraction and Control of Traffic Engineered Networks", IETF draft, work-in-progress draft-ietf-teas-actn-framework-04.txt, February 16, 2017.

[3] R. Casellas, R. Muñoz, R. Martínez, and R. Vilalta, "Applications and Status of Path Computation Elements [Invited]", *IEEE/OSA J. of Optical Communications and Networking. v. 5, n. 10*, October 2013.

[4] R. Casellas, et al. "A Control Plane Architecture for Multi-Domain Elastic Optical Networks: the view of the IDEALIST project", *IEEE Communications Magazine,* Vol. 54, No. 8, pp. 136-143, August 2016.

[5] A. Farrel, ed. "An Architecture for Use of PCE and PCEP in a Network with Central Control", IETF draft, work-in-progress, https://tools.ietf.org/html/draft-ietf-teas-pce-central-control-01, December 5, 2016.

[6] E. Crabbe, I. Minei, S. Sivabalan, R. Varga, "PCEP Extensions for PCE-initiated LSP Setup in a Stateful PCE Model", IETF draft, work-in-progress, https://tools.ietf.org/html/draft-ietf-pce-pce-initiated-lsp-09, March 7, 2017

[7] D. Dhody, Y. Lee, "PCEP Extension for Distribution of Link-State and TE Information", IETF draft, work-in-progress, draft-dhodylee-pce-pcep-ls-07, February 16, 2017.

[8] Y. Lee, editor "A Yang Data Model for ACTN VN Operation", IETF draft, work-in-progress, https://tools.ietf.org/html/draft-lee-teas-actn-vn-yang-03, March 9, 2017