End-to-End Service Provisioning in Carrier-Grade Ethernet Networks: The 100 GET-E3 Approach

Xiaomin Chen*, Admela Jukan* and Thomas Fischer**

*Technische Universität Carolo-Wilhelmina zu Braunschweig, {xiaomin, jukan}@ida.ing.tu-bs.de}
**Nokia Siemens Networks, Munich, Germany

Abstract—This paper presents and discusses issues of multi-domain, end-to-end service provisioning in Ethernet core networks from the perspective of the 100GET-E3 project, a German subproject under the umbrella of the CELTIC research initiative 100GET ("100 Gigabit Ethernet Technologies"). We propose several multi-domain service provisioning scenarios in IEEE’s Provider Backbone Bridging Traffic Engineering (PBB-TE) networks.

I. INTRODUCTION

Today, 95% of all Internet data traffic either originates or terminates as Ethernet [1]. Ethernet technologies have been commercialized and deployed widely due to their merits of simplicity, ubiquity, interoperability, as well as cost-effectiveness. Recently, efforts have started to extend the reach of Ethernet from LAN to WAN, on the one hand, to replace a good portion of SDH/SONET interfaces, while on the other hand to offload core IP routers and allow for high-bit rate traffic to be switched below OSI Layer 3. This has become an attractive proposition for many carriers, not only due to the Ethernet’s intrinsic cost-effectiveness, but also due to its technological maturity to run at data rates of up to 10Gb/s and even 100 Gb/s in future. While the Internet will continue to serve as pre-dominant source of content and applications, carrier-grade Ethernet is becoming one of the most promising technologies for future transport networks.

Standardization bodies and organizations have worked on the enhancement of original Ethernet to facilitate realization of carrier-grade Ethernet. Efforts are under way to develop higher speed Ethernet interface standards (IEEE 802.3ba) [2] and define Ethernet services with required Quality of Service (QoS), as well as service reliability and security [3, 4]. Moreover, OAM for Ethernet in the first mile was defined by IEEE (802.3ah)[5], while the so-called Connectivity Fault Management and Protection Switching were specified to check continuity and manage connectivity status for enhanced availability (802.1ag)[6]. Finally, Ethernet services (G.Imp8011.2 /Y1307.2) [7] and UNI as well as Ethernet transport NNI (G.8012/Y.1308)[8] were proposed by ITU-T to provide standard definition of carrier-grade Ethernet services and give the reference demarcation devices for inter-domains.

The research initiative 100GET ("100 Gigabit Ethernet Technologies, in Germany supported by the Federal Ministry of Education and Research) has responded to the challenge of the future Ethernet transport network, and has brought together industry and academic researchers to address a simple, packet-optimized and converged network from access to the core. As part thereof, the 100GET-E3 subproject spotlights the underlying 100 Gb/s DWDM transmission technology as well as all relevant methods, protocols and technologies for planning, management and control of multi-domain and multi-layer (Ethernet-over-DWDM) core networks. The cooperative research activities are carried out by academia, telecom industry and network operators. The activities in cluster E3 which address the multi-domain issues of service provisioning in Provider Backbone Bridging Traffic Engineering (PBB-TE) networks are the main focus of this paper.

This paper is organized as follows. In Section II, we briefly present the 100GET architecture and vision. Section III describes the generic network scenario for end-to-end service provisioning with multiple networks and discusses potential signaling mechanisms within carrier Ethernet networks. In Section IV, we discuss the complexity associated with multi-domain service provisioning in carrier Ethernet networks. In Section V, we conclude the paper and present the directions for further research.

II. 100GET ARCHITECTURE

One of the main objectives of 100GET is to develop a carrier-grade Ethernet/DWDM transport technology with cost/performance ratio superior to the traditional IP/SDH /DWDM. This is not only motivated by the ever-growing complexity of managing separate legacy networks with different elements and maintenance tools, but also by the complicated data framing and mapping which have been increasing the operation cost, while limiting adaptation to packet-oriented traffic. To reach these goals, the 100GET network architecture should offer flexible transport granularity, high adaptability to various access networks, as well as ability of providing end-to-end services with guaranteed performance and quality.

