A Survey of Inter-Domain Peering and Provisioning Solutions for the Next Generation Optical Networks

Mohit Chamania *Student Member, IEEE,* and Admela Jukan *Senior Member, IEEE,*

***Abstract*—The emergence of carrier grade transport technologies has led to a paradigm shift in inter-domain routing which became an important feature of the transport layer based on optical transmission and switching. While the new technologies have capabilities to provide end-to-end guaranteed quality of service (QoS), the lack of inter-operability between different technologies, administrative areas and control planes makes inter-domain peering and provisioning below the conventional Internet Protocol (IP) layer a challenge. In this survey, we analyze various multi-domain routing models for emerging Layer 2 and WDM switched networks which have been proposed till date, and based on that survey, we highlight some open issues and future challenges pertaining to scalability, reliability, multi-domain QoS, control plane interworking and dynamic peering.**

***Index Terms*—Inter-domain, QoS, protection, routing, signaling, path computation, control plane.**

I. Introduction

THE UBIQUITOUS presence of the Internet coupled with the increasing demand for dedicated large-scale private networks has made it imperative that multiple carriers and domains interconnect with each other in a scalable manner. At the same time, a multitude of emerging switching and routing technologies at different tiers in the OSI model have led to diversity in carriers' choice of networking infrastructures. While conventional end-to-end solutions required each domain to exchange data using a common (IP) Protocol, the emergence of the control plane and next generation applications with high-demands for bandwidth have resulted in inter-domain peering and provisioning paradigms below the IP layer. This new paradigm not only encompasses reconfigurable optical network technologies, including optical switching and routing based on principles of Wavelength Division Multiplexing (WDM), but also the newly emerging Layer 2 (L2) transport technologies, such as carrier-grade Ethernet [1].

While inter-domain routing in the Internet is a widely researched subject, issues of inter-domain routing, peering and provisioning at configurable layers below the IP layer is a relatively new concept. Although multiple research initiatives are addressing the inter-domain issues in the optical layer, each of them is unique in its approach and perspective. For instance, the high-performance computing research community (Grid computing) is studying inter-domain control and management issues with special emphasis on application requirements for inter-domain provisioning [2], [3], [4], [5]. On the other hand, the Internet Engineering Task Force (IETF) standard drafts are addressing signaling extensions to the current IP protocol suite to be used by the reconfigurable optical layer, and more recently by the Ethernet layer, such as Generalized Multi Protocol Label Switching (GMPLS) and GMPLS controlled Ethernet Label Switching (GELS) [6]. Yet another group of industry and academic researchers is concerned with the inter-carrier relationships and addressing specific issues of policies and service level agreements [7]. In summary, efficient provisioning of high-bandwidth connections between the multiple domains separated by technologies, administrative rules, and control and signaling concepts is an open challenge.

To address this challenge, we study the emerging requirements and mechanisms for inter-domain provisioning and present a review of the various technologies and solutions in use today. As result of the survey, we reveal open issues and discuss potential avenues and research trends.

The survey is organized as follows. In Section II, we describe the various functions required in an inter domain routing model. In section III, we identify and describe main features and mechanisms used for topology dissemination, path computation, quality of service, signaling, protection and control plane interaction in Layer 2 and WDM networks. Section IV presents an illustrative, hypothetical inter-domain architecture for carrier-grade Ethernet networks using the different inter-domain routing concepts examined here. Section V highlights some open issues to be addressed in future works and Section VI concludes the paper.

II. Motivation

Fig. 1 illustrates a typical inter-domain network scenario in the layers below IP. A few characteristics are imminently apparent from this example. First, there has not been a de-facto technology, which has replaced all other transport solutions. In fact, the technologies shown here co-exist and include WDM, SDH/SONET and, most recently, carrier-grade Ethernet. Second, physical interconnectivity requires the knowledge and compatibility among the technologies used in the domains to be able to route traffic between them. In other words, in order to route traffic across multiple domains, the protocols and interface have to consider specific information, such as number and color of free wavelengths, available aggregate bit rate at each hop, and adaptation for framing and packet processing at the domain's edges. Third, the presence of multiple vendor-specific control planes along with IETF and ITU-T standards, such as GMPLS and Automatically Switched Optical Networks (ASON), require interoperability of signaling and control in interconnecting multiple domains. Finally, in a highly competitive market,

Manuscript received 28 May 2007; revised 26 August 2008.

The authors are with Technische Universität Carolo-Wilhelmina zu Braunschweig (e-mail: {chamania,jukan}@ida.ing.tu-bs.de).

Digital Object Identifier 10.1109/SURV.2009.090104.

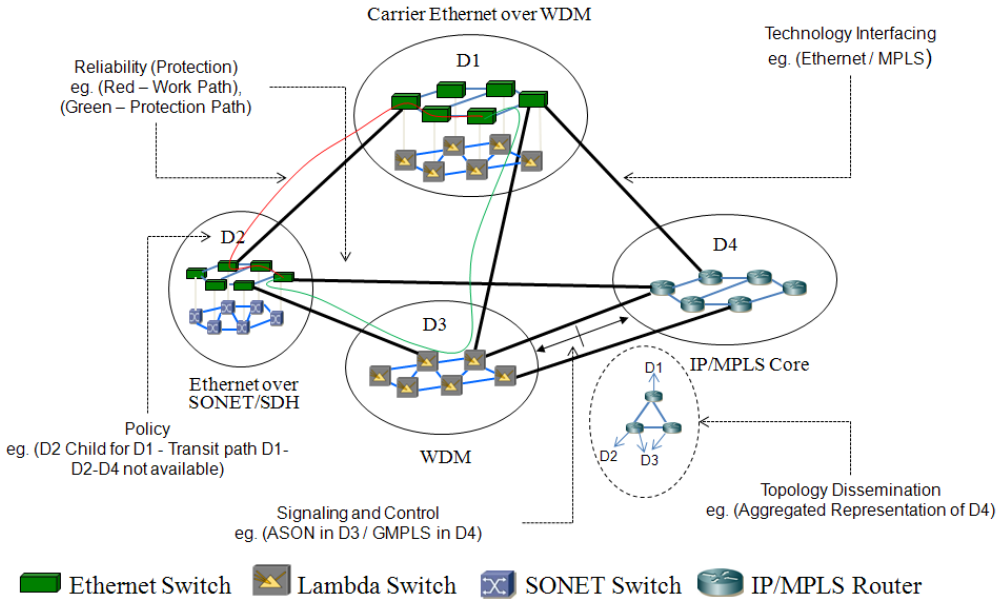


Fig. 1. A generic inter-domain network scenario.

carriers like to exercise control on the use of their resources in a multi-domain scenario. Therefore, decisions such as routing, admission control, cost etc. are not only driven by network parameters but are also influenced by the policy in place between different carriers. Let us first summarize the factors involved in the problem area.

Topology Dissemination and Routing

The sheer size of the global network poses a challenge for routing in multi domain networks. While path vector protocols such as Border Gateway Protocol (BGP) have been used extensively for IP networks, the need to provision paths with multiple quality-of-service (QoS) constraints, such as the path signal quality and cost, requires that the routing protocol have the necessary information about the end-to-end path. Different domains typically do not share complete topology and link state information with each other due to administrative reasons. Therefore, a mechanism is required which can facilitate QoS routing without divulging complete topology information.

Service Differentiation

Different applications and services such as Internet Protocol Television (IPTV), Triple Play, Storage Area Networks (SAN's), Virtual Private Networks (VPN's) etc. have different demands on the network. Given a diverse set of applications working on the same network, a multi-domain routing model should be able to set up paths which satisfy a given set of constraints, usually defined in terms of bandwidth offered, signal quality, end-to-end delay and jitter. While most carrier transport technologies have mechanisms to deliver QoS, different technologies and measurements metrics in different domains and restricted information exchange between domains hinders the set up of end-to-end QoS paths.

Policy Control

Policy considerations in routing which are implemented using the Border Gateway Protocol (BGP) at the IP layer are also required in the lower layers. For instance, in GMPLS parlance every connection is a label switched path (LSP), which among other parameters is defined based on the traffic granularity supported by the existing interfaces, such as WDM multiplex. For every such LSP, the policy may not only define a set of end-points but also a set of domains that may be involved in provisioning of this service, similar to what can be inferred from the AS_PATH attribute of the BGP. Additional policy-based parameters may include issues of network access control and security.

Reliability

All factors mentioned before require reliable network operation, beginning with technology interfacing over control plane to differentiated reliability requirements by the customers. Especially, with customers demanding guaranteed service from network providers, it is necessary to have mechanisms to recover from path failures by switching to the alternative paths and routes. While most protection schemes in place today work under the umbrella of a single control plane, efficient mechanisms are in demand to protect paths in an end-to-end manner in a multi-domain network.

Technology Interfacing

Technology interfacing not only implies the detailed knowledge of the technologies used in the neighboring domains for compatibility, but may determine the path routing based on that knowledge. For instance, for a path provision along the domains D2-D3-D1 in Fig.1, D3 may be chosen due the fact that it is a pure WDM domain, compatible to carry any carrier-grade Ethernet technologies.

It should also be noted that although both D1 and D2 carry the same type of traffic (Ethernet), a simple interconnect may not be always compatible due to the existence of proprietary solutions. Technology interfacing can be used to describe interworking between different transmission technologies at the same layer (for example Ethernet and ATM) as well as interworking between technologies at different layers, such as SONET-over-WDM. [1]

Signaling and Control

Each network domain typically runs an independent control plane. To enable end-to-end connectivity, it is necessary to have a signaling mechanism to exchange information between domains. With each vendor choosing a different control plane technology and some with proprietary control planes, signaling mechanisms that are compatible with every control plane are currently not in place. Some control plane solutions, such as GMPLS, are designed to provide a unified control plane for various switching technologies and hierarchies. However, much work is needed to realize a single standardized control plane serving the administrative and technological needs of different carrier grade systems.

In the following sections, we provide a detailed survey of some of these requirements as well as solutions which have been proposed within the research studies of inter-domain peering and provisioning.

III. Main Features of Inter-Domain Provisioning

The different functionalities required to develop an inter-domain provisioning paradigm can be classified as routing, QoS provisioning, signaling, protection, and inter-domain control plane interactions. Each of these features are inter-linked with each other and together form a complete inter-domain provisioning framework. We now discuss each of these features in detail.

A. Routing and Path Computation

Routing in the inter-domain scenario poses a challenge, not only due to the size of the network but also due to the reluctance of different providers to share information. Providers use information like topology, link weights etc. as a competitive edge and are therefore reluctant to exchange this information. A few theoretical approaches have been proposed to solve the problem of inter-domain routing in a distributed manner without exchanging large amounts of information. Shrimali et al. [8] present a theoretical approach which can achieve a proportionally fair solution between two peering domains while exchanging only minimal information between domains. While the proposed solution is novel and can achieve fair solutions, it does not consider the case where domains route transit traffic. Another distributed theoretical approach is presented in [9] which splits a global optimization problem using Lagrangean decomposition. In contrast to [8], the proposed solution in [9] requires exchange of flow vectors for incoming and outgoing links for each domain in the network, which in turn poses challenges on synchronization of the information. In addition, given the size of the present networks, the amount of information exchange required enforces practical constraints on scalability. While [8] does identify the possibility of a domain advertising incorrect flow vectors, both works can be combined for a complete analysis on the effects of inaccurate information advertisement by domains in the network.

Currently deployed inter-domain routing schemes are based on more practical approaches. Inter-domain routing can be distinctly classified into two distinct functions: 1) Exchange of information between domains to determine location of hosts in the network and abstract topology information to reach these destinations and 2) A mechanism to determine the exact inter-domain path in order to serve a request for an inter-domain connection.

1) Topology Information Dissemination: Topology information is required to facilitate routing in a network. To this end, various algorithms for topology information dissemination have been used to automate the process. Most of these algorithms depend on link state updates to determine an abstract network topology. In a multi-domain model, not only would the scale of the topology make it impossible to always share complete topology information, but also different carriers would not like to share their topology information with their competitors. To overcome this, different classes of solutions have been presented, including topology aggregation mechanisms, path vector protocols and partial/full topology information dissemination.

