GMPLS-Based Dynamic Provisioning and Traffic Engineering of High-Capacity Ethernet Circuits in Hybrid Optical/Packet Networks*

Xi Yang[1], Chris Tracy[2], Jerry Sobieski[2], Tom Lehman[1]

[1]Information Sciences Institute East, University of Southern California
3811 Fairfax Drive, Suite 200, Arlington, VA 22203; Email: {xyang, tlehman}@isi.edu

[2] Mid-Atlantic Crossroads, University of Maryland
8400 Baltimore Ave, Suite 102, College Park, MD 20740; Email: {chris, jerrys}@maxgigapop.net

Abstract - **Rapid progress in deployment of national and regional optical network infrastructures holds the promise to provide abundant, inexpensive bandwidth to scientific communities. The expectation is that this will provide a catalyst for innovation across a range of eScience applications. However, there remains considerable architecture and research work in order to allow such applications to access a common shared high-speed infrastructure efficiently and flexibly. This paper presents an experimental solution currently being designed and tested over the Dynamic Resource Allocation via GMPLS Optical Networks (DRAGON) and the Internet2 Hybrid Optical and Packet Infrastructure (HOPI). This implementation provides deterministic network services to high-end eScience applications in the form of Ethernet "circuits" using GMPLS to dynamically provision user specific VLAN based configurations. This requires addressing several research issues associated with multi-layer networks, interdomain routing, signaling, and multi-constraint path computation. The larger vision is that of a shared infrastructure consisting of multiple network technology layers over which an intelligent control plane can provide for a unified service model. The instantiation described here of Ethernet over WDM provides an ideal platform for both experimental research and, real services, and a migration path to a production infrastructure.**

I. INTRODUCTION

Hybrid optical and packet network architectures are of increasing interest because of their potential to provide a more inexpensive and flexible network infrastructure. The vision is for a common transport infrastructure which can be utilized to provide the base connectivity for a traditional router based IP network in addition to enabling advanced services for high end users or for further traffic engineering of the IP network. These advanced services can be provided by enabling direct access to network services at various network layer/technologies such as Ethernet and SONET. Here we address a mechanism to dynamically provide advanced services across the Ethernet layer available on the DRAGON[1] and HOPI[2] networks. The HOPI architecture includes high capacity Ethernet switches which are placed on top of a static WDM based optical infrastructure. The DRAGON infrastructure includes an Ethernet layer on top of a dynamically switched all-photonic core. The advantages of service provisioning at the Ethernet layer includes inexpensive and high capacity connectivity based on a well understood technology. However, it faces several challenges in design of the network control plane. Firstly, demands of eScience applications are very flexible in bandwidth amount, time and topology. This requires dynamic provisioning of almost arbitrary bandwidth granularity over a wide range of Ethernet switch vendor types and across multiple administrative domains. This in turn requires dynamic resource advertising, routing, and signaling protocols. Secondly, while all circuits share a common network infrastructure in a dynamic and efficient manner, each individual circuit must have reserved, dedicated bandwidth. A flexible resource reservation mechanism is needed to pack Ethernet circuits into wavelength channels while guaranteeing that each circuit gets its desired amount of bandwidth. Thirdly, we need a User Network Interface (UNI) for service-level presentation of network resources to facilitate user-friendly access. For example, a cluster application may request for a GigE circuit with multiple given access ports at ingress and egress switches. Such requirements should be captured and supported by the control plane.

II. SOLUTION APPROACH

Our solution of dynamic provisioning and traffic engineering of Ethernet circuits is based on the Generalized Multi-Protocol Label Switching (GMPLS) protocol suite. Resource advertising, routing and signaling components are extended from the GMPLS standards to support end-to-end Ethernet circuits. A key component is the Virtual Label Switch Router (VLSR), a control plane entity that speaks GMPLS protocols, including OSPF-TE[3] and RSVP-TE[4], on behalf of the underlying non-GMPLS-capable Ethernet switches. OSPF-TE is our choice of resource state

*This work is supported in part by National Science Foundation under Grants Numbers 0335300 and 0335266.

advertising protocol. In order to dynamically manage the Ethernet resources, we advertise the aggregate available bandwidth on each wavelength-link together with the set of available Ethernet VLAN tags via OSPF-TE. Provisioning actions are instantiated via use of RSVP-TE signaling in order to set up Label Switched Paths (LSP) with the requested bandwidth and a proper VLAN tag. At each Ethernet switch, VLSR translates RSVP-TE signaling messages into local switch commands and creates the desired VLAN-ports associations along with the requested bandwidth guarantees. Whenever an Ethernet circuit (or LSP) is set up or torn down, the bandwidth and VLAN tag information is updated via distribution of OSPF-TE Link State Advertisements (LSAs) in order to maintain proper link states across the network. Figure 1 depicts VLSR configuration.