Our E3 architecture comprises of four network segments, namely Customer Network, Access Network, Aggregation Network and Core Network, as shown in Figure 1. Services carrying data originating from customer networks are multiplexed and differentiated at the access and aggregation networks. Legacy services such as voice, video and data from residential or business end-clients in the customer networks are first multiplexed by DSLAM connected to the edge Carrier Ethernet Switches (CES) of the Carrier Ethernet transport

1 In PBB-TE, this edge switch is called BEB (Backbone Edge Bridge).
layer in Access Network. The CES incorporates multi-layer processing, data encapsulation and packet forwarding. Ethernet-based services such as E-Line, E-LAN enter the carrier Ethernet via an access switch. ISPs (Internet Service Providers) peering utilize Broadband Remote Access Server (BRAS) in IP layer to connect with core IP/MPLS network, and then adapted to the core carrier Ethernet. Services provisioned through 3P-PE can be differentiated at the edge of IP/MPLS network before transported through carrier-grade Ethernet layer. In Figure 1, assumed but not shown here is an automatic control plane based provisioning that enables timely configuration of the network infrastructure.

Two functional layers are shown in this architecture: IP layer and carrier Ethernet transport layer. In this architecture, IP layer is deployed over the carrier Ethernet layer directly. Carrier Ethernet transport layer extends to access networks and other network domains with CES as edge equipments. IP/MPLS network in the IP layer is connected to CESs in transport layer via its edge elements LER (Label Edge Router), making it possible to bypass LSRs (Label Switch Routers) within carrier network. Customer Ethernet services from end-clients with multiple rates can be adapted dynamically by Ethernet interfaces and L2-switches, which can support 10M/100M Ethernet service backward compatibly in legacy Ethernet and 1G/10G/40G in Ethernet transport networks or even as high as 100G in future carrier Ethernet networks [8].

Traditional SDH/SONET services, with speeds as high as OC-192 can be used in combination with the configurable optical layer, while the IP services with multiple granularities can be forwarded from IP layer to the carrier Ethernet layer over label switch routers.

End-to-end service provisioning in the carrier Ethernet can be implemented with pre-established tunnels by using one of the three carrier-grade Ethernet forwarding mechanisms: PBB-TE, VLAN-Switching and T-MPLS. As a replacement of MAC learning-based forwarding in the traditional Ethernet environment, all three carrier grade Ethernet technologies introduce new mechanism to label tunnels for data forwarding. Ethernet tunnels of low traffic granularity can be encapsulated and aggregated to one tunnel with higher traffic granularity; this is precisely what enables flexible granularities traffic in 100GET.

PBB-TE is also known as MAC-in-MAC which encapsulates the native MAC frame into a backbone MAC frame [9]. Forwarding is based on a 48bit backbone destination MAC address (B-DA) and a 12 bit backbone VLAN ID (B-VID), so that transport network can be highly scalable with 16 million independent service instances (I-SID). For deterministic transport quality, PBB-TE disables some typical Ethernet mechanisms. MAC learning is disabled by using an external control plane to pre-configure tunnels. STP (Spanning Tree Protocol) and flooding mechanism are shield due to their inefficient properties. PBB-TE is backward compatible with the existing legacy Ethernet networks, which is an important feature given that Ethernet devices have been deployed everywhere. At the same time, PBB-TE can also accommodate the emerging IP-centric networking technologies.

The second standard, i.e., T-MPLS [10] (ITU-T G.8110.1), pushes an additional header in front of the Ethernet frame in form of a 20 bit per port swappable label to address a pre-defined backbone tunnel. Although highly compatible with IP networks, T-MPLS is not well compatible with some non-IP services. The third technology for carrier Ethernet transport network is VLAN-Switching which extends VLAN tagging and doubles the tagging field [11]. Although simple, the stacking of VLAN tag will cause overloading in Carrier-Grade Ethernet and limit its scalability by the number of available tags with an upper bound of 4096.