Topology Aggregation Mechanisms

In inter-domain provisioning scenarios, topology aggregation is used for two main reasons - scalability and confidentiality. A domain's topology is represented by an aggregated logical network topology as illustrated in Fig. 2. For instance, each domain in the multi-domain topology shown in Fig. 2(a) is represented as a single node in Fig. 2(b) while in Fig. 2(d), each domain is reduced to a full mesh between its border nodes. By aggregating the topology of a domain, not only is the actual topology inside a network hidden from other domains, but the end-to-end path computation problem is performed over a reduced number of virtual nodes, with each domain calculating the real path corresponding to the advertised virtual path inside the domain.

Previous work addressed the topology aggregation mechanisms within three large categories: (i) single node/symmetric node, (ii) star, and (iii) full mesh topology. In the single node/symmetric node scheme, a domain is represented by a single node (Fig. 2(b)) and although very scalable, this scheme does not show the resources available inside the domain and is therefore inefficient and results in suboptimal paths and higher blocking probabilities. In the star topology (Fig 2(c)), the logical topology consists of all the border nodes in the network connected to a single virtual node. The size complexity of the star topology is $O(|M|)$, where M is the set of border nodes in a domain. The branches of the star can be given virtual costs to depict the cost of traversing through the network. Given that each node in the star is connected to a virtual node, it may be useful for representing available bandwidth, but is not ideal

[1]It should be noted that multi-layer issues, which can also be addressed in the context of multi-domain provisioning and peering are outside the scope of this paper and deserve a separate study.

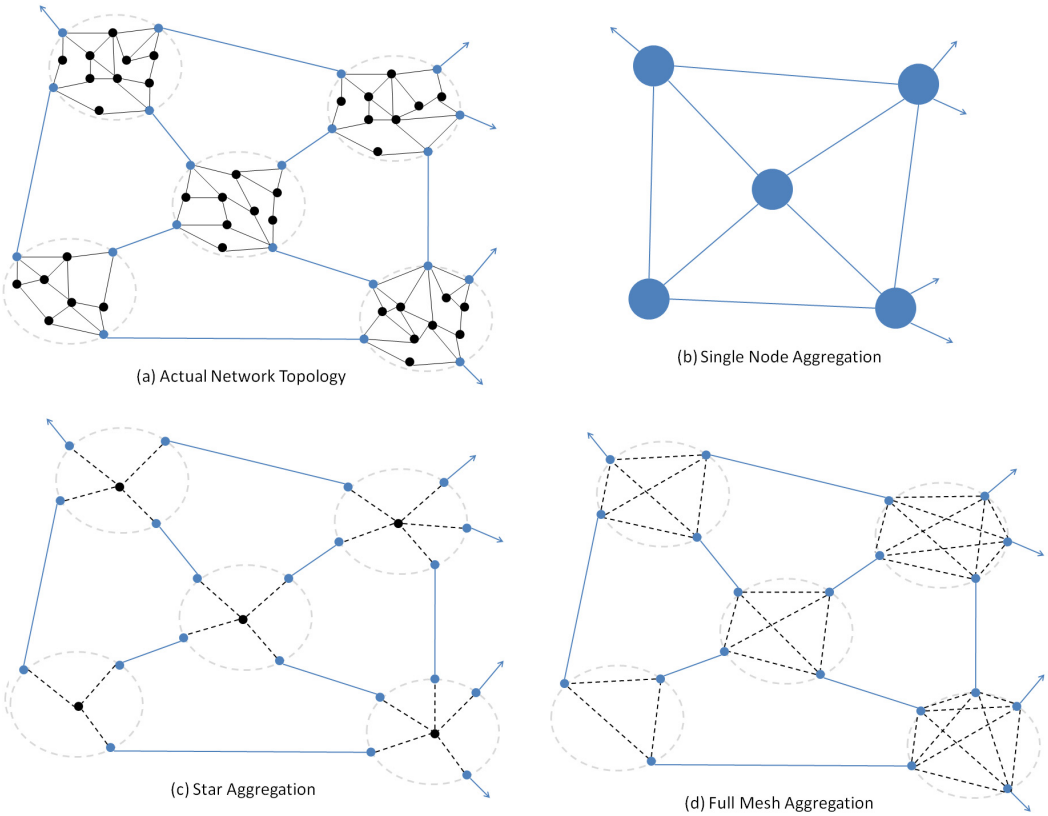


Fig. 2. Examples of different topology aggregation schemes

for representing wavelength connectivity between two border nodes. Therefore the star topology is not widely used as an aggregation scheme for optical networks. Finally, the full mesh topology aggregation scheme (Fig. 2(d)) is the representation of a domain as a fully connected mesh graph of its border nodes. The interconnections inside the mesh graph advertise the amount of resources available when going from one border node to another and is therefore a better mechanism than the single node and the virtual star topology aggregation scheme.

Aggregated topologies are exchanged between domains, which are then used to calculate optimal domain hops from a source to a destination domain. Mechanisms proposed in [10], [11], [12] depict how aggregated topologies can be used to calculate QoS inter-domain paths, and even calculate diverse autonomous system (AS) paths in a network. The aggregated topology can either be distributed to trusted domains, or can be sent to a central server. The central server can process the relevant information and then send reachability information to all domains. We now describe some aggregation schemes which have been proposed for different layer-1 and layer-2 networks.

Two studies [13], [14] compare the performance of the single node aggregation scheme and the full mesh in a WDM network. It is found that the single node abstraction is not efficient when applied to a WDM network, even in the presence of wavelength conversion at the edges of domains. The two studies describe a full mesh topology aggregation scheme for a physical layer inter domain routing scheme, with each virtual edge between two border nodes containing cost and wavelength availability information. The size complexity of the update message is $O(|M|^2)$, as there are $(|M| * |M-1|)$ links in the full mesh representation. The wavelength availability information between border nodes in the full mesh aggregation scheme leads to lower blocking probabilities when compared to the single point topology. It should also be noted that the inter domain update frequency increases in a full mesh aggregation scheme as link capacities as advertised in the full mesh are affected with intra domain traffic. Sanchez-Lopez et al. [15] propose an enhancement of the full mesh topology aggregation scheme for ASON networks. The full mesh topology now includes information about the availability of multiple paths between a pair of border nodes with propagation delay information, wavelength availability and count (in a multiple fibre case) for each path. The additional information is useful in setting up protection paths in the domain and also determining the end-to-end delay in the network.

Wan et al. [16] use the bi-directional shuffle-net topology as an aggregated topology in WDM optical networks. The border nodes of a domain are mapped on to a shuffle-net topology using a genetic algorithm. The topology is formed by reducing the full mesh representation of the domain to the bi-directional

shufflenet which has a smaller size complexity than the full mesh and better performance than the symmetric star approach as shown in this paper.

Finally, a unique topology aggregation scheme is presented in [17] useful for setting up end-to-end link disjoint paths in a label switched network. The complexity of the proposed aggregated topology is of the order of $O(|M|^4)$, and is built upon the full mesh aggregation scheme. The algorithm proposed by Suurballe [18] is used to determine link disjoint path pairs in the directed graph. In the proposed algorithm, optimal disjoint path pairs are calculated in a two step process, where the optimal end-to-end path is evaluated first, and then the link disjoint pairs are calculated by reversing the directions of the links used in the optimal path and then re-calculating the optimal path. In order to support the algorithm, each link in the full mesh aggregation is associated with another full mesh aggregated topology, which indicated the path costs in case the concerned path is used in the first optimal path computation cycle.

Path Vector Protocols

Path vector protocols have been proposed for optical networks based on the success of BGP at layer 3. The OBGP (Optical BGP) protocol [19], [20], [21], which is an adaptation of the BGP protocol at the optical layer is used to determine lightpath availability across multiple domains in a network. The OBGP protocol is used to join different administrative domains and to create smaller management clouds inside a large WDM domain to reduce the complexity of the routing and wavelength assignment problem. Each optical switch at the edge of a WDM cloud runs an OBGP server instance, and exchanges wavelength availability information between domains. OBGP servers advertise the reachability to other OBGP servers along with a set of wavelengths that can be used for the connection. Other OBGP servers use this advertisement and update their reachability information. For example, if domain C is reachable via domain B on wavelengths $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ and domain B is reachable via domain A on wavelengths $\lambda_2, \lambda_3, \lambda_4$ on the same edge router, then domain C is reachable from domain A on wavelengths $\lambda_2, \lambda_3, \lambda_4$.

The protocol can also be configured to allow for policy enforcement, just as in traditional BGP; however, a separate signaling mechanism is required to signal the inter-domain path setup request. St. Arnaud et al. [19] describe the motivation behind the use of BGP in optical networks, and give a basic framework for deploying OBGP. Francisco et al. [20] define the additional messages required in the BGP protocol to support routing for inter-domain light-paths. Another path vector protocol called Constraint-based Optical Path-vector Routing Protocol (COPRP) for inter domain connection setup has been proposed by O. Yu [22]. COPRP has a similar working model to OBGP but allows for advertisement of multiple routes to the same destination, which are calculated based on different sets of constraints to such as minimum number of inter-domain hops, optimized load-balance, minimum cost, delay etc.

Yannuzzi et al.[23] highlight the lack of traffic engineering (TE) information sent in OBGP messages and also highlight that an OBGP message can only find a single path. They propose an extended path vector protocol, and propose the use of an Inter Domain Routing Agent (IDRA) to advertise reachability information. The use of a centralized IDRA implies that reachability information as advertised by the domain for different domains is processed at a single point and can give better representation of the available resources inside a domain. The protocol also supports the advertisement of multiple paths to a destination which is not supported by BGP/OBGP.

Path vector protocols are a tried&tested approach for inter-domain route computation, with BGP being the de-facto algorithm used for inter-domain routing at the IP layer. However, changes in a BGP system take a long time to converge, and BGP is therefore not suited for systems where available resources change frequently. While path vector mechanisms such as OBGP may be applied to optical networks which have a very large granularity and low update frequency, they may suffer from stability and scalability problems in layer 2 networks which can experience frequent change in resources.

Partial Topology Information Dissemination

Partial topology information dissemination is used to advertise resources in the network without disclosing the full topology of the domain. Partial topology information dissemination is useful as it allows the domain to control the amount of topology information it shares with different domains. The work presented in [24] uses such a scheme for optical networks by using centralized servers in each domain which communicate with each other and control the exchange of control plane information between the two domains. The servers dictate the extent of domain information exchange with connected domains based on the relationship between the interconnected domains. The information exchange is controlled to define the relationship between two domains, which can be peer-to-peer or a hierarchical overlay.

Full Topology Information Dissemination

Full topology information models have not been used extensively, as service providers are reluctant in sharing topology information. Also, in a large scale multi-domain network, full topology dissemination will increase the computational complexity and the volume of inter domain signalling in the network. For completeness, performance of a global shortest path first (SPF) algorithm with full topology information dissemination is compared with a per domain SPF algorithm for light-path setup in a WDM network in [25]. The authors observe an increase in blocking probability when applying per-domain routing and wavelength assignment (RWA) schemes as compared to a global RWA scheme, as expected, but conclude that full topology dissemination is not desirable due to issues with scalability.

2) Constraint Based Path Computation: While topology dissemination mechanisms as described in the previous section can determine loose routes spanning multiple domains, it is necessary to have a mechanism to determine defined hop-by-hop routes with respect to constraints such as bandwidth, total delay, signal quality, protection and cost, while efficiently utilizing the network resources to prevent over provisioning. Path Computation Elements (PCEs) have

been proposed as a major class of centralized solutions for route computation; in this paper we discuss the applicability of the PCE framework to multi-domain networks. Signalling based approaches have also been deployed to determine paths in a multi-domain networks.

Path Computation Elements (PCEs)

A path computation element (PCE) is a network entity which holds the topology information and is polled by various nodes to determine the path from a source to a destination. The PCE uses the loose domain hop information obtained from topology information dissemination mechanisms described earlier to calculate optimal end-to-end inter-domain paths. The use of a path computation element eliminates the need for every node within the network to compute the path, and all link state information is sent only to the PCE. A single domain can have multiple PCEs to facilitate load sharing and avoid single point of failures. In a multi-domain scenario, PCEs in different domains may share information with each other to compute paths. PCE based routing architectures of use in multi-domain networks can be classified into two major groups: (a) Peer-to-peer and (b) Hierarchical.

