The First Application-driven Lambda-on-Demand Field Trial over a US Nationwide Network

Yukio Tsukishima, Akira Hirano, Naohide Nagatsu, Takuya Ohara, Wataru Imajuku, Masahiko Jinno, Yoshihiro Takigawa, Kazuo Hagimoto
NTT Network Innovation Laboratories, NTT Corporation, 239-0847 Yokosuka, Japan
Phone: +81-468-59-8594, Fax: +81-468-59-5541, E-mail: tsukishima.yukio@lab.ntt.co.jp
Luc Renambot, Byungil Jeong, Jason Leigh, Tom DeFanti, Alan Verlo
Electronic Visualization Laboratory, University of Illinois at Chicago M/C152, 851 South Morgan Street, Chicago, IL60607 USA
Phone: +1-312-996-3002, fax: +1-312-412-7585, E-mail: luc@evl.uic.edu
Linda Winkler
Argonne National Laboratory, 9700 South Cass Avenue, Argonne, IL60439 USA
Phone: +1-630-252-7236, fax: +1-630-252-5986, E-mail: lwinkler@mcs.anl.gov

Abstract: The *Lambda-on-demand* functionality with link aggregation to accommodate the dynamic bandwidth demands of an ultra-high-resolution visualization application was realized in over a US nationwide photonic network for the first time.
©2006 Optical Society of America
OCIS codes: (060.4250) Networks; (999.9999) GMPLS

1. Introduction

A photonic network comprised of Photonic Cross Connects (PXC) with Generalized Multi-Protocol Label Switching (GMPLS) capability is a promising candidate for the design of next-generation high-performance networks. Extensive efforts [1, 2] are well underway to validate such networks. These reports mainly focus on issues such as multi-vendor interoperability and implement such photonic networks with equipment that fully supports the GMPLS protocol. However, little is done to actually evaluate their performance using real-time bandwidth-intensive applications- which is a key area of focus in the work described in this paper. The National Science Foundation's OptIPuter project [3] has been pursuing the development of data-intensive LambdaGrid applications that utilize the networks as wide-area systems busses connecting clusters of storage, computing and visualization services. The Scalable Adaptive Graphic Environment (SAGE) developed by the Electronic Visualization Laboratory (EVL), the OptIPuter's visualization middleware, can easily scale up to Tbps rendering capacity and is therefore a prime candidate for aggressively evaluating high-performance networks. In such an application, and in LambdaGrid applications in general, dynamic capacity-on-demand without explicit messaging between the applications and the photonic control plane network will highly desirable, since applications will be able to focus on application-specific processing rather than configuring network topologies and characteristics.

In this paper, we present the results of the dynamic lambda-on-demand functionality against the SAGE visualization application over a US nationwide testbed. We used two 10GbE links from Chicago to San Diego. Three PXC systems with GMPLS capability were installed in Chicago and San Diego to form a triangular topology. In addition, we developed control servers to monitor the application traffic on the first lambda and to invoke the automatic activation of the second lambda. The control server with the PXC control-plane (C-plane) automatically configured the number of lambdas and L2 connections to transparently meet the dynamic traffic changes in SAGE.

2. Network configuration and SAGE middleware system

The geographical network configuration for this field trial is shown in Figure 1. It consists of two 10GbE links over the US. One of link traverses the National Lambda Rail (NLR) [4] via the StarLight [5] facility in Chicago, through to Seattle, Los Angeles, and eventually San Diego. 10GbE LAN-PHYs were directly transported on DWDM links by Cisco 15808 and ONS15454 gear. Another 10GbE was mapped on an OC-192 channel with the help of Multi-Protocol Label Switching (MPLS) on a Juniper T-640. The 10GbE LAN-PHY extends from Chicago, through Los Angeles, to San Diego.