Among the three carrier-grade Ethernet technologies proposed so far, in our 100 GET-E3 architecture we recommend PBB-TE, as it is developed on base of traditional Ethernet forwarding mechanism which is expected to be highly backward compatible, facilitating seamless evolution to carrier-grade Ethernet.

III. END-TO-END SERVICE PROVISIONING

Efficient end-to-end service provisioning with QoS guarantees to end-users is of highest importance to the success of the carrier-grade Ethernet. The term of end-to-end may refer to a number of different scenarios. For a given service provider, end-to-end may imply provisioning a service across its national backbone to interconnect two Ethernet metro networks of a
specific customer. For another, end-to-end may include portions of metro and access networks or National Research and Education Networks (NREN). In this section, we first draw our attention to the emerging issues of end-to-end service provisioning in the carrier-grade Ethernet based on PBB-TE. To this end, we discuss a generic network scenario in which we illustrate three typical end-to-end services. After that we discuss issues and challenges in inter-domain service provisioning, which we plan to address in 100GET-E3. They include scenarios with diversified transport technologies and multiple domains.

A. Generic multi-domain network scenario

Clearly, end-to-end service provisioning will inevitably involve multiple technologies, multiple granularities and multiple domains. The notion of “multiple granularities” is referred to the case where a service goes through networks with different transport rates (e.g., line rates of OC-1 and OC-3), while the definition of a domain depends on different parameters. Usually a domain is administrative, distinctively characterized by its network topology (e.g., mesh, ring or star). In addition, a domain can be characterized by routing policy, or visible address space. Networks with different transport plane technologies can also be assigned to different domains. For instance, one domain maybe SONET/SDH and the other is carrier-grade Ethernet transport based on optical transport network (OTN).

Let us assume an end-to-end service request between two customer networks C1 and C2, as shown in the generic multi-domain network scenario in Figure 2. In order to establish connectivity for end-to-end service, the two customer networks have options to traverse multiple domains. Domains D1 and D3 offer PBB-TE carrier Ethernet transport services over the optical layer directly (e.g., OTN standard [ITU-T G.709]). On the other hand, domain D2 uses the traditional SONET/SDH transport network. Domain D4 is IP/MPLS based core network deploying Label Switched Routers (LSRs). We assume in this example that the two carrier Ethernet networks D1 and D3 deploy the same technologies, but this does not have to be the case in general (e.g., PBB-TE in D1 and T-MPLS in D3). We here also do not distinguish between the customer types, which maybe Ethernet-based VLAN, IP/MPLS networks or IP-only networks, and we assume that they are “simple” Ethernet users with edge Ethernet switches connected to the carrier core networks. The customer networks can also be residential video/data home networks or an enterprise network. Three parameters can be taken into account when considering network segments that services traverse, i.e., traffic locality, granularity and technology. Traffic locality determines the end-points of the connections. Granularity determines the traffic rates of the service in each segment, while technology refers to the type of transport network technology. By varying the combination of these three parameters, we can find a number of examples of possible end-to-end services, such as four tunnels illustrated in Figure 2 and denoted as R1, R2, R3 and R3’.

As it can be seen from Figure 2, R1 interconnects the customer sites through two PBB-TE domains, i.e., D1 and D3. Route R2 includes only the IP/MPLS core. Route R3 and R3’ illustrate a case and its extension. In R3 an Ethernet tunnel is established which traverses two domains, i.e., D2 and D3. In 100GET-E3 networks, R1 is considered the best option for end-to-end service provision in which all carrier Ethernet networks utilize PBB-TE technology. Other route options will result in a number of approaches for end-to-end service and need a separate consideration, which is beyond the scope of this paper.