In a peer-to-peer model, PCEs of neighbouring domains create peering relationships and interact with each other to exchange routing information via mechanisms of the control plane. To establish a multi-domain route, PCEs are probed in a sequential fashion to determine the availability of the path. If available, the return message of the signal is the next hop information between different domains, as discussed in [26]. The PCE peering mechanisms are not only useful in the circuit switched WDM networks, but as shown in [27], they can be used in a similar way to set up inter-domain paths in Optical Burst Switching (OBS) networks.

Yannuzzi et al. [24] propose the use of intelligent PCEs which customize inter-domain route computation and link state announcements according to the relationship between two domains. For example, peering PCEs can be configured to share information as peer-to-peer domains when connected at the network core, while domains near the edge of the network are configured to create a hierarchical routing overlay at the edge of the network.

Okumus et al. [28] proposes a hierarchical PCE approach, which uses a single PCE for each domain and a centralized global PCE. The latter uses aggregated information from each domain to calculate the optimal inter-domain path. However, a single PCE as the centralized controller has limited scalability and is also a single point of failure in the network, and should be duplicated for reliability. Multiple hierarchical levels were first introduced in the Asynchronous Transfer Mode (ATM) Private Network to Network Interface (PNNI) [29] architecture. The multi-level hierarchy (up to 104 levels) allows the system to encompass a large number of domains, and the hierarchical structure allows for automatic loose route discovery in the network. The PNNI control plane has been used extensively in ATM networks and has inspired the ASON control plane. The ITU ASON draft [30] extends the PCE hierarchy to be applied in ASON networks, by increasing the number of tiers of aggregation. Multiple PCEs at one level of hierarchy are connected to a parent PCE, which aggregates wavelength availability information from the child PCEs and represents a larger domain area. Similarly domain areas are further clustered under a single PCE to form larger domain areas, thereby increasing the scalability of the system as shown in Fig. 3. Similar hierarchical PCE models have also been proposed in [31] which uses hierarchically distributed PCEs to create IPTV trees for multi-domain users. The use of hierarchical PCEs reduces the computational load at each PCE in creating the tree. Matsuura et al. [32] propose a framework in which hierarchical path computation servers can be used to calculate inter-domain paths in GMPLS networks. In this work, a domain is assumed to be the leaf in the hierarchy, and is represented as a single node at the next hierarchical level. In this manner, inter-domain paths can be computed in a scalable manner. However, as domains are reduced to single nodes at higher hierarchical levels, there is loss of optimality in the computed path. It should also be noted that inter-domain signalling between two PCEs and inter-domain path setup signalling are different functions. In fact, in a dynamic multi-domain network, it is possible that the route determined during path discovery using PCEs may not be available during path reservation calls due to the delay encountered through signalling.

The use of PCE for controlling Layer 1 and 2 multi-domain networks has been proposed in manifold works surveyed here, i.e., [24], [27], [28], [29], [31], [33], [34], [35]. A brief overview of the proposed approaches is shown in Table I. In recent times, the IETF has led standardization efforts for the PCE framework in MPLS/GMPLS networks. The PCEP [36] protocol defines the protocol for communication between a requester and a PCE. The same protocol is also used to achieve inter-PCE communication. The Backward Recursive Path Computation (BRPC) protocol [37] has been proposed to compute optimal inter-domain paths in a multi-domain network. The BRPC assumes that an inter-domain sequence is provided from the source to the destination before the actual path computation. Using this sequence, the PCE protocol is used to contact contiguous domains in a sequential manner. On reaching the PCE of the destination domain, the destination PCE returns a tree of possible paths to the destination from all of the possible ingress nodes. Each PCE extends this tree to the ingress nodes inside the respective domains and sends it to the previous PCE. On reaching the source PCE, the optimal end-to-end path from the received set of paths is selected for the inter-domain path setup. It should be noted that only the optimal paths to the ingress border nodes are considered in the tree, thus ensuring that the proposed mechanism is scalable. To hide the topology inside the domains, a path-key mechanism is proposed in [38]. Here, each path as computed by a PCE for an inter-domain path setup is mapped to a key. The key is sent in the PCE message, and is stored in the Extended Route Object (ERO) in the Resource Reservation Protocol (RSVP) during path setup. At the ingress border node of each domain, the path key is sent to the PCE to determine the path inside the domain, thus maintaining topology confidentiality.

Proposals have also been made to extend the PCE framework to facilitate multi-layer routing. Requirements for multi-layer PCE frameworks are defined in [39] and a few basic frameworks for the same are proposed in [40].

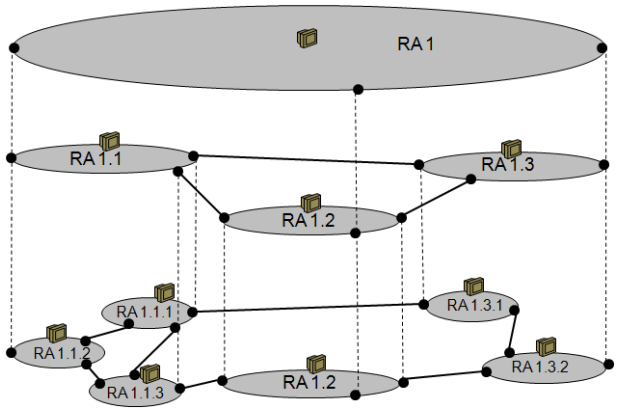


Fig. 3. Routing area hierarchy using PCEs in ASON [30]

The PCE framework proposed in the IETF drafts is compared against per-domain signalling schemes in [41]. Various metrics are used to study the performance of the PCE framework against the signalling with crankback framework. The results show that the PCE framework can set up LSPs with lower costs as compared to the signalling framework, while also reducing the possibility of failure during path computation.

The use of the path computation element also opens up the possibility of pre-computation. Pre-computation schemes compute solutions a priori for a large set of possible parameters. These parameters are then used during path computation on arrival of a request, thus increasing response time and scalability and reducing the computational load on the PCE. Orda et al. [51] propose a pre-computation scheme which can be applied to multi-level PCE hierarchies which are proposed in the ASON and the PNNI control planes. The proposed mechanism defines pre-computation schemes for both additive QoS metrics (delay) as well as bottleneck weight metrics (bandwidth). Pre-computation schemes combined with the PCE framework can significantly increase the performance of constraint based routing in multi-domain networks. However, a more detailed analysis highlighting parameters such as pre-computation overhead, frequency of computation cycles and signaling load is required to analyze the benefits of the pre-computation schemes.

Per-Domain Path Computation

The per-domain path computation approach computes optimal path segments in a limited visibility. In other words, the loose domain hop information is used to determine the best egress border node inside a domain and the inter-domain routing protocols are used to calculate the optimal path to this node. This process is carried out in each intermediate domain to obtain and end-to-end inter-domain path. It should be noted that distributed shortest path computation approaches can be realized using signaling protocols such as RSVP [42] or using PCEs.

In a per-domain approach, each border node announces the reachability information for different destinations inside the domain, similar to intra-domain routing models. However, unlike traditional intra-domain algorithms, the lack of frequent updates in an inter domain system leads to inaccurate network state information and hampers performance of traditional distributed algorithms. To illustrate this trend, Zhou et al. [43] use simulations to observe the performance of routing and wavelength assignment algorithms under inaccurate network state information conditions. Three RWA schemes (random fit, first fit and most used) are used to set up multi domain light-paths, and imperfections in the link states are implemented by using a timer based link state update. It is seen that algorithms such as first fit and most used which perform better than random fit in networks with accurate state information perform very poorly under inaccurate state conditions.

Bypass based routing has been proposed for routing in optical [44], [45] and layer 2 [46] networks. In bypass based routing, prior to the network state update, links are identified that are most likely to get congested. An explicit bypass path is defined for these links and sent with the path setup message to ensure that path setup is not hindered. Sanchez-Lopez et al. [15] have demonstrated the use of a bypass based routing algorithm [45] for routing in inter-domain WDM networks, and have compared the performance of bypass based routing against first fit shortest path routing in multi-domain networks. Two different aggregation schemes are chosen to represent the domain topology, and simulations show that bypass based methods reduce the blocking probability as compared to first fit shortest path routing for both schemes. Signaling approaches are also used to increase the efficiency of light-path setup in multi-domain networks. Salvadori et al. [47] present different signaling based approaches to set up end-to-end QoS light-paths, and compare the performance of these paradigms against the use of QoS aware routing algorithms to setup light-paths in a multi-domain WDM network. One of the proposed signaling approaches to set up impairment aware light-paths has comparable performance with the QoS constraint based routing approach in terms of blocking probabilities. Although path setup times in signaling approaches is higher than QoS routing, QoS routing requires frequent inter-domain updates with QoS parameters and require high computational power due to their complexity.

TABLE I
OVERVIEW OF WORK ON PCE-BASED INTER-DOMAIN ROUTING APPROACHES

Reference	Description	Hierarchical PCE	Peer to Peer PCE
24	Model for routing using intelligent peering PCEs in a multi-tier network; hierarchical overlay topologies near the edge, and peer-to-peer topologies in the core.		X
27	Multi domain routing in Optical Burst Switched (OBS) networks; uses specific nodes to run complex routing algorithms;		X
28	Use of PCE to set up LSP's in the network; use of hierarchical labelling to aggregate Differentiated Service (DiffServ) flows between two domains;	X	
29	The ITU draft for hierarchical aggregation of routing areas using PCEs;	X	
31	Hierarchical PCE ordering to create multi-domain multicast trees;	X	
33	Inter domain routing model for Grid computing. PCE and Authentication, Authorization and Accounting (AAA) servers are included to form a new service layer, which then instructs the control plane to perform various tasks;		X
34	Uses a hierarchical aggregation to evaluate a loose order of domains to be traversed and then calculates actual end-to-end path; seen to perform better than ad-hoc peering model.	X	
35	Model for inter-domain routing. Two layers of hierarchy, all domain-specific PCEs are connected to one central PCE;	X	

Rank accounting methods [48], [49] have also been used in conjunction with per-domain signaling based approaches to enhance light-path setup in multi domain WDM networks. In a multi domain scenario, end-to-end wavelength availability information is not available, and a set of candidate wavelengths is sent from the source with the path setup message. Wavelengths that are not available at intermediate nodes are removed from the setup message as it progresses to the destination. While only one of the available wavelengths is selected by the destination to setup a light-path, the availability information of other wavelengths is also sent to the source in the signaling message, allowing the source to rank wavelengths according to availability. In [48], simulations were used to show the performance of the rank accounting method against random allocation. The rank accounting method proposed used only K out of W available wavelengths to route inter-domain traffic. The rank accounting algorithms proposed in [48] always showed better performance than random allotment of K wavelengths for inter domain traffic, and had lower blocking probabilities than random allotment with all W wavelengths for inter domain traffic at low loads. Also, change in performance of the rank accounting algorithm was observed against the value of K, and it was seen that the rank accounting algorithms achieved almost similar performance to random allocation of all wavelengths for K=3 with W=8, and K=8 with W=16, and therefore giving the optimal number of wavelengths to be utilized in inter-domain path setups. The rank accounting algorithms proposed in [49] also showed significant reduction in blocking probabilities when compared to random allotment methods.

Crankback signaling is also used to increase the efficiency of traditional signaling mechanisms. In crankback signaling, if a path setup request fails at a given node, the signaling process intimates the upstream node to use an alternate path if available. The performance of PCE based approaches against RSVP with crankback signaling was evaluated in [41]. It was observed that the PCE based approach consistently calculated paths with lower cost as compared to the crankback based schemes, and could admit more connections. The number of crankbacks also increased with the increase of load in the system. An extension of the crankback signaling scheme is also used in [50], where the crankback signal is used to send path parameters upstream to determine optimal multi domain traffic engineered (TE) LSP paths in the network. The crankback scheme was shown to outperform the per-domain scheme and was comparable to the multi-domain visibility scheme for lower loads. A comparison of the proposed scheme with other per domain distributed path computation schemes is shown in [50]. Simulation results showed a 40% increase in LSP formation and a mild reduction when compared to LSP formulation with multi-domain visibility.