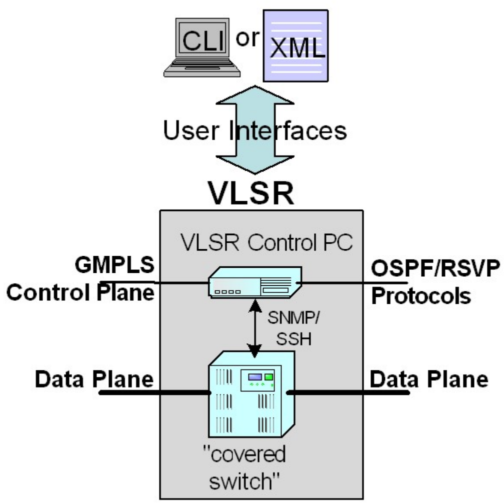

Figure 1 VLSR Architecture

The information advertised by VLSR OSPF-TE is collected into a traffic engineering database by another entity called Network-Aware Resource Broker (NARB). The NARB is an entity which represents a domain and runs as a listener to the intradomain routing protocol. The NARB includes a Resource Computation Engine (RCE) which utilizes the traffic engineering database information to perform constrained based path computation. When an Ethernet circuit is requested, the bandwidth and VLAN tag parameters in the traffic engineering database are used to constrain the end-to-end path computation. The path computation is performed by a RCE that employs a two-hierarchy recursive path computation scheme to conduct sophisticated constraint-based routing algorithms over a multi-layer, multi-domain traffic engineering link-state topology. Upon successful path computation for a requested Ethernet circuit, and Explicit Route Object (ERO) is provided to the RSVP-TE signaling engine to initiate provisioning. Another function of the NARB is to perform inter-domain traffic engineering routing exchanges. There is one NARB per domain which maintains a peering with NARBs in adjacent domains. These NARBs exchange topology information in order to allow interdomain path computation and signaling. This topology exchange can be based on actual topologies or "abstracted" topologies as defined by administrative configuration. Figure 2 shows the dynamic ethernet provisioning environment enabled by operation of the DRAGON control plane across the HOPI and DRAGON infrastructures. As depicted DRAGON is a multi-layer environment, while HOPI is a single ethernet layer from a provisioning perspective. Users are provided with three options to initiate a provisioning request

- Peer to Peer GMPLS - in this mode users run a full suite of the GMPLS protocols and have a control and data plane connection to an edge VLSR. This can be accomplished on an intra or inter domain basis.
- UNI Interface - in this mode the user only needs to run RSVP in a "UNI" style mode.
- Web Service - in this mode the user uses an XML format to request services from the appropriate control plane element via an application interface.

Edge control flexibility is provided by use of a "Local ID". This feature allows a user to specify the type and number of egress ports which are mapped to the ends of a LSP. There are four types of Local IDs defined which allow a LSP to be terminated in user defined and flexible combinations of one port, multiple ports, tagged ports, or untagged ports. The Local ID configurations can be different at either end of the LSP. In this manner an end to end LSP can be used in multiple application specific configurations such as connecting a single host on one side of the LSP to a cluster of computers on the opposite LSP end.

II. OPEN RESEARCH ISSUES

Multi-layer Networks - As noted earlier, the DRAGON network is a multi-layer GMPLS network. This means that the WDM optical layer and the Ethernet layer can be provisioned dynamically via the GMPLS protocols. This provides for much flexibility as well as several complexities for which solutions are still in the research phase such as multi-layer path computation and signaling. Our general approach to date has been based on hierarchical provisioning of LSPs. In this manner the optical layer is provisioned first and a new traffic engineering link (of higher layer switching capability) is then advertised back into the link state topology. This is subsequently utilized for new dynamic provisioning actions. Future plans include a contiguous multi-layer provisioning action where a single RSVP signaling session is utilized to instantiate an end to end LSP. Another method for future investigations includes the stitching together of separate LSPs to form an new end to end "LSP". Provisioning across multi-layer, multi-region network topologies presents many research and real world challenges. Research issues include development of