Finally, note that Figure 2 does not explicitly communicate switching hierarchies of every domain. For instance, a PBB-TE tunnel in D3 can be modeled as if it belonged to two sub-domains: one in the Ethernet layer, and the other in the WDM layer. Also, some level of traffic aggregation is always necessary at the network edges towards the customer. For instance, a higher level tunnel (e.g., MPLS) can request a lower level tunnel (PBB-TE), and the two networks can run two separate routing and tunnel setup instances.

As the purpose of this paper is to deal with end-to-end service provision in carrier-grade Ethernet in which data is predominantly transported in pre-established tunnels, we will next draw our attention to the tunnel setup within the R1, and after that we will discuss more complex issues that can arise from other tunnel examples. Recall that route R1 includes two PBB-TE carrier grade Ethernet networks, each running directly over the configurable optical layer.
B. Signaling flow for tunnel setup

In a typical multi-domain scenario, each domain is running its own control plane which carries challenges to the inter-domain interoperability and signaling for tunnel setup. For timely configuration and efficient resource utilization, control planes have to be capable of recognizing and handling local and global traffic, adapting to different networks with various switching technologies, service types and traffic granularities. Control planes of various network domains can be separated or integrated within the common framework. For instance, every PBB-TE domain shares the address space within the domain and can be either running its own control plane instance or integrating with the carrier Network Management System (NMS) instead or in addition to it. For integrated control plane frameworks, GMPLS is a viable candidate, since it is designed to support a variety of network technologies. In fact, IETF has already proposed a number of Ethernet-specific extensions to GMPLS [12] and also solutions for applying GMPLS for backbone MAC-based forwarding in carrier-grade Ethernet such as PBB-TE [13]. However, more extensions and new solutions are still needed to deal with interoperating among multiple domains in end-to-end services provisioning with carrier Ethernet networks.

In multi-domain carrier Ethernet scenarios, two basic approaches are possible for signaling of end-to-end tunnel setup, i.e., sequential and concurrent. These two examples of signaling flow for R1 tunnel setup are shown in Figure 3. Figure 3(a) shows the sequential setup example, where the tunnel setup in D3 is triggered only after the tunnel has been successfully setup in D1. First, the customer network C1 sends the request to the edge switch of D1, which further triggers the tunnel setup procedure in the same domain. While we illustrate this request with the familiar PATH message, with a note that RSVP may not be the only choice for signaling flow here and future carrier Ethernet signaling procedure may adopt different, or rather simpler reservation protocols. After receiving the RESV confirm message, PATH message is sent to the edge of D3 to trigger the tunnel setup procedure. Once both tunnels are setup successfully, PATH message is forwarded to the destination customer network C2, which will finally send RESV message to C1 as a confirmation of the successful end-to-end tunnel setup.

Figure 3(b) shows the concurrent tunnel setup example. The tunnel setup procedures are carried out in D1 and D3 at the same time, i.e., concurrently. In this scenario, multiple tunnels may exist in both domains towards the required destination. The edge switches of two domains should be able to exchange the traffic engineering and forwarding information (“Negotiation”) such that an optimal combination of tunnels can be selected. After C2 has sent the RESV confirmation back to C1, the signaling for end-to-end tunnel setup has finished. Finally note that these examples have not discussed the means to actually route the data. For instance, sequential mode may only work for a small number of domains, whereas the interconnections between each domain is known in advance. Concurrent signaling may be more successful for scenarios with large number of domains. In these scenarios, multiple path messages can be sent to multiple domains, and the most successful one can be chosen for the final end-to-end path setup. In that case, some form of coordination can be implemented in the example (b), similar to what has been proposed in [14].

