Effective Elasticity for Data Centers Interconnection in Multi-Domain WAN: Information Modelling and Routing

Paola Iovanna[(1)], Fabio Ubaldi[(2)], Francesco Giurlanda[(3)], Sandro Noto[(3)], Alessandro Priola[(3)], Luis M. Contreras[(4)], Victor Lopez[(4)], Juan Pedro Fernandez-Palacios Gimenez[(4)]

[(1)] Ericsson, via Giuseppe Moruzzi 1, 56214, Pisa, Italy. paola.iovanna@ericsson.com
[(2)] Ericsson, via Giuseppe Moruzzi 1, 56214, Pisa, Italy.
[(3)] CoRiTeL, via Giuseppe Moruzzi 1, 56214, Pisa, Italy.
[(4)] Telefonica, Core Network And Transport I+D, Ronda de la Comunicación, 28050.

Abstract *The paper proposes a novel information model to represent elastic capability of the network, E2E and intra-domain routing algorithms. Simulation results on relevant network show the solution efficiently combines DC and traditional operator services, with no impact on the complexity.*

Introduction

Cloud computing is one key driver for the evolution of telecommunication networks that is sum up by the “as a service” concept promoted by ETSI NFV group[1]. This paradigm allows simplifying the organization and manipulation of resources and network functions at different levels to create smart and fast services.

Cloud computing will evolve following two main trends[2]. The first one is the “centralization” of some control-plane related network functions in centrally placed big data centers (DCs) to allow resources pooling and the reduction of operation costs by concentrating site visit and competence. The second trend, instead, is to decentralize the location of some delay-sensitive and data-intensive network functions (typically data-plane related) in small edge DCs placed near the access networks in order to reduce the transport cost and increase the performance. As a consequence, it is a key requirement to provide ubiquitous and elastic network connectivity among distributed DCs connected by means of a transport network that usually consists of heterogeneous transport domains made of nodes from different vendors, employing different technologies and managed in different ways including domains managed just by Network Management System (NMS), or by distributed control plane.

DCs interconnections have a behaviour that is completely different from the traditional telecom and Internet traffic[3]. Thus, the traditional transport network architecture is unsuitable since it is too static if compared to the elastic and virtualized DC resources connected to it[4]. A solution for avoiding that is enabled by Software Defined Networking (SDN) paradigm that thanks to the capability of separating control and data-plane, facilitate the abstraction of the network resources and simplify the multi-vendor and multi-technology interworking. Network abstraction is a well-known concept that is deeply investigated in standardization bodies[1,5], where different abstraction methods are proposed[6]. However, such solutions lack of a flexible and elastic management of the network resources that is an important requirement for DC services due to the bursty behaviour of the traffic (like Web Gaming, MapReduce applications). However, the elasticity and flexibility features represent a crucial aspect for the network operators that have already deployed network infrastructures for traditional services. So, the trend is to combine them with the new DC services in order to efficiently use the network resources and maximize the return of investment. Such concept is expressed as “multiservice” networks paradigm.

This paper defines a novel information model to be exposed as part of the network abstraction that describes the elastic capability of the underlying transport network. Moreover the paper defines a solution that provides an efficient E2E service routing algorithm and an efficient intra-domain resource management that allows optimizing the transport cost. The proposed algorithms make the network able to follow the dynamic behavior of DC traffic while combined it with traditional traffic. The proposed solution is tested on a realistic network scenario provided by Telefónica. It shows a more efficient transport resource usage with respect to non-elastic methods (up to 10% of more service being routed), while the abstraction is kept stable. In addition different intra-domain resource management methods are proposed that can be suitable for different domain control systems. The paper is organized as follows: first the solution is detailed, then the simulation environment and the results are presented and finally conclusions are given.

Proposed Solution

The proposed solution is set on the reference architecture shown in Fig. 1.

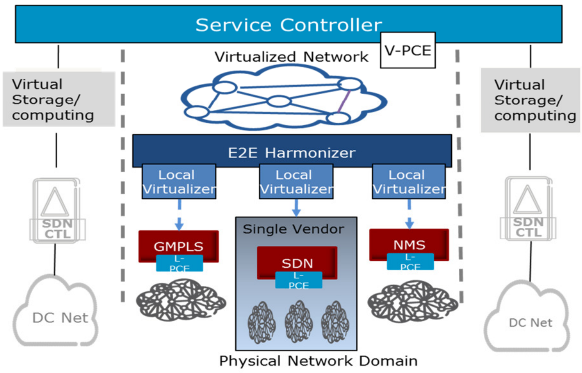

Fig. 1: Reference Architecture

Such architecture[7] is based on two main principles. The first one is to provide a sort of umbrella on the top of the transport layer that "translates" the technology parameters of each network domain into a set of agnostic parameters for provideing a homogenous abstraction view.The second principle is the separation of the E2E routing performed at service level on the virtualized view with respect to the intra-domain resource management.