B. Quality of Service (QoS)

QoS demands from the network can be a combination of various parameters. For Layer 1 and Layer 2 networks, the demands can be defined in terms of bandwidth, delay, jitter and optical signal quality. Common QoS demands for bandwidth and delay can be satisfied by constraint based path computation mechanisms as discussed in the previous section. However, technology and vendor specific parameters like signal attenuation, signal impairment etc. are not standardized between different domains and therefore difficult to evaluate. Also, the lack of standard measurement mechanisms in these networks implies that same QoS guarantees as defined by different carriers can be different in implementation. Therefore, end-to-end QoS guarantees are not easily achieved. It should be noted that the existence of non-standardized service classes in different networks also leads to difficulties in defining inter-domain QoS parameters [52].

Different works target optimization of different parameters to provide a feasible QoS mechanism in multiple domains. Problems related to measurement of end-to-end optical signal impairment in a multi-domain network are outlined in [47]. The author proposes a signaling based approach which records the impairment in a path in the RSVP (Resource Reservation Protocol) PATH message. A controlled flooding mechanism is proposed to discover a suitable QoS path in the multi-domain network. An alternate mechanism can be to use link state updates to distribute impairment parameters and use QoS constraint-based routing to determine optimal paths at the source. This mechanism requires large and frequent link state updates between domains due to fluctuations in impairment values. The authors use simulations to illustrate that both mechanisms show comparable blocking probabilities, and while setup time is larger in the constrained broadcast signaling mechanism, the reduced number of link state updates

when compared to the QoS routing mechanism make it an attractive alternative to set up QoS paths in WDM networks.

End-to-end delay affects the performance of many upcoming applications such as VoIP, SAN etc. Delay is typically acceded to two major categories: 1) Transmission and queuing delay, and 2) propagation delay. Carrier grade transport technologies can deliver guaranteed bandwidth and delay bounds in the network, and therefore it is possible to determine end-to-end delay by adding the edge-to-edge delays experienced in various domains. However, new routing mechanisms are required in order to determine paths with optimal end-to-end delays. In long haul fiber networks, propagation delay constitutes a significant component of the end-to-end delay. A topology aggregation mechanism proposed in [15] announces multiple edge-to-edge light-paths in a WDM domain along with the propagation delay encountered in each path. Such an update can be used to determine end-to-end delay in WDM networks.

C. Tunnel/Circuit Signaling

Bandwidth and delay constraints are extensively studied parameters for delivery of QoS. Today, most carrier grade transport technologies like MPLS-TP, PBB-TE, SONET/SDH etc. allow set-up of paths with bandwidth guarantees in the network. Different mechanisms are then used to create end-to-end guaranteed bandwidth tunnels in a multi-domain network. While path computation mechanisms can evaluate constraint based paths inside a network, a framework is required to signal QoS paths inside the network. The RSVP protocol [42] is used extensively for signaling paths inside GMPLS networks. In cases such as the per-domain path computation, signaling and path computation run simultaneously to compute and set up TE paths in the network. Three different path setup mechanisms, namely 1) Contiguous LSP, 2) Stitched LSP, and 3) Nested LSP, have been standardized for MPLS networks. [53] defines the per-domain path setup mechanism for each of these three approaches, which can be implemented using a signaling protocol such as RSVP. A combination of these approaches can also be used to set up inter-domain paths.

The contiguous LSP setup allows the source to set up an end-to-end TE LSP till the destination. While this is ideal, it is required that the border nodes of different domains participate in the LSP setup. The other mechanisms (stitched, nested) do not require the same.

In the stitched LSP setup, each domain sets up independent LSPs inside their own domains and mappings are created to merge the different LSPs at the border nodes. In this fashion, the head end node of each segment inside a domain is aware of the state of the path setup inside the domain. In the stitched TE LSP mechanism, each domain creates a new LSP for the request inside the domain and therefore has fixed capacity as that of the requested connection.

In case of nested LSP setup, the required TE LSP segment in reserved inside an already existing TE LSP by reserving bandwidth in the LSP. Existing LSPs which are used to transport TE LSPs can also be advertised as forwarding adjacencies in the MPLS domain.

Nested LSPs can also be used to implement advance reservation of capacity for inter-domain connections inside a domain. Nested TE LSP groom traffic inside a single LSP and can therefore have a flexible mechanism to admit traffic inside the nested tunnel. Nested LSPs also maintain fewer states in the network and reduce the signaling required during path setup inside the network. On the other hand, stitched LSP approaches can be used to distribute inter-domain load in the topology. Also, stitched LSP approaches allows for freedom in choice of path protection mechanisms used.

D. Multi Domain Path Protection and Restoration

Path protection is a widely studied phenomenon in carrier transport networks, with each technology having some mechanism to recover from node/link failures in the network. However, as protection mechanisms are triggered by the control plane and are specific to the technology used, inter-domain path protection is non-trivial. Path protection in a multi-domain scenario can be broken down into the following major categories:

1) End-to-end link disjoint path protection
2) Per domain link disjoint path protection
3) Multi-domain Shared Path Protection
4) Virtual path protection

1) End-to-End Link Disjoint Path Protection: Discovery of end-to-end link disjoint paths in a multi domain network is typically hindered due to the absence of full topology information across domains. Ricciato et al. [54] outline two broad categories of possible approaches, namely PCE-based and signaling approaches to determine end-to-end disjoint LSPs in a multi-domain network. In a PCE-based approach, all the PCEs in the different domains collaborate beforehand to determine the optimal end-to-end link disjoint path, thereby always giving the optimal path for the system. Two signaling based approaches are presented, with one approach forming the primary path and recording the path in a RRO (Record Route Object) in the RSVP RESV message, and the second LSP RSVP reserve request carrying the primary route details in a new object thus enabling domains to determine a link disjoint path inside the domain. The second approach includes computation of both paths at the same time. In such a mechanism, a pair of ingress and egress border nodes is selected, and two link disjoint paths are set up inside a domain towards the destination.

Note that both signaling based schemes do not guarantee an optimal path pair as the selection of the egress border nodes inside a domain does not depend on the other domains. In addition, the reservation of a path followed by discovery of a link disjoint pair may lead to scenarios where a path cannot be setup even with the existence of link disjoint pairs, as illustrated in Fig. 5. While reserving a path from s to d, if the path s-x-b-d is reserved first, then a link disjoint path cannot be reserved between s to d, even though link disjoint paths s-a-b-d and s-x-y-d exist in the network between s and d.

Sprintson et al. [17] proposed a novel topology aggregation scheme which can be used to determine end-to-end link disjoint paths. The scheme is built upon the algorithm proposed in [18] to determine the optimal link-disjoint path pair in a network. To determine the optimal path pair in the network,

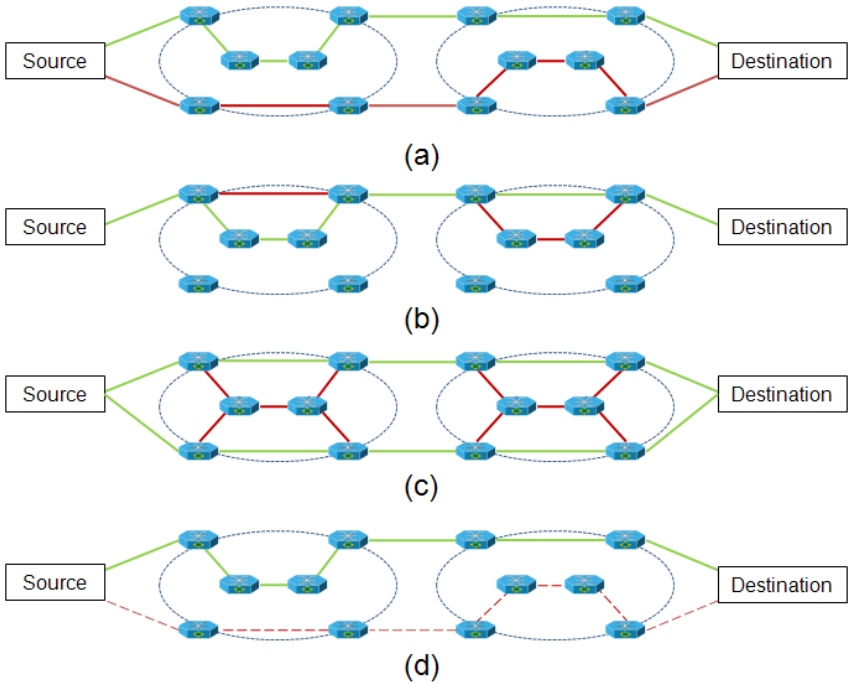


Fig. 4. (a) End-to-end link disjoint protection path (red), (b) Per domain path protection (red), (c) Shared Path Protection (red) and (d) Virtual path protection (red, dashed)

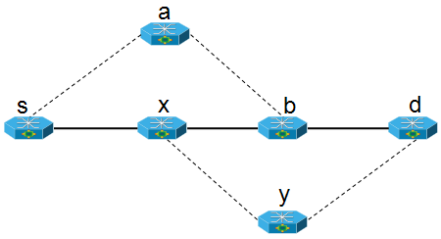


Fig. 5. Sample "Trap" Topology [54]

the scheme proposed in [17] first calculates the optimal path from source to destination, and then reverses the direction of the links used in the optimal path to determine the optimal protection path. The aggregation scheme proposed in [18] uses the full mesh aggregation to represent the optimal advertised path between two edge nodes in the domain. For each path p, a separate full mesh topology is advertised, which is evaluated by reversing the links used in p which is then used to determine disjoint path pairs. With each path in the full mesh representation requiring another full mesh representation for determination of the link disjoint pair, the size of the topology information updates of the order of $O(|M|^4)$ as compared to $O(|M|^2)$ in a full mesh scheme (M is the number of border nodes).

Another possible approach in the calculation of end-to-end disjoint paths can be the calculation of diverse AS path chains. The model proposed in [10], [11], [12] proposes a centralized service which can calculate multiple AS level paths between two domains. These mechanisms can be extended to calculate different AS chains such that working and protection path is routed in different ASs. End-to-end disjoint paths can be set up easily if all ASs traversed in the working path are not present in the protection path. However, in the case when there are no two AS chains without common AS hops, new mechanisms will be required to ensure the link disjointedness of the working and protection paths.

2) Per-Domain Path Protection: End-to-end link disjoint paths are difficult to set up, and require switching of paths at the source in case of failure. Such a mechanism also leads to higher switchover times, as a failure trigger has to traverse multiple domains to switch over at the source. Per domain protection have lower switchover times as the switchover is restricted to a single domain. Per domain path protection schemes for optical light-path setup are proposed in [13], [14], [26] which utilize the local control plane to provide per-domain path protection.

Although per domain protection paths are easier to control and require significantly lower signalling information, a failure in an inter-domain link requires that a new end-to-end path be setup. A per domain path protection mechanism is proposed in [55] for MPLS. In a scenario where two domains may be interconnected by multiple links, the choice of protection LSPs is made such that end-to-end protection can be implemented with minimal signaling.

3) Multi-Domain Shared Path Protection: While both end-to-end link-disjoint as well as per-domain path protection schemes uniquely reserve a backup path for a working path, shared path protection attempts to reduce the capacity reserved for backup paths by sharing the reserved backup capacity for different paths. Shared path protection mechanisms have been extensively studied for intra-domain path protection and are classified into two major categories 1) Static routing - where an exact demand matrix or a bounded demand matrix is assumed for shared path protection and 2) Dynamic routing - where requests are considered upon arrival with any knowledge of the traffic request matrix. As they do not assume any traffic demand matrix, dynamic routing shared path protection mechanisms are useful in the multi-domain path protection scenario.

Truong et al. [56] surveyed different shared path protection schemes in the multi-domain context. As only limited information exchange is allowed between domains, end-to-end shared path protection is not considered in these works and only segment based or overlapping segment based path protection is used for inter-domain path protection. Different protection schemes are classified in two categories 1) Multiple intra-domain Protection and 2) Hierarchical Routing with Topology Aggregation. The multiple intra-domain protection is similar to the per-domain path protection scheme, where each segment is assumed to span a domain in the inter-domain path and is protected by a shared path protection scheme. However, such schemes cannot provide protection against failure of border nodes or of inter-domain links. In order to protect inter-domain links, a virtual domain consisting of border nodes of the interconnected domains is introduced. While this mechanism can protect the inter-domain link, it does not provide protection against failure of border nodes.