IV. ISSUES OF MULTI-DOMAIN PROVISIONING IN CARRIER GRADE ETHERNET

Today’s PBB-TE networks provide end-to-end services in pre-configured tunnels with unique labels, also known as Eth-LSP (Ethernet Label Switched Path) [15]. The tunnels are provisioned with guaranteed bandwidth, QoS performance and fast recovery. To establish Eth-LSPs, two types of Ethernet switches are used in PBB-TE networks: BEB (Backbone Edge Bridge) and BCB (Backbone Core Bridge), each equipped with two components, i.e., I-component and B-Component [16]. Both of these two components carry a distinctive label, as shown in Figure 4. In the I-component, backbone MAC address <B-DA, B-SA> is used along with the instance ID (I-SID), assigned according to the type of service. In the B-Component, a different label is inserted in the form of <T-DA, T-SA, T-IID>, where T-DA/T-SA are backbone tunnel (B-tunnel) MAC addresses and T-IID identifies the tunnel instance type. Given this hierarchical label space, we next analyze three possible scenarios for multi-domain end-to-end provision such as in case of R1 from Figure 2.

---

2 Based on switch architecture proposed in 802.1Qay
A. Multi Domain Scenarios

1) Scenario I: Point-to-Point Concatenated Tunnels

In this scenario, an end-to-end service denoted as \( S_1 \) is provisioned from C1 to C2 with \(<\text{B-DA}, \text{B-SA}>=<\text{D-I2}, \text{A-I3}>\), as shown in Figure 5(a). Bandwidth required by \( S_1 \) is \( b \), while the capacities of the B-Tunnel-1 (in D1) and B-Tunnel-2 (in D3) are \( b_1 \) and \( b_2 \) respectively, and \( b_3 \leq b_1, b_3 \leq b_2 \). In other words, \( S_1 \) is to be setup as a concatenation of two point-to-point tunnels in two domains. Note that every tunnel in carrier Ethernet network is bidirectional, which implies that \( S_1 \) is also established from C2 to C1 with \(<\text{B-DA}, \text{B-SA}>=<\text{A-I3}, \text{D-I2}>\). T-space label is swapped between two domains, while B-space label is invariant from the head node to the tail node. Traffic is stripped off the tunnel header (<T-DA, T-SA, T-IID>) of D1 in BEB-B and enters to the edge bridge of D3, i.e., BEB-C, where the new tunnel header is added. Unaware of the transformation of the B-Tunnel layer (with T-space label), customers are offered a single tunnel (with B-space label) for end-to-end service provisioning.

2) Scenario II: Point-to-Multipoint Concatenated Tunnels

In this example, we make the same assumptions as previously, where the bandwidth required by \( S_1 \) is \( b \) but only the D1 can provide this capacity by single tunnel (e.g. B-Tunnel-1 in D1). In other words, multiple tunnels may be required in D3 towards the destination. This raises the issues of concatenation of the tunnels B-Tunnel-1 in D1, and B-Tunnel-2 and B-Tunnel-3, with capacities \( b_2 \) and \( b_3 \) respectively in D3, as shown in Figure 5(b). The two tunnels in D3 are established with an aggregate capacity equal or larger of the B-Tunnel-1, i.e., \( b_2+b_3 \geq b \), \( b_2 < b \), \( b_3 < b \). In this scenario, the end-to-end service is provisioned as a point-to-multipoint concatenated tunnel, within the same B-space label (<B-DA, B-SA>). To end-users, the service is still carried in “one” tunnel.