Each DC domain exposes proper virtualization of computing and storage resource to the "Service Controller" (SC). According to the separation between service and transport layers, "E2E Harmonizer" is in charge to expose the virtualized view, obtaining by the "Local Virtualizer" (one for each domain) that collects domain connectivity (e.g. paths connecting border nodes). SC routes services between DC by using the virtualized view provided by "E2E harmonizer".

The proposed solution defines an information model to represent the elastic capability of the domain that is exposed in the abstract view. For each couple of border node of a given domain a virtual link (VL) is exposed in the abstract view. VL includes elastic parameters that represent the elastic capability that domain is able to provide for the border nodes connected by VL. Each domain path connecting two border nodes is described by three parameters. B_i is the free bandwidth already reserved by the domain control to *i-th* path. E_i is the additional bandwidth value that can be provided by links crossed by *i-th* path. $SB_j^{(k)}$ is the bandwidth required to guarantee the quality of *j-th* service with *k-th* priority class. The sum of B_i and E_i is the capacity of the path (C_i). The parameters that represent the elastic capability are: *vB* that is the virtual bandwidth reserved on VL, *vE* that is the virtual additional bandwidth available on VL, $vSB^{(k)}$ that is the virtual bandwidth that the domain can provide to the *k-th* service class. Such parameters are computed using the formulas shown in Eq. (1-3).

$$vB = \max(B_i) \quad (1)$$

$$vE = \max(B_i + E_i) - vB \quad (2)$$

$$vSB^{(k)} = \max(B_i + E_i - \sum SB_{j,i}^{(q)}) \quad (3)$$

$SB_{j,i}^{(q)}$ is the bandwidth of previous allocated *j-th* service on *i-th* path with *q-th* class of service where the priority is higher or equal priority to *k-th* service class.

The parameters are used by SC to compute E2E service routes. In the paper it is proposed the following algorithm: first a shortest path algorithm (Dijkstra is used for simulation) is executed on the abstract topology filtered by VL that does not satisfy the formula in Eq. (4).

$$SB_j^{(k)} \le vSB^{(k)} \quad and \quad B_r \le vB \quad (4)$$

B_r is the current bandwidth request by new service with *k-th* priority class. If a route is not found, the same shortest path algorithm is executed filtering the VL that does not satisfy the less strict Eq. (5).

$$SB_j^{(k)} \le vSB^{(k)} \quad and \quad B_r \le vB + vE \quad (5)$$

If a path is not found after this step the service request is refused. When the request is accepted, each "Local Virtualizer" involved in the computed E2E service route, selects a physical intra-domain path. Three different methods are proposed as intra-domain resource management. *"Minimum Path Selection" (MIN-PS)* selects for each service with *k-th* priority class the path with minimum available bandwidth able to satisfy the service bandwidth request $SB_j^{(k)}$. *"Maximum Path Selection " (MAX-PS)* selects for each service with *k-th* priority class the path with maximum available bandwidth able to satisfy service bandwidth request $SB_j^{(k)}$. *"Hybrid Path Selection" (HYB-PS)* combines the previous two methods. A *k-th* priority class is chosen and for each service with priority high or equal to it, the "MAX-PS" method is used; otherwise the "MIN-PS" is selected.

Simulation Environment

The reference network topology, provided by Telefónica, is the Spanish IP/WDM backbone, shown in Fig. 2. It consists of five regional domains based on IP/MPLS technology (Fig. 2a) and one national domain (Fig. 2b) providing optical interconnections between the regional domains, based on IP and optical technology.

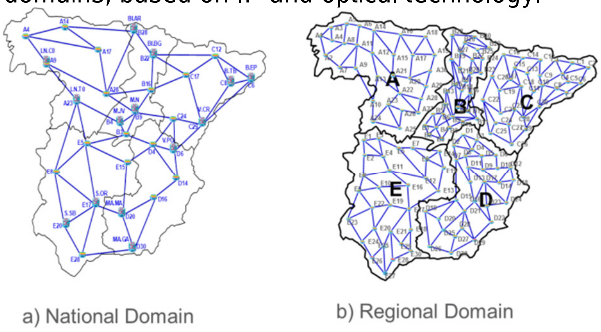

Fig. 2: Reference Topology

Three types of services are considered*: "Low-Latency" DC traffic (LL-DC)* is the highest priority service that represents user interactive DC services (like Web Gaming) where the requirement is to maintain a low delay. *"High-Throughput" DC traffic (HT-DC)* is the medium priority service that represents data intensive DC services (like MapReduce based one) where the network as to assure a high and constant throughput when it is required by the service. *"Internet" traffic (INT-TR)* is the lowest priority service (based on Best-Effort) that represents the traditional telecom traffic. Existing models are used to generate service traffic for both DC service[8] and for Internet traffic[9]. Service traffic is then splitted in transport request using ECMP algorithm[10].