Hierarchical routing with topology aggregation (HiTA) techniques use aggregated topology representation of the different domains in the network to create a simple inter-domain virtual topology. On this virtual topology, intra-domain shared path protection schemes such as P-cycles, Shared Segment Protection (SSP) and Overlapping Shared Segment Protection (OSSP) are applied. Once capacities are determined, signalling protocols are used to set up the actual working path and the backup path in the different domains. These techniques can exploit the available inter-domain information for better working and backup path selection. However, the lack of intra-domain information during the calculation of inter-domain paths can lead to setup failures.

To overcome this drawback, a novel technique known as Map and Route OSSP is proposed which proposes that only certain intra-domain paths be used to transfer inter-domain traffic between border nodes. Each intra-domain path is advertised as a virtual path between a border node, and this mapping allows for exact mapping of working and backup paths during the inter-domain path computation.

Results show that multiple intra-domain protection schemes suffer from the lack of global information while being locally optimal. On the other hand, traditional HiTA techniques can create scalable inter-domain protection mechanisms but suffer from lack of intra-domain information. The Map and Route OSSP is a possible combination of the two techniques which also gives intra-domain information in the inter-domain aggregated topology and can therefore be used to obtain an optimal trade off between the two techniques.

4) Virtual Path Protection: Virtual path protection refers to the case where a tradeoff is created between a protection scheme which reserves a protection path before a fault happens and a scheme where a fault triggers discovery and set up of the protection path. In such a scheme proposed in [33] for optical networks, a protection path with available resources is discovered but not reserved. A probe packet is sent by the source at regular intervals to ensure the availability of the protection path. If the protection path becomes unavailable due to the setup of another connection, the source node tries to discover another available protection path. In an event of a failure in the network, a new path setup request is sent on the discovered protection path. The setup request is subject to failure depending on the availability of a protection path and the frequency of the probe message. Also, the time required to set-up the backup path must be taken into consideration.

E. Control Plane Interactions

In a multi-domain scenario, domains operate on different control planes. To share control information between domains, an inter-domain control plane interaction mechanism is required. The first standardization efforts to address inter-domain issues in optical networks came from the Global Lambda Integrated Facility (GLIF, www.glif.is). Most of the last GLIF meetings of the Control Plane working group were almost entirely dedicated to the inter-domain issues and interoperability. In the area of Grid networking, significant past research, as described in [57], has been conducted to address the applicability of inter-domain control protocols and architectures which existed in the area of optical networking since late nineties. Recently, more efforts are underway to address the integration of Gigabit Ethernet (GbE) technologies within carrier networks [58]. An open standardization issue is the design of a mechanism to provision and control the network resources within the multiple Ethernet domains.

In the absence of a standardized inter-domain control plane interaction mechanism, two broad mechanisms have been proposed to achieve inter-domain control plane interactions, i.e., hierarchical and peer-to-peer (flat).

1) Hierarchical Control Plane Architecture: Hierarchical control plane architectures have been proposed to unify inter domain control planes while also supporting existing control planes without changing them. A uniform control plane is overlaid over the existing control plane, and routing and other signaling information is exchanged on this control plane which then instructs the domain-internal control plane to perform the desired actions. Variations of this control plane architecture have been introduced in [26], which proposes the use of the 2-tier control plane architecture in a multi domain optical network for grid computing, and [22] who implements a similar architecture with a Path Computation Element (PCE) embedded into the upper tier control plane. Similarly, Lehman et al. [59] propose a two-tier control plane

architecture as shown in Fig. 6. The domain controllers interact with each other as web services to utilize the existing security mechanisms in these services. The E-NNI, I-NNI and UNI are used for interactions between different domain controllers, domain controllers and the control plane, and end user with the domain controller respectively. The UNI (User Network Interface) has been developed to facilitate requests and establishment of connections dynamically between a user and a provider. The OIF [60] defines the UNI as a mechanism which defines the interfaces between the user and the provider, and gives information about the signaling mechanisms, discovery features and service offerings to the user. Similarly, the so-called Grid-UNI has been standardized within the Open Grid Forum, defined as a specialization of the UNI for scientific applications [61]. The E-NNI (External Network to Network Interface) defines standards to exchange interface, signaling, discovery and policy information between two different administrative domains and the I-NNI (Internal Network to Network Interface) defines standards for information exchange between different network domains in the same administrative domain.

A service plane based hierarchical architecture is proposed in [10]. The service plane serves as a uniform communication plane between domains, where the service plane is used to exchange topology information for routing. The service plane also includes functions for admission and policy control for each domain. However, the paper assumes the existence of a mechanism to facilitate signaling between control planes of different domains and uses this mechanism to exchange PCE based path computation signaling messages as well as RVSP messages for path setup.

2) Peer-to-peer (Flat) Control Plane Structure: A flat control plane structure consists of a single tier control plane in each domain, with translators at the domain boundary. As shown in Fig. 7, the translator at the edge of domain D1 converts outbound inter-domain signal to a pre-determined protocol decided upon by the interconnected domains, which is then translated into the protocol used in domain D3. Such a mechanism is easy in its deployment as the translator only operates at the edges and a single domain can talk to different domains using different protocols. Flat control plane mechanisms are proposed in the DRAGON [5] and CHEETAH [62] frameworks, which introduce the Virtual Label Switch Router to solve adaptation issues for non-GMPLS compliant network elements within multiple domains. Inter domain control plane messaging uses the GMPLS protocol, which are translated to the local domain's control plane protocol for inter-operability.

Recently, the GMPLS protocol suite is being enriched by new functionalities equivalent to those defined for packet-based MPLS, described in IETF drafts RFC4105 [63] and RFC4216 [64] with their key components of reachability and TE information, path computation and LSP signalling. RFC4726 [65] specifies the framework for establishing traffic engineered LSP across multiple MPLS/GMPLS domains. Path computation can take place at the ingress Label Switched Router (LSR) for an end-to-end LSP, or a per-domain LSP or at a separate path computation element (PCE) in the domain, as described in [53], [66]. LSP-signalling can be performed in form of contiguous end-to-end LSPs [67], [68], domain wise tunnelling with nested LSPs [69] or segment based LSPs [70]. The RSVP-TE protocol extensions necessary to control and select one of these mechanisms is described in [71]. The crankback signalling mechanism is also added to the MPLS/GMPLS signalling [72]. For increased reliability, different methods are proposed in [73] to set up recovery LSPs using per domain path computation.

Standardization efforts have also been made to inter-connect the ASON and the GMPLS control planes. Okamoto et al. [74] describe the challenges in interconnecting GMPLS and ASON control planes to allow automatic end-to-end path setup. They propose the use of an interworking function, which can translate the signalling at the border nodes to facilitate interworking between the GMPLS and the ASON frameworks. An interworking function is proposed for the GMPLS border nodes which can support different inter-domain connection request scenarios. A field trial of the proposed function is also presented in their work.

IV. An Example: Multi-Domain Carrier-Grade Ethernet

We now discuss the relative advantages of the different approaches and features on a hypothetical inter-domain Provider Backbone Bridging-Traffic engineering (PBB-TE) routing architecture. It is important to note that as of today there is no standard or previously published work on inter-domain provisioning within carrier-grade Ethernet domains and that the example is for illustration only. Let us assume the network which consists of multiple PBB-TE domains as shown in Fig. 8, where P stands for provider domains carrying Ethernet traffic and D stands for core domains running carrier-grade PBB-TE routers and switches. Our architecture should be capable of setting up end-to-end working and protection TE paths (tunnels) and guarantee QoS across the network. We assume that domains are not allowed to share complete topology information with each other and use either partial topology dissemination or mesh-like topology aggregation schemes. Route determination can be performed by using either distributed signaling protocols or PCE based architectures. We choose a PCE based approach over a signaling approach as the PCE approach always gives a better path as compared to a signaling based approach.

Let us assume a hierarchical PCE architecture similar to ASON networks (Fig. 8). The hierarchy implies that each tier is looking at a reduced view of the network, which leads to a simplified path computation problem. A hierarchical PCE model also implies that link state updates are sent to the PCE, and each PCE in turn sends a more compact link state update to the higher level PCE. The controlled flow of information implies that link state updates can be more frequent and can allow for announcement of multiple paths between border node pairs calculated based of different constraints such as delay, bit error rate, jitter etc. It should be noted the PCEs at the highest hierarchy peer with each other to exchange topology information. In addition, PCE approaches are also efficient in formation of end-to-end link disjoint paths as compared to other proposed signaling and topology aggregation mechanisms.

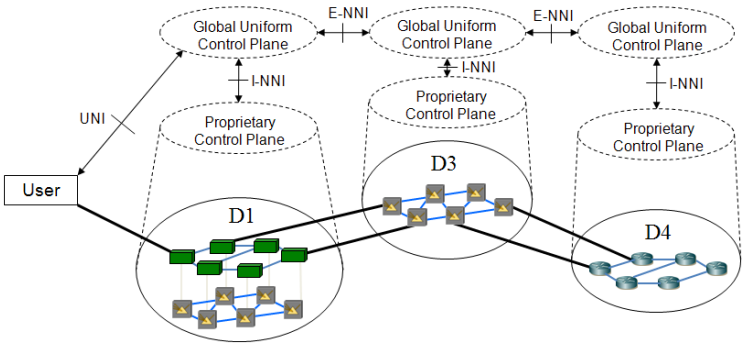


Fig. 6. Hierarchical control plane architecture as proposed in [59]

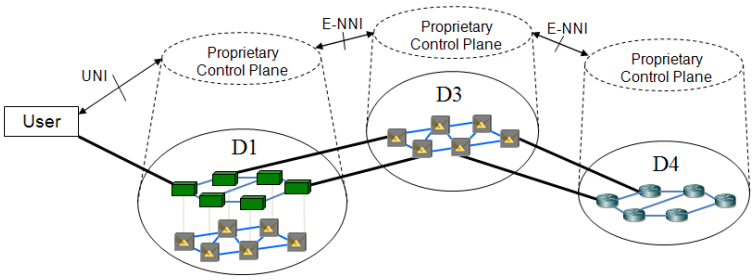


Fig. 7. A flat control plane architecture

During a path setup, all PCEs at the highest level of hierarchy talk to each other in order to determine the optimal path before setup. It should be noted that in highly dynamic network operations, high delays between the PCE signaling for path determination and signaling for path setup may lead to requests being blocked in the network.

Let us assume that an end-to-end tunnel with QoS guarantees has to be set up between edges E1 and E12. To determine the path to setup the red tunnel as shown in Fig 8, edge E1 in domain D1 queries the PCE 1.a for a route to the customer node in provider domain P4. As the destination E12 is not in the same domain, the request is forwarded to higher level of hierarchy, i.e., PCE 1. From prior information exchange, i.e., routing update which is exchanged periodically with PCE 2, PCE 1 knows that the destination is in the routing area defined by PCE 2. PCE 1 queries PCE 2 for path information to the destination from the edge routers, and then determines the optimum path. This information is then sent back to E1 which initializes the signaling process illustrated in Fig. 9. The path information does not contain hop-by-hop path details, and the domain PCE is consulted at the ingress for the detailed intra-domain path. In order to set up the tunnel, E1 first signals a PBB-TE tunnel initialization message to E3. The signaling process blocks resources at all intermediate nodes along the path. E3 forwards the path message to E4, and waits for a confirmation from E4 before replying to E1. Similarly, E4 sends the setup message to E5 which forwards it to E11 and waits for the confirmation message. E11 forwards it to E12, and E12 sends a reply to E11. Once the tunnel E11-E12 is formed, E11 sends a confirmation to E5. E5 sends the confirmed reservation response to E4, and creates a mapping between the local tunnel to the inter domain tunnel. The mapping creates a relation between the backbone destination address and the backbone VLAN identifier (B-DA, B-VID) of the local tunnel i.e. (MAC E5, VID 2) to the (B-DA, B-VID) of the inter domain tunnel i.e. (MAC E12, VID 4). A similar mapping is created at E3 by creating a mapping from (MAC E3, VID 1) in D1 to (MAC E5, VID 2) in D2. Data transfer can be started when signaling is terminated at E1.