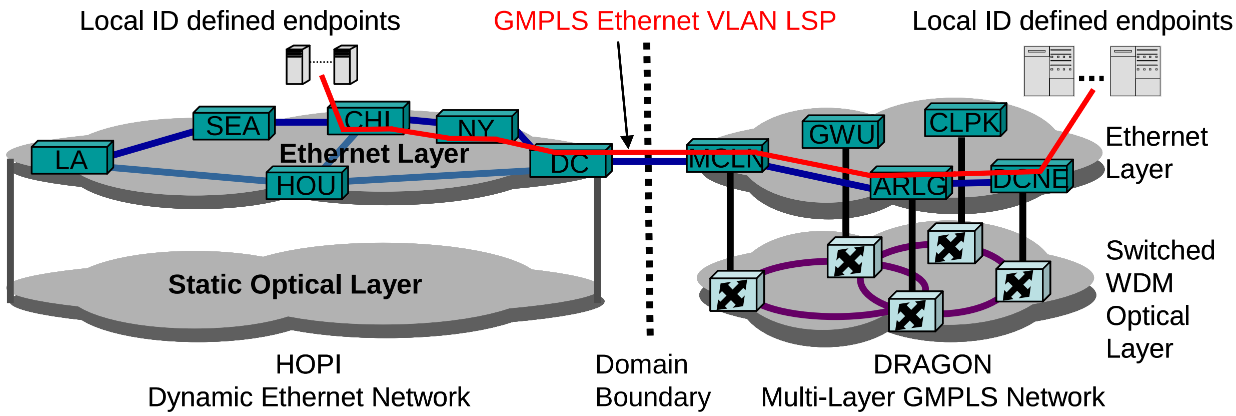

Figure 2 Dynamic Ethernet Provisioning Environment

algorithms to compute paths across heterogeneous technology multi-layer networks. Real world challenges include dealing with vendor specific implementations, limitations, or partial standards compliance.

Advanced Ethernet Services - The ethernet service provisioning capability described in this paper allows for a VLAN based "circuit" to be provisioned between two end points across the network. The Local ID concept does provide a degree of flexibility for the end points in terms of number of connection points and options for tagged versus untagged VLAN ports. However, when using tagged VLANs, there are limitations associated with the global nature of the VLAN tag space. There are several research areas underway to address this limitation. One is based on ethernet VLAN tag/label swapping technique known as GMPLS-controlled Ethernet Label Switching (GELS) which has been discussed in IETF[5]. Another method is to use a function like Ethernet QinQ to enable VLAN tunneling. Incorporation of these types of features into a heterogeneous interdomain control plane involves several areas of research and complication beyond just the capability being available on an individual switch. This is an active area of research for us, and we expect to integrate some of these advanced ethernet services as features available via the GMPLS control plane. The dynamic application of Quality of Service (QoS) parameters as part of the provisioning process is another area where value could be added these Ethernet services. This would include going beyond the current application of basic bandwidth guarantees that we have currently implemented. Use of features like priority queues to provided latency or jitter control could be of interest for some services. Other topics for future investigation include true GMPLS point-to-multipoint and multipoint-to-multipoint provisioned ethernet services.

Hybrid Network Environments - An important part of the overall vision is that of a hybrid network environment where end points can rapidly decide and seamlessly move between the routed IP service and the more deterministic dedicated path service. There is much work required to determine procedures and develop mechanisms to allow end systems to operate in this environment. This includes providing tools and APIs to interact with the circuit network, resource discovery mechanisms to determine peer system availability, and local system adaptations to accommodate any end system specific addressing or configuration requirements.

VI. SUMMARY

We present in this paper the design and implementation details of an experimental control plane architecture for dynamic provisioning and traffic engineering of Ethernet circuits in hybrid optical and packet networks. Our field tests have demonstrated that using this control plane, end users can easily provision end-to-end Ethernet circuits of arbitrary bandwidth up to 10 Gbps over a multi-vendor, multi-domain, common shared optical and Ethernet network infrastructure. These capabilities are being used to provide services to eScience users. A broader use for this capability is envisioned in the future including additional applications and traffic engineering for general purpose IP networks.

REFERENCES

[1] DRAGON Project, “Dynamic Resource Allocation via GMPLS Optical Networks,” http://dragon.maxgigapop.net.
[2] HOPI Project, "The Hybrid Optical and Packet Infrastrucure Project", http://networks.internet2.edu/hopi.
[3] D. Katz, D. Yeung and K. Kompella, “Traffic engineering extensions to OSPF version 3,” IETF Internet Draft, Work in Progress, draft-ietf-ospf-ospfv3-traffic-05.txt, March 2005.
[4] L. Berger (ed), “Generalized MPLS Signaling - RSVP-TE Extensions,” IETF RFC 3473, January 2003.
[5] D. Papadimitriou et al., "Framework for GMPLS-controlled Ethernet Label Switching", draft-dimitri-gels-framework-00.txt, February 2006