Solution is compared with a non-elastic solution that allocates the peak for LL-DC traffic (80% of Ethernet link capacity), the mean-rate for HT-DC traffic and the current bandwidth request for INT-TR.

Results

Fig. 3 shows the comparison of the proposed elastic method, with the non-elastic solution in each domain. Proposed solution accepts more service requests (up to 10% of more accepted traffic request) preserving the abstraction complexity. In addition, Fig. 4 shows the path load distribution comparing the three methods of intra-domain path selection. The results show that "MIN-PS" splits the paths in two classes: the paths used to the current traffics loaded near the maximum capacity and the paths reserved for the new traffics, thus the "MIN-PS" method preserves the waste of the network resources.

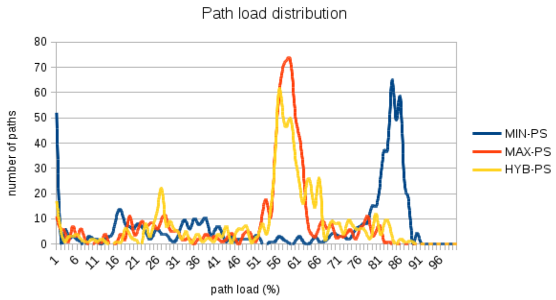

Fig. 4: Path selection comparison

Conclusions

The paper proposes a novel information model representing the elastic capability of the network. In addition, the paper proposes E2E service routing algorithms and intra- domain resource management methods that use such information model. The solution is able to efficiently follow the bursty dynamic of DC services and combines it with the traditional telecom services as shown by simulation results on Telefónica network.

References

[1] ETSI_PORTAL

[2] F. Bari et al, "Data Center Network Virtualization: A Survey", IEEE Communication Survey and Tutorials, Vol. **15**, no. 2, pp. 909-928 (2013).

[3] T. Benson et al., "Network Traffic Characteristics of Data Centers in the Wild", Proc. IMC 2010, pp.267-280

[4] L.M.Contreras et al., "Toward Cloud-Ready Transport Networks". IEEE Communications Magazine, pp. 48-55, September, 2012.

[5] ACTN WEB PAGE

[6] A. Fischer et al., "Virtual Network Embedding: A Survey", IEEE Communication Survey and Tutorials, Vol. **15**, no. 4, pp. 1888-1906 (2013).

[7] P.Iovanna et al., "Main Challenges on WAN due to NFV and SDN: Multi-layer and Multi-domain Network Virtualization Routing", Invited paper ONDM 2015.

[8] Y. Han et al., "Flow-Level Traffic Matrix Generation for Various Dat Center Networks", Proc. NOMS 2014, pp1-6

[9] Y. Zhang et al., "A Simplified Traffic Generating Method for Network Reliability Based on Self-Similar Model", Journal of Communications, Vol. **8**, no. 10, pp. 629-636 (2013)

[10]T. Benson et al., "MicroTE: Fine Grained Traffic Engineering for Data Centers", Proc. CoNext2011, Article 8

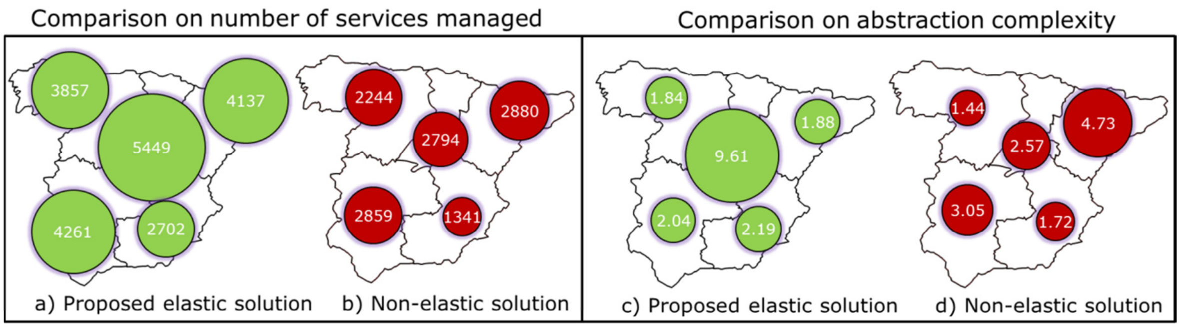

Fig.3: Elastic simulation results