In case the PCE cannot find a suitable path in a domain, the tunnel reservation is cancelled and a message is sent upstream to release the reserved resources. If neither confirmation nor a cancellation message is received, reserved resources are released after a preset timeout.

The example depicts a QoS path set up by creating TE paths inside each domains and then ”stitching” them together. Different domains in the network may be operating on different control planes. While PCEs interact with each other using the PCEP signaling protocol as defined in [75], signaling between different control planes is facilitated by using the E-NNI architecture to interconnect different domains at the edges to form a flat control plane structure.

Although mechanisms are in place to perform inter-domain peering and provisioning, as seen in this example, a number of challenges still remain. For instance, in a pure PBB-TE architecture, destination nodes are identified by their MAC addresses. While routing in each domain only requires information about MAC addresses in the local domain, global MAC

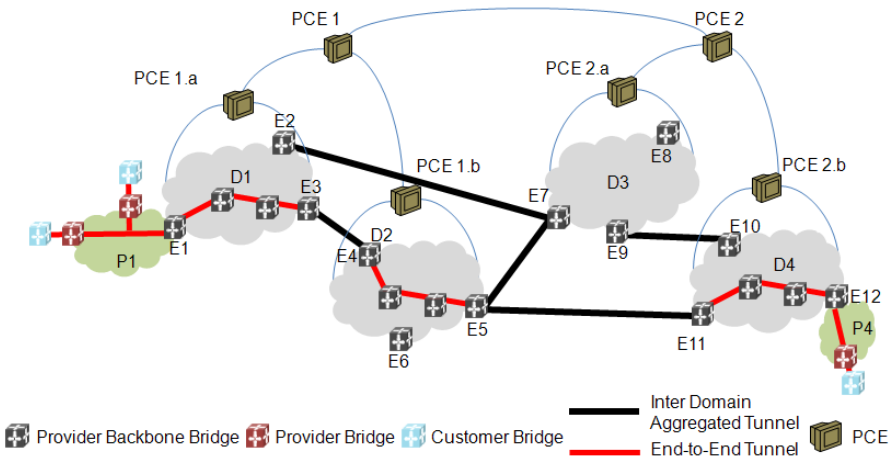


Fig. 8. An example PBB-TE network with PCE-based hierarchical routing areas

address information is required to during setup of inter-domain paths. The flat addressing scheme of Ethernet MACs will lead to large lookups in such a scenario, thereby limiting the scalability of the system. Other open issues, such as protection and QoS, are summarized in the next section.

V. Open Issues

A. Routing Scalability and Inter-domain Path Computation

The success of the BGP protocol in the Internet makes a compelling case for use of path-vector protocols in large-scale networks. The path vector protocol has the highest scalability of all the approaches discussed above. The reduced information as advertised by path vector protocols implies that the information can be spread easily and is less susceptible to fluctuations in the network. However, the path vector protocols do not address problems with multi-layer-multi-technology path determination. It should also be noted that as path-vector algorithms have large recovery times from routing information failures, which also need to be taken into consideration.

Aggregated topology schemes are not as scalable as advertised inter-domain capacity and other TE parameters may change due to fluctuations in intra-domain traffic. On the other hand, aggregated topology schemes provide a more accurate description of resources available in various domains for inter-domain traffic and are therefore better suited for QoS routing. To increase scalability of topology aggregation schemes, mechanisms must be put in place to reduce fluctuations in the advertised resources. A possible mechanism can be to reserve resources for inter-domain connections in advance, so as to ensure that available inter-domain capacity is not affected by changes in intra-domain traffic.

Distributed routing algorithms and signaling approaches, although scalable, are susceptible to inefficient resource utilization. On the other hand, PCE based approaches are known to provide better path selection than distributed routing approaches. The routing area aggregation approach proposed for ASON networks [29] allows for routing scalability, and reduces the signaling load in the network. PCE based architectures are also ideally suited for multi-layer routing problems. However, PCEs are a single point of failure in the network, thus requiring failover mechanisms.

Scalability has been recognized as an important challenge for PCE design. Future work will have to explore in detail the potential of pre-computation schemes in conjunction with path computation schemes to improve the performance in the network, both in the control plane and the data plane.

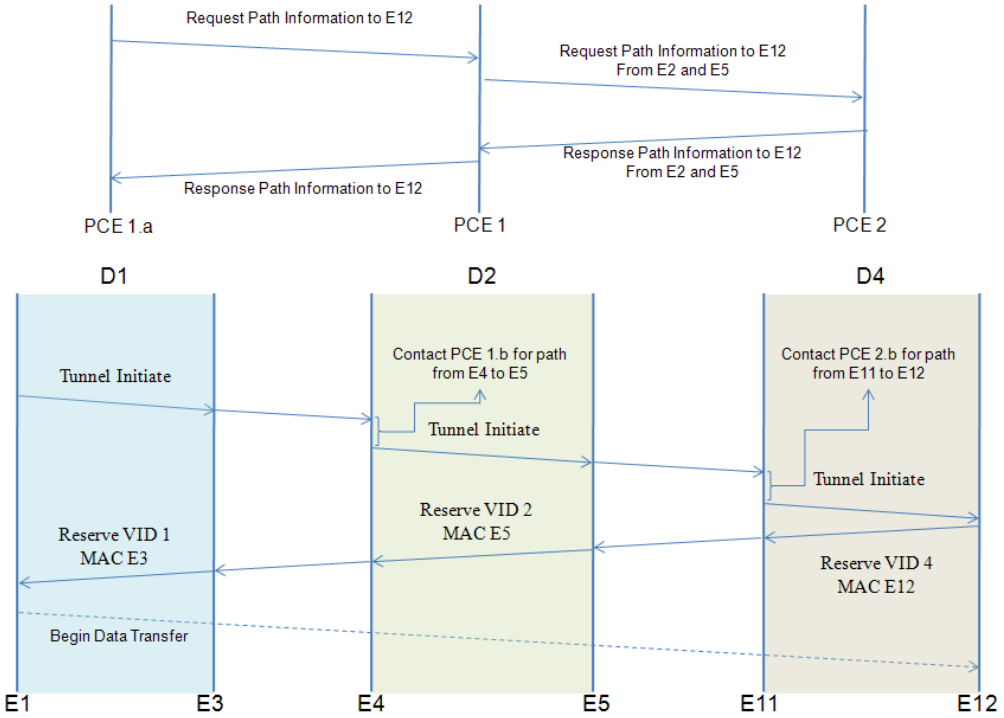
B. Reliability

We have discussed different protection mechanisms that can be deployed in an inter-domain scenario. It should be noted that these mechanisms may be deployed simultaneously in the network, with the choice of protection mechanism governed by the cost of deployment and the criticality of the connection. From the presented schemes, shared path protection and per domain dedicated path protection schemes are the most scalable and promising mechanisms for inter-domain path protection. While end-to-end link disjoint path protection is desirable, the high cost involved in discovery and setup of end-to-end link disjoint paths prove a major hindrance for the same.

Shared path protection mechanisms can be applied successfully for layer 2 networks, where data can be groomed easily into shared segments. However, for layer-1 optical networks without wavelength conversion, shared path protection schemes are not desirable as the constraint of wavelength continuity increases the amount of information exchange required between domains, thus reducing scalability. Per-domain dedicated path protection schemes may be better suited for inter-domain all optical path protection. However, issues related to signal attenuation due to shared paths has not been addressed in present works.

C. Admission Control and Billing

In a multi-domain network, it is imperative to have an admission control and billing model, which can decide if a



Tunnel Mapping <B-DA, B-VID>

Node	Incoming <B-DA, B-VID>	Outgoing <B-DA, B-VID>
E3	<MAC E3, VID 1>	<MAC E5, VID 2>
E5	<MAC E5, VID 2>	<MAC E12, VID 4>

Access Control

Node	Incoming Node	Incoming <B-DA, B-VID>	Rule
E1	Source	-	Append <MAC E3, VID 1>
E4	E3	<MAC E5, VID 2>	Allow to continue into domain
E11	E5	<MAC E12, VID 4>	Allow to continue into domain
E12	-	<MAC E12, VID 4>	Remove Tag and forward to destination

Fig. 9. End-to-end tunnel setup in example PBB-TE network

connection request is authentic and has sufficient privileges to set up a connection across the domain. While each domain typically has proprietary Authentication, Authorization and Accounting (AAA) mechanisms in place today, uniformity in the billing model will ensure smoother operations and easier accountability. A uniform AAA mechanism can possibly lead to establishment of dynamic Service Level Agreements (SLA) against static SLAs currently in place. To do so, inter-domain AAA interactions must be standardized. Hierarchical control plane proposals such as [26] propose the implementation of AAA in the global control plane layer. As against AAA in flat control planes, hierarchical AAA implementations ensure that there is uniformity in the working of AAA services, and AAA interactions between peering domains are well defined. In [76], a method is proposed for secure reservation and admission which ensures that (a) users pay for resources that are available, (b) users cannot block resources without paying for them, (c) requests to the various elements of the provider network are suitably authenticated (d) successful reservation of a resource implies that this resource is available so possession of a reservation implies admission. The key proposed innovation of this architecture is the ability to handle decentralized access control through the use of credentials.

D. Multi-domain QoS

The lack of standardization of service classes in the network require that QoS parameters be discovered during path setup. Different mechanisms have been proposed to discover certain QoS parameters in a path in a network. While some approaches [47] use a signaling based mechanism to measure parameters in the network, other approaches include QoS constraints in the SFP [47] algorithm, and broadcast QoS parameters of the advertised paths.

The first step in providing end-to-end QoS requires standardization of QoS service classes and measurement mechanisms. After standardization, discovery of QoS paths in the network will require that each domain announce the presence of these QoS paths. While this increases the number of updates in the network, a hierarchical routing area approach can be used to reduce the signaling overhead in the network.

Billing and monitoring of QoS services in multi-domain scenario poses another challenge in the context of current networks. Network Management Systems (NMS) of neighbouring domains can be interfaced to exchange operational information useful in monitoring specific inter-domain paths. However, the diversity in existing NMSs implies that this scheme may not be scalable. Billing and monitoring services require that a trust mechanism must exist between the source domain and all the intermediate domains along the inter-domain path. A possible mechanism can be to introduce an authentication mechanism where a domain can authenticate with non-neighbouring domains in the network. One such mechanism has been proposed in [81], where the source domain has a trust agreement with each remote (transit/destination) domain. Public certificates are used to sign PCEP messages to authenticate the source, while special tokens are included to validate authorization for requested path in each remote domain. The PCEP message also carries a token from each of the remote domains to the source, which are later introduced in the RSVP messages to couple authorization used in path computation with authorization for path setup. By forming a trust relationship with non-neighbouring domains, billing as well as monitoring information can then be shared securely between domains via the control plane.

E. Control Plane Interworking

Ideally, the same control plane architecture should be deployed over all domains for multi-domain networks. A single control plane will ensure seamless integration between domains and is also easier to upgrade. Interfacing between multiple domains is not the ideal solution as it can lead to complex interworking functions required to translate control plane messages between domains and will be cumbersome to manage and upgrade. The hierarchical control plane approach can simultaneously support the development of a uniform control plane while easing the transition by allowing proprietary control planes to interface with the global control plane. However, in the absence of a clear choice of a single global control plane, peering mechanisms must be developed to interface the dominant choices of control planes. In the present scenario, where a few largely accepted control plane architectures such as ASON and GMPLS are present with other proprietary control planes, hierarchical interfacing approaches must be used to deploy GMPLS or ASON as a uniform overlay, while peering approaches must be developed to perform interworking between the GMPLS and ASON control planes. This simultaneously supports the development of GMPLS and ASON as future control planes for multi-domain networks while proprietary control planes to interface easily with these architectures.

F. Dynamic Peering

Peering in inter domain networks has been facilitated by the use of physical interconnects between two domains and by exchange points. Exchange points (IX) have been widely used to interconnect multiple domains in a network with reduced cost. When multiple domains want to establish peering agreements amongst themselves, new connections must be laid to interconnect the border nodes of the two domains. This implies that a large number of back-to-back links (physical circuits) must be drawn to interconnect all domains. Exchange point architectures convert the physical mesh topology to a star topology, and virtual circuits can be formed over this physical topology to generate a logical mesh interconnection.