Figure 2 describes the experimental setup in more detail. We have developed three PXC prototypes equipped with an 8 x 8 Planar Lightwave Circuit (PLC) optical switch (see Figure 3). It supports various kinds of interfaces including both a Network Node Interface (NNI) and User-Network Interface (UNI); and has a supervisory and control unit for management and signaling communication with GMPLS capability within the C-plane. The interface cards can accommodate both 10GbE LAN-PHY and OC-192 (10GbE WAN-PHY). We installed three PXC systems, two in Chicago and one in Calit2/UCSD. Each PXC is equipped with Network Element Manager (NE-Mgr) which signals other NE-Mgrs via GMPLS to establish/delete a lambda. Each of the NE-Mgrs control PXC switches by extended Generic Switch Management Protocol(GSMP). L2 switches in each site are equipped with control servers which control their respective NE-Mgr via GMPLS/OUNI, and L2 switch via the Simple Network Management Protocol (SNMP) over Telnet, and a remote control server via a

proprietary protocol. L2 switches accept two GbE traffic from/to SAGE Transmitter (Tx) / Receiver (Rx). These NE-Mgrs and control servers are connected with each other on a routed control-plane (C-plane) network which is independent from 10GbE data-plane (D-plane) network.

SAGE [7] is a middleware system for managing visualization and high-definition video streams for viewing on ultra-high-resolution displays such as EVL's 100 Megapixel LambdaVision tiled LCD display wall. SAGE consists of Free Space Manager, SAGE Application Interface Library (SAIL), SAGE Receiver, and User Interface (UI client) as shown in Fig. 4. The Free Space Manager gets user commands from UI clients and controls pixel streams between SAIL and the SAGE Receivers. SAIL captures output pixels from applications, and streams them to appropriate SAGE Receivers. A SAGE Receiver can retrieve multiple pixel streams from different applications, and displays streamed pixels on multiple tiles. Remote visualization applications (such as 3D rendering, remote desktop, video streams, and large 2D maps) stream their rendered pixels (or graphics primitives) to SAGE, allowing for any given layout onto the displays (e.g. the output of arbitrary M by N pixel rendering cluster nodes can be streamed to X by Y pixel display screens).

3. Lambda-on-demand with GMPLS, SNMP, and Telnet

A schematic diagram for lambda-on-demand with GMPLS, SNMP, and Telnet is shown in Figure 5. Function#1 is to periodically monitor the amount of application traffic on the first lambda by SNMP polling between ingress control server and ingress L2/L3 system. Function#2 is to decide when to add or delete the second lambda, or to gather the status via the ingress control server. Function#3 is to check the status of network resources, such as that of the L2/L3 systems' interfaces and lambdas, by ingress and egress control servers. The cycle of Function#1-#3 is repeated regularly independent from Function#4-#8. Function#4-#8 is started when the decision to add or delete the second lambda is issued. Function#4 is to set up the second lambda via GMPLS. Function#5 is to check if the establishment of the second lambda was successful. Function#6 will be cancelled if the second lambda is disabled. Function#6 is to activate the link on the second lambda and to bundle links on first and second lambdas via link aggregation (IEEE802.3ad). Function#7 is to release the link on the second lambda from bundling status and to deactivate the link. Function#8 is to delete the second lambda. The parameters in this field trial are as follows. The second lambda is designed to be engaged if a series of 2 traffic measurements, monitored for four seconds on the first lambda, exceeded 800Mbps. The second lambda was disengaged if the same measurements yielded bandwidth below 350Mbps. A round-robin algorithm is used to distribute packets into the links.

4. Dynamic lambda on demand with L2 harmonization

Figure 6 presents an example of our packet loss records as a function of elapsed time observed during our field trial. As an initial condition, we have one 10GbE connection UP between SAGE Tx and Rx. Then we started a small streaming application on SAGE which pushed traffic below threshold as shown in the picture. Then we launched two additional streaming applications which required enough total bandwidth to invoke the activation of another lambda. To avoid undesirable oscillation of additional lambda activation, we set a decision guard time of eight seconds (as described earlier). During setup, L1 (lambda) reconfiguration preceded L2 reconfiguration to avoid smooth transition in combination with L2 reconfiguration. As a result, the additional lambda yielded complete elimination of packet loss as shown in Figure 6. We confirmed excellent streaming during most of the test time. In the transition when the new lambda was setup or torn down, we observed slight packet loss increase but it was recovered within a few seconds. All the L1/L2 reconfiguration was completed in a few seconds in our field trial. The change of traffic volume in these applications controlled the number of lambdas automatically without requiring any explicit messaging between the application and the C-plane of photonic network.