3) Scenario III: Shared Tunnels

In Figure 5(c), two services \( S_1 \) and \( S_2 \) are requested between C1 and C2, crossing two PBB-TE domains D1 and D3. Here, service provider in D1 is able to set up tunnels for each service as requested, which is similar to the point-to-point concatenated tunnels in Scenario I. For efficient resource utilization, however, partly shared tunnels can be considered. Let us assume that bandwidth required by \( S_1 \) and \( S_2 \) is \( b \) and \( b' \) respectively. The capacity of B-Tunnel-1 is \( b_1 \), \( b_1 \geq b+b' \). Tunnels B-Tunnel-2 and B-Tunnel-3 in D3 are setup with the capacity of \( b_2 \) and \( b_3 \) respectively, here, \( b+b' \geq b_2, b+b' \geq b_3 \), \( b_2 \geq b, b_2 \geq b' \), \( b_3 \geq b \), \( b_3 \geq b' \). Services \( S_1 \) and \( S_2 \) with different B-Space label (<B-DA, B-SA>) can share the same B-Tunnel with same T-Space label <B-B2, A-B1, Va> in D1. Data carried by these two services are encapsulated at BEB-A and added their own <B-DA, B-SA> labels. After that, at the B-component of the BEB-A, they are aggregated in the same CBP with MAC address A-B1 to B-Tunnel-1. At the egress of D1, BEB-B strips off the label <B-B2, A-B1, Va>, forwards them to the edge bridge of D3 where \( S_1 \) and \( S_2 \) are provisioned as two tunnels. Note that mechanisms of shared forwarding can be used in this and in similar scenarios, which is the subject of our future studies [12].

B. Open Issues

A) Technology domains

More complex routes may include a combination of circuit-oriented transport technologies, including SONET/SDH network, reconfigurable DWDM and new carrier-grade Ethernet networks, with various forwarding technologies such as VLAN-Switching, T-MPLS or PBB-TE. The simplest way to address the technology heterogeneity is to define one common underlying transport layer, and specify that any gateway equipment must support this layer. Otherwise, it may be necessary to define protocol and bandwidth adaptation functions. For example, the Virtual Label Switch Router has been proposed in DRAGON and CHEETAH to solve adaptation issues for non-GMPLS compliant network elements [17]. From the service provider’s point of view, however, any additional interfaces or adaptation functions may lead to an increased operational costs and investments.

B) Path selection

Previous examples have shown that it is important to choose a best-suited route that can leverage bandwidth utilization and traffic granularity adaptation between two domains according to some specific objectives. Parameters of resource allocation, bandwidth, latency, etc, for intra-domain connections have to be advertised for inter-domain traffic, while at the same time preserving the integrity of every domain and respecting the scalability of signaling [18]. In addition, the requirements for scalability and resource visibility for each domain are different which also put challenge to routing and resource allocation.
Expected are functionalities equivalent to those for packet-based MPLS, which provide a framework for establishing traffic engineered LSP across multiple MPLS/GMPLS domains [20]. Similar framework is needed in carrier-grade Ethernet for routing and path computation.

C) Control plane
The complexity of the multi-domain control plane architecture can be described by three basic parameters: a) technology properties of individual domain, b) properties of traffic within each domain (domain internally, local), and c) properties of domain external traffic distribution (inter-domain). Whereas the control plane in today’s carrier networks is exclusively used for creation of services within one domain, the configuration management involving requests for connections and services in multiple domains are typically implemented through network management applications. For efficient carrier Ethernet to become reality, issues of dynamic but simple control plane and scalability are still open and have to be addressed.

V. SUMMARY
In this paper, we presented the conceptual 100GET-E3 architecture and proposed three multi-domain service provisioning scenarios in PBB-TE networks as well as discussed issues of end-to-end service provisioning in carrier-grade Ethernet networks with multiple granularities and domains. The cluster E3 of project 100GET is taking pioneering steps in innovating next generation transport network based on Ethernet technologies (Ethernet-over-DWDM) and this paper drafted the first discussions and possible directions within some of the topical areas.

ACKNOWLEDGMENT
This work has been supported by the BMBF project grant 100GET-E3.

VI. REFERENCES
[17] Q. Song, Z. Li, I. Habib “CheetaH VLSR Design and Deployment in GMPLS Optical Networks", GLOBECOM '06, IEEE, pp. 1-6

(a) Point-to-Point Concatenated Tunnels
(b) Point-to-Multipoint Concatenated Tunnels
(c) Shared Tunnels

Figure 5 Multi-domain End-to-end Provision Scenarios