Exchange points have been traditionally used to create peering agreements which are static in nature. Exchange point architectures have been developed for various technologies like MPLS, Ethernet etc, which form tunnels between connected domains. However, the exchange point architecture is ideal to create dynamic inter domain peering in the layers below IP. Exchange point architectures have been proposed for MPLS [77] and GMPLS [76] networks which utilize the ability of these networks to set up LSPs which can then be used as forwarding adjacencies. MPLS/GMPLS indicates a freedom in the choice of transmission technology used by the providers connecting to the exchange points, and also removes proximity constraints as posed by older Ethernet exchange points. Next generation commercial exchange point models have been proposed [76], [78], [79], [80] which allow trading of network resources between interconnected parties. Such a framework can be extended to form dynamic service level agreements (SLA) for short sporadic durations, and pre-reservation of resources can effectively lead to the formation of a commodities market situation.

VI. CONCLUSIONS

In this paper, we presented a survey of issues and challenges of inter-domain peering and provisioning of use in L2 and WDM based on optical transport networks. Specifically, we identified issues of inter-domain topology dissemination, routing, delivery of QoS, signaling for connection setup, protection and control plane interactions as key factors. A review of the existing literature was presented to give an understanding of the problems existing with present technologies and present challenges in inter-domain provisioning. To understand the challenges arising in the evolving technologies, such as carrier-grade Ethernet, we presented a hypothetical architecture for multi-domain PBB-TE networks and showed that the proper mix of features and technologies is necessary when forming an inter-domain model. Finally, we highlighted open issues pertaining to scalability, reliability, multi-domain QoS, control plane interactions and dynamic peering which need to be addressed in the future. Our main findings are that scalable mechanisms are required to perform routing in multi-domain networks and that pre-computation schemes need to be explored in conjunction with path computation schemes such as the PCE to improve the performance in the network. We highlighted the problems arising due to the lack of standardized QoS metrics which needs to be addressed in the future. To

Acronym	Abbreviation
AAA	Authentication, Authorization and Accounting
AS	Autonomous Systems
ASON	Automatically Switched Optical Networks
ATM	Asynchronous Transfer Mode
B-DA	Backbone Destination Address
BGP	Border Gateway Protocol
BRPC	Backward Recursive Path Computation
B-VID	Backbone Virtual Local Area Network identifier
COPRP	Constraint-based Optical Path-vector Routing Protocol
DiffServ	Differentiated Service
E-NNI	External Network to Network Interface
GbE	Gigabit Ethernet
GELS	GMPLS controlled Ethernet Label Switching
GLIF	Global Lambda Integrated Facility
GMPLS	Generalized Multi Protocol Label Switching
HiTA	Hierarchical Routing with Topology Aggregation
IDRA	Inter Domain Routing Agent
IETF	Internet Engineering Task Force
I-NNI	Internal Network to Network Interface
IP	Internet Protocol
IPTV	Internet Protocol Television
ITU-T	International Telecommunication Union Telecommunication Standardization Sector
IX	Exchange Point
L2	Layer 2
LSP	Label Switched Path
LSR	Label Switched Router
MAC	Media Access Control
MPLS	Multi Protocol Label Switching
MPLS-TP	Multi Protocol Label Switching Transport Profile
NMS	Network Management Systems
OBGP	Optical Border Gateway Protocol
OBS	Optical Burst Switching
OIF	Optical Internetworking Forum
OSSP	Overlapping Shared Segment Protection
PBB-TE	Provider Backbone Bridging Traffic Engineering
PCE	Path Computation Element
PNNI	Private Network to Network Interface
QoS	Quality of Service
RRO	Route Record Object
RSVP	Resource ReSerVation Protocol
RWA	Routing and Wavelength Assignment
SAN	Storage Area Network
SDH	Synchronous Digital Hierarchy
SLA	Service Level Agreement
SONET	Synchronous Optical Network
SPF	Shortest Path First
SSP	Shared Segment Protection
TE	Traffic Engineering
UNI	User Network Interface
VID	Virtual Local Area Network Identifier
VLAN	Virtual Local Area Networks
VPN	Virtual Private Network
WDM	Wavelength Division Multiplexing

speed up deployment in existing networks, we point to the fact that control plane interactions need to be standardized and a uniformly accepted control plane framework should be identified.

VII. Acknowledgements

This work has been supported by the German Federal Ministry of Education and Research (BMBF) under support code 01BP0771; Eureka-100GET. We thank the anonymous reviewers for their insightful comments and suggestions.

Appendix

See Acronym list above.

References

[1] IEEE802.1Qay, *Provider Backbone Bridge Traffic Engineering*, http://www.ieee802.org/1/pages/802.1ay.html

[2] *Global Lambda Integrated Facility: Control Plane Working Group*, http://www.glif.is/working-groups/controlplane/

[3] *CA*net 4 Research Program Update - UCLP Roadmap for creating User Controlled and Architected Networks using Service Oriented Architecture*, www.canarie.ca/canet4/uclp/UCLP_Roadmap.doc

[4] *Interdomain Reservation and Authorization on OSCARS/DCN Version 0.3*, www.internet2.edu/dcresearch/DCN-reserv.arch.v5-1.pdf

[5] T. Lehman, J. Sobieski, B. Jabbari, "*DRAGON: A Framework for Service Provisioning in Heterogeneous Grid Networks*," IEEE Communications Magazine, pp. 84-90, March 2006

[6] D. Fedyk, H. Shah, N. Bitar, A. Takacs, "*GMPLS control of Ethernet PBB-TE*", IETF Internet Draft, http://www.ietf.org/internet-drafts/draft-ietf-ccamp-gmpls-ethernet-pbb-te-01.txt, July 2008

[7] *The North American Network Operators' Group*, www.nanog.org

[8] G. Shrimali, A. Akella, A. Mutapcic, "*Cooperative Inter-Domain Traffic Engineering Using Nash Bargaining and Decomposition*," IEEE INFOCOM 2007.

[9] A. Tomaszewski, M. Pioro, M. Mycek, "*Distributed Inter-Domain Link Capacity Optimization for Inter-Domain IP/MPLS Routing*," IEEE GLOBECOM 2007.

[10] R. Douville, J.L. Le Roux, J.L. Rougier, S. Secci, "*A Service Plane over the PCE Architecture for Automatic Multidomain Connection-oriented Services*," IEEE Communications Magazine, vol. 46 pp. 94-102, June 2008

[11] S. Secci, J.L. Rougier, A. Pattavina, "*AS Tree Selection for Inter-Domain Multipoint MPLS Tunnels*," IEEE ICC 2008, pp.5863-5868

[12] S. Secci, J.L. Rougier, A. Pattavina, "*On the Selection of Optimal Diverse AS-paths for Inter-domain IP/(G)MPLS Tunnel Provisioning*," Telecommunication Networking Workshop on QoS in Multiservice IP Networks, 2008. IT-NEWS 2008.

[13] Q. Liu, M. A. Kok, N. Ghani, V. M. Muthalaly, M. Wang, "*Hierarchical Inter-Domain Routing in Optical DWDM Networks*," IEEE INFOCOM 2006, pp. 1 - 5

[14] Q. Liu, M. A. Kok, N. Ghani, V. M. Muthalaly, M. Wang, "*OPN08-02: Inter-Domain Provisioning in DWDM Networks*," IEEE GLOBECOM 2006, pp. 1 - 6

[15] S. Sanchez-Lopez, X. Masip-Bruin, E. Marn-Tordera, J. Sol-Pareta, J. Domingo-Pascual "*A Hierarchical Routing Approach for GMPLS based Control Plane for ASON*," IEEE ICC 2005

[16] P. Wan, W. Jiao, X. Wu, "*A Novel Topology Aggregation Method for Hierarchical Routing in ASON Network*," IEEE GLOBECOM 2007, pp. 2275-2279

[17] A. Sprintson, M. Yannuzzi, A. Orda, X. Masip-Bruin, "*Reliable Routing with QoS Guarantees for Multi-Domain IP/MPLS Networks*," IEEE INFOCOM 2007, pp.1820 - 1828

[18] J. Suurballe, "*Disjoint Paths in a Network*," Networks, 4:125-145, 1974

[19] B.S. Arnaud, R. Hatem, W. Hong, M. Blanchet, F. Parent, "*Optical BGP Networks*", www.canarie.ca/canet4/library/c4design/opticalbgpnetworks.pdf

[20] M. J. Francisco, L. Pezoulas,C. Huang, I. Lambadaris, "*End-To-End Signaling and Routing for Optical IP Networks*," IEEE International Conference on Communications 2002, vol. 5, pp. 2870-2875

[21] M. Blanchet, F. Parent, B. St-Arnaud, "*Optical BGP (OBGP): InterAS lightpath provisioning*," IETF Internet Draft, http://tools.ietf.org/id/draft-parent-obgp-01.txt, March 2001

[22] O. Yu, "*Intercarrier Interdomain Control Plane for Global Optical Networks*," IEEE ICC 2004, vol. 3, pp. 1679 - 1683

[23] M. Yannuzzi, X. Masip-Bruin, G. Fabrego, S. Sanchez-Lopez, A. Sprintson, A. Orda, "*Toward a New Route Control Model for Multidomain Optical Networks*," IEEE Commun. Mag., vol. 46 pp. 104-111, June 2008

[24] M. Yannuzzi, S. Sanchez-Lopez, X. Masip-Bruin, J. Sole-Pareta Jordi-Domingo-Pascua, "*A Combined Intra-Domain and Inter-Domain QoS Routing Model for Optical Networks*," Optical Network Design and Modeling 2005.

[25] T. Saad, H. T. Mouftah, "*Inter-domain wavelength routing in optical WDM networks*,"Telecommunications Network Strategy and Planning Symposium. 11th International IEEE NETWORKS 2004, June 2004, pp. 391- 396

[26] L. Gommans, F. Dijkstra, C. de Laat, A. Taal, A. Wan, B. van Oudenaarde, T. Lavian, I. Monga, and F. Travostino, "*Applications Drive Secure Lightpath Creation across Heterogeneous Domains*," IEEE Commun. Mag., vol. 44, no. 3, pp. 100-106, Mar. 2006

[27] I. Baldine, P. Mehrotra, G. Rouskas, A. Bragg, D. Stevenson, "*An Intra and Inter-Domain Routing Architecture for Optical Burst Switched (OBS) Networks*," International Conference on Broadband Networks 2005, vol. 2, pp. 1073 - 1082

[28] T. Okumus, H. Junseok, H. A. Mantar, S. J. Chaplin, "*Inter-domain LSP setup using Bandwidth Management Points*," IEEE GLOBECOM 2001, pp. 7 - 11

[29] S. Sanchez-Lopez, J. Sole-Pareta, J. Comellas, J. Soldatos, G. Kylafas, M. Jaeger, "*PNNI-based Control Plane for Automatically Switched Optical Networks*," IEEE J. Lightwave Technol., vol. 21, pp. 2673 - 2682, Nov. 2003

[30] G.7715/Y.1706, "*Architecture and Requirements for Routing in ASON*", ITU-T Recommendation Draft

[31] H. Matsuura, N. Morita, I. Nakajima, K. Takami, "*Hierarchically Distributed PCE for Flexible Multicast Traffic Engineering*," IEEE GLOBECOM '07, pp. 2439-2444

[32] H. Matsuura, T. Murakami, K. Takami, "*An Interdomain Path Computation Server for GMPLS Networks*," IEEE Trans. Commun., vol.E88-B, no. 9, pp.3329-3342, Aug 2005

[33] O. T. Yu, T. A. DeFanti, "*Collaborative User-Centric Lambda-Grid over Wavelength-Routed Network*," ACM/IEEE Conference on Supercomputing 2004

[34] P. Torab, B. Jabbari, Q. Xu, S. Gong, "*On Cooperative Inter-Domain Path Computation*," IEEE ISCC 2006, pp. 511 - 518

[35] D. Dimitrijevic, B. Maglaris, R. R. Boorstyn, "*Routing in Multidomain Networks*," IEEE/ACM Transactions on Networking, vol. 2, no. 3, pp. 252-262, 1994