5. Conclusion

Dynamic lambda-on-demand functionality with harmonized re-configuration of corresponding L2 connections with link aggregation was achieved for the first time, with the OptIPuter SAGE visualizing applications over a US nationwide network testbed. Two 10GbE lambdas were automatically configured to provide end-to-end L2 connections by the newly developed control server that successfully negotiated L1/L2 connections. We confirmed excellent streaming for both configurations; single and two lambdas configuration, fulfilling the capacity demand of SAGE applications without any messaging between the application and C-plane of photonic network.

6. References

[1] S. Okamoto et al., *Tech. Digest of OFC2005*, paper PDP40 (2005).
[2] T. Otani et al., *Tech. Digest of ECOC2003*, paper Tu3.4.3 (2003).
[3] OPTIPuter http://www.optiputer.net/index.html
[4] National Lambda Rail http://www.nlr.net/
[5] StarLight http://www.startap.net/starlight/
[6] T. Watanabe et al., *Tech. Digest of ECOC2004*, Th3.6.3, (2005).
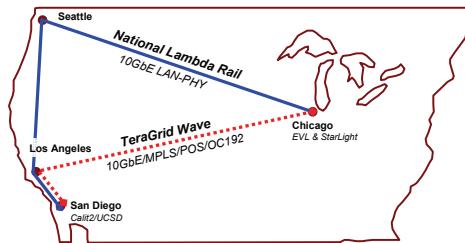[7] L. Renambot et al., *Proceedings of WACE 2004*, Nice, France (2004).
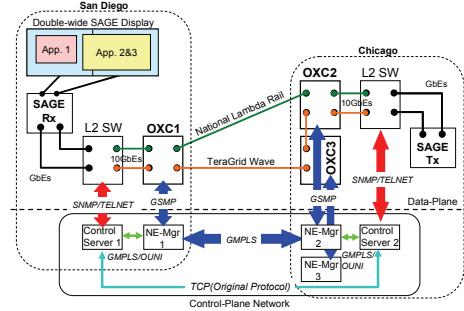
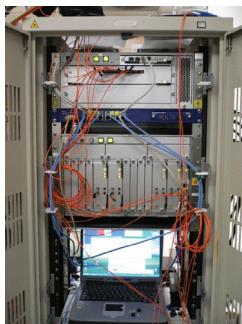Fig.1. Network Configuration

Fig.2. Experimental Setup

Optical Switch Unit

10GbE I/Fs & control units

Console

Fig.3. **OXC system with GMPLS control**

SAGE 55-tiled display system in EVL

SAIL : Sage Application Interface Library

Fig.4. **SAGE
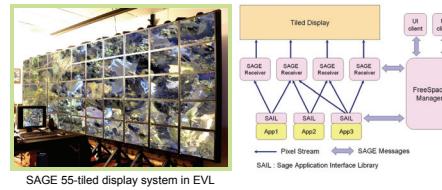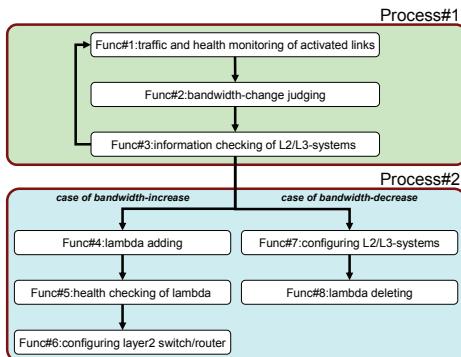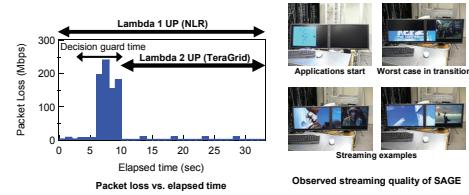(The Scalable Adaptive Graphics Environment)**

Fig.5. **Diagram for Lambda on demand function**

Fig.6. **Packet loss and streaming quality of SAGE**