[36] JP. Vasseur, JL. Le Roux, "*Path Computation Element (PCE) Communication Protocol (PCEP)*," IETF Internet Draft, http://www.ietf.org/internet-drafts/draft-ietf-pce-pcep-13.txt, Aug. 2008

[37] JP. Vasseur, R. Zhang, N. Bitar, JL. Le Roux, "*A Backward Recursive PCE-based Computation (BRPC) Procedure To Compute Shortest Constrained Inter-domain Traffic Engineering Label Switched Paths*," IETF Internet Draft, http://tools.ietf.org/id/draft-ietf-pce-brpc-09.txt, April 2008

[38] JP. Vasseur, A. Farrel, "*Preserving Topology Confidentiality in Inter-Domain Path Computation Using a Key-Based Mechanism*," IETF Internet Draft, http://www.ietf.org/internet-drafts/draft-ietf-pce-path-key-03.txt, May 2008

[39] E. Oki, "*PCC-PCE Communication and PCE Discovery Requirements for Inter-Layer Traffic Engineering*," IETF Internet Draft, http://www.ietf.org/internet-drafts/draft-ietf-pce-inter-layer-req-07.txt, April 2008

[40] E. Oki, T. Takeda, J-L Le Roux, A. Farrel, "*Framework for PCE-Based Inter-Layer MPLS and GMPLS Traffic Engineering*," IETF Internet Draft, http://tools.ietf.org/id/draft-ietf-pce-inter-layer-frwk-07.txt, June 2008

[41] S. Dasgupta, J. C. De Oliveira, J. P. Vasseur, "*Path Computation Element Based Architecture for Interdomain MPLS/GMPLS Traffic Engineering: Overview and Performance*," IEEE Network, pp.38-45, July 2007

[42] R. Braden, L. Zhang, S. Berson, S. Herzog, S. Jamin, "*Resource ReSerVation Protocol (RSVP)*," IETF RFC 2205, http://tools.ietf.org/rfc/rfc2205.txt, Sept 1997

[43] J. Zhou, X. Yuan, "*A Study of Dynamic Routing and Wavelength Assignment with Imprecise Network State Information*," IEEE ICPP Workshop on Optical Networks 2002

[44] X.Masip-Bruin, R.Munoz, S.Sanchez-Lopez, J.Sole-Pareta, J.Domingo-Pascual, G. Junyent, "*An Adaptive Routing Mechanism for reducing the Routing Inaccuracy Effects in an ASON*," ONDM 2003, pp. 333-349

[45] X.Masip, S.Sanchez-Lopez, J.Sole, J.Domingo, D.Colle, "*Routing and Wavelength Assignment under Inaccurate Routing Information in Networks with Sparse and Limited Wavelength Conversion*," IEEE GLOBECOM 2003

[46] X.Masip-Bruin, S.Sanchez-Lopez, J.Sole-Pareta, J.Domingo-Pascual, "*QoS Routing Algorithms under Inaccurate Routing Information for Bandwidth Constrained Applications*," IEEE ICC 2003

[47] E. Salvadori, Y. Ye, A. Zanardi, H. Woesner, M. Carcagni, G. Galimberti, G. Martinelli, A. Tanzi, D. La Fauci, "*Signalling-Based Architectures for Impairment-Aware Lightpath Set-Up in GMPLS Networks*," IEEE GLOBECOM 2007, pp. 2263-2268

[48] T. Tachibana, H. Harai, "*End-to-End Lightpath Establishment Based on Rank Accounting in Multi-Domain WDM Networks*," IEICE TRANS. COMMUN., vol. E89-B, no.9 sept. 2006, pp. 2448-2457

[49] T. Tachibana, H. Harai, "*Aggressive Rank Accounting for End-to-End Lightpath Establishment in Multi-Domain WDM Networks*," IEEE ICC 2006, vol. 6, pp. 2459 - 2464

[50] F. Aslam, Z. A. Uzmi, A. Farrel, M. Pioro, "*Inter-Domain Path Computation using Improved Crankback Signaling in Label Switched Networks*," IEEE ICC 2007, pp. 2023 - 2029

[51] A. Orda, A. Sprintson, "*Precomputation Schemes for QoS Routing*," IEEE/ACM Transactions on Networking, vol. 11, no. 4, pp. 578-591, Aug. 2003.

[52] P. Jacob, B. Davie, "*Technical Challenges in the Delivery of Interprovider QoS*," IEEE Communications Magazine, vol. 43, issue. 6, pp. 112- 118, 2005

[53] J.P. Vasseur, A. Ayyangar, R. Zhang, "*A Per-domain path computation method for establishing Inter-domain Traffic Engineering (TE) Label Switched Paths (LSPs)*," IETF RFC 5152, http://www.rfc-editor.org/rfc/rfc5152.txt, Feb. 2008

[54] F. Ricciato, U. Monaco, D. Ali, "*Distributed Schemes for Diverse Path Computation in Multi-domain MPLS Networks*," IEEE Communications Magazine, Feature Topic on Challenges in Enabling Inter-Provider Service Quality on the Internet, 43(6):138-146, June 2005.

[55] Huang, D. Messier, "*A Fast and Scalable Inter-Domain MPLS Protection Mechanism*," Journal of Communications and Networks, vol. 6, no. 1, Mar. 2004 pp. 1-8

[56] D.L. Truong, B. Jaumard, "*Recent Progress in Dynamic Routing for Shared Protection in Multidomain Networks*," IEEE Communications Magazine, vol. 46 pp. 112-119, June 2008

[57] A. Jukan, G. Karmous-Edwards, "Optical Control Plane for the Grid Community," IEEE Surveys and Tutorials, accepted for publication in 2007.

[58] D. Papadimitriou, N. Sprecher, J. Cho, L. Andersson, D. Fedyk, D.Allan, P. Busschbach, A. Takcs, T. Eriksson, D. Caviglia, "*A Framework for GMPLS-controlled Ethernet Label Switching*," IETF Internet Draft, http://quimby.gnus.org/internet-drafts/draft-dimitri-gels-framework-00.txt, Feb. 2006

[59] T. Lehman, Xi Yang, C.P. Guok, N.S.V. Rao, A. Lake, J. Vollbrecht, N. Ghani, "*Control Plane Architecture and Design Considerations for Multi-Service, Multi-Layer, Multi-Domain Hybrid Networks*," IEEE High-Speed Networks Workshop 2007, pp. 67-71

[60] *Optical Internetworking Forum*, www.oiforum.com

[61] *Open Grid Forum*, www.ogf.org

[62] Q. Song, Z. Li, I. Habib, "*Cheetah VLSR Design and Deployment in GMPLS Optical Networks*," IEEE GLOBECOM 2006, pp. 1-6

[63] J-L. Le Roux, J-P. Vasseur, J. Boyle, "*Requirements for Inter-Area MPLS Traffic Engineering*," IETF RFC4015, http://www.ietf.org/rfc/rfc4105.txt, June 2005

[64] R. Zhang, J.-P. Vasseur, "*Requirements MPLS Inter-Autonomous System (AS) Traffic Engineering (TE) Requirements*," IETF RFC 4216, http://www.ietf.org/rfc/rfc4216.txt, Nov 2005

[65] A. Farrel, J-P. Vasseur, A. Ayyangar, "*A Framework for Inter-Domain Multiprotocol Label Switching Traffic Engineering*," IETF RFC 4726, http://www.ietf.org/rfc/rfc4726.txt, Nov 2006

[66] A. Farrel, J-P. Vasseur, "*A Path Computation Element (PCE)-Based Architecture*," IETF RFC4655, http://www.ietf.org/rfc/rfc4655.txt, Aug. 2006

[67] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, G. Swallow, "*RSVP-TE: Extensions to RSVP for LSP Tunnels*," IETF RFC 3209, http://www.ietf.org/rfc/rfc3209.txt, Dec. 2001

[68] L. Berger, "*Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions*," IETF RFC 3473, http://www.ietf.org/rfc/rfc3473.txt, Jan 2003

[69] K. Kompella, Y. Rekhter, "*Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic Engineering (TE)*," IETF RFC 4206, http://www.ietf.org/rfc/rfc4206.txt, Oct 2005.

[70] A. Ayyangar, K. Kompella, J-P. Vasseur, A. Farrel, "*Label Switched Path Stitching with GMPLS TE*", IETF Internet Draft, http://tools.ietf.org/wg/ccamp/draft-ietf-ccamp-lsp-stitching/draft-ietf-ccamp-lsp-stitching-06.txt, April 2007

[71] A. Farrel, A. Ayyangar, J-P. Vasseur, "*Inter domain MPLS and GMPLS Traffic Engineering - RSVP-TE extensions*," IETF RFC 5151, http://www.rfc-editor.org/rfc/rfc5151.txt, Feb. 2008

[72] A. Farrel, Satyanarayana, A. Iwata, N. Fujita, G. Ash, "*Crankback Signaling Extensions for MPLS and GMPLS RSVP-TE*," IETF RFC 4920, http://www.rfc-editor.org/rfc/rfc4920.txt, July 2007

[73] T. Takeda, A. Farrel, Y. Ikejiri, JP. Vasseur, "*Analysis of Inter-domain Label Switched Path Recovery*," IETF RFC 5298, http://tools.ietf.org/html/rfc5298, Aug. 2008

[74] S. Okamoto, H. Otsuki, T. Otani, "*Multi-ASON and GMPLS Network Domain Interworking Challenges*," IEEE Communications Magazine, vol. 46 pp. 88-93, June 2008

[75] J. Ash, J-L. Le Roux, "*Path Computation Element (PCE) Communication Protocol Generic Requirements*," IETF RFC4657, http://www.ietf.org/rfc/rfc4657.txt, Sept. 2006

[76] S. Tomic, A. Jukan, "*GMPLS-based Exchange Points: Architecture and Functionality*," Proceedings of the Workshop on High Performance Switching and Routing 2003 (HPSR 2003), Torino, Italy, June 2003.

[77] I. Nakagawa, H. Esaki, K. Nagami, "*A design of Next Generation IX using MPLS Technology*," IEEE Symposium on Applications and the Internet, 2002.

[78] V. Prevelakis, A. Jukan, "*How to buy a Network: Trading of Resources in the Physical Layer*," IEEE Commun. Mag., Dec. 2006.
[79] F. Dijkstra, C. de Laat, "*Optical Exchanges*," Proc. of GRIDNETS 2004, Oct 2004.
[80] S. Tomic, A. Jukan, "*MPFI: The Multi-provider Network Federation Interface for Interconnected Optical Networks*," IEEE GLOBECOM 2002, Nov. 2002, Taipei, Taiwan.
[81] S. G. Polito, M. Chamania, A. Jukan, "Extending the Inter-domain PCE Framework for Authentication and Authorization in GMPLS Networks," accepted in IEEE ICC 2009.

Mohit Chamania is currently pursuing his PhD at the Technical University Carolo-Wilhelmina of Braunschweig, Germany. Prior to his PhD studies, he pursued his B.Tech and M.Tech in Mechanical Engineering at the Indian Institute of Technology, Bombay. His research interests include: inter-domain routing, Ethernet and Optical technologies.

Admela Jukan is W3 Professor of Electrical and Computer Engineering at the Technical University Carolo-Wilhelmina of Brunswick (Braunschweig) in Germany. Prior to coming to Brunswick, she was research faculty at the Institut National de la Recherche Scientifique (INRS), University of Illinois at Urbana Champaign (UIUC) and Georgia Tech (GaTech). From 2002-2004, she served as Program Director in Computer and Networks System Research at the National Science Foundation (NSF) in Arlington, VA. She received the M.Sc. degree in Information Technologies and Computer Science from the Polytechnic of Milan, Italy, and the Ph.D. degree (cum laude) in Electrical and Computer Engineering from the Vienna University of Technology (TU Wien) in Austria. Dr. Jukan is the author of numerous papers in the field of networking, and she has authored and edited several books. She serves as a member of the External Advisory Board of the EU Network of Excellence BONE. Dr. Jukan has chaired and co-chaired several international conferences, including IFIP ONDM, IEEE ICC and IEEE GLOBECOM. She is a Senior Member of the IEEE.