12th Optoelectronics and Communications Conference/16th International Conference on Integrated Optics and Optical Fiber Communication (OECC/IOOC2007)
Technical Digest, July 2007, Pacifico Yokohama

12A2-5

The First Optically-Virtual-Concatenated Lambdas Over Multiple Domains in Chicago Metro Area Network Achieved Through Interworking of Network Resource Managers

Yukio Tsukishima[1], Akira Hirano[1], Atsushi Taniguchi[1], Wataru Imajuku[1], Masahiko Jinno[1], Yoshinori Hibino[1], Yoshihiro Takigawa[1], Kazuo Hagimoto[1], Xi Wang[2], Luc Renambot[2], Byungil Jeong[2], Ratko Jagodic[2], Sungwon Nam[2], Jason Leigh[2], Tom DeFanti[2], Alan Verlo[2]
[1]NTT Network Innovation Laboratories, NTT Corporation, 1-1 Hikarinooka 239-0847 Japan
Tel: +81-46-859-8594, Fax: +81-46-859-5541, Email: tsukishima.yukio@lab.ntt.co.jp
[2]Electronic Visualization Laboratory, University of Illinois at Chicago USA, E-mail: xiwang@uic.edu

Abstract
Optically virtual concatenated parallel lambdas over multiple-domains that are provided using reservation-messaging between network resource managers for a high-end parallel visualizing application are experimentally shown for the first time in the Chicago metro area network.

Introduction

High-end visualizing applications [1] are targeting parallelism using tiled displays and cluster computers to ensure unlimited scalability. Inevitably there will be a need for load-balanced parallel lambdas to carry their traffic. Then, these lambdas possibly traverse multiple network domains as needed. On the other hand, such applications are quite sensitive to latency deviations among these lambdas. Therefore, we have proposed Optical Virtual Concatenation (OVC) [2,3] to de-skew these lambdas. There have been no reports of the demonstration of load-balanced & OVC-enabled parallel lambdas over multiple-domains.

A network resource manager (NRM) with the capability to accept reservation requests of lambdas, to compute available routes, and to provide lambdas over a photonic network domain comprising Photonic Cross Connects (PXCs) in response to reservation requests from those applications will be a key device in the lambdaGrid network [4]. It has the capability to dynamically configure parallel lambdas in collaboration with other network devices.

In this paper, we present the first field trial of the interworking of NRMs and OVC over two photonic network domains against advanced reservation requests for network resources from a visualization middleware, which is the Scalable Adaptive Graphic Environment (SAGE) [1] developed by the Electronic Visualization Laboratory (EVL). Two PXCs equipped with OVC interfaces [2,3] were installed in the Chicago metro area network. We jointly defined and implemented a SOAP-compatible web service interface (WSI) for the NRMs interworking and NRM-SAGE interworking to exchange network resource information and reservation-relevant messages to achieve the OVC capability. In the field trial, the SAGE application sent an advanced reservation request for two parallel GbE lambdas with OVC over multiple domains. For the first time the successful interworking of two different types of NRMs achieved OVC between two lambdas that yielded complete recovery from the deterioration in the rendering performance resulting from relative latency deviation between the two GbE lambdas.

Messaging among SAGE, NRMs and OVC

Figure 1 illustrates the experimental configuration. Domain #1 and Domain #2 over I-WIRE are multi-vendor PXC networks. GbE link #1 (GbE #1) traverses through Layer 2-switch #1 (L2SW #1), PXC #1, PXC #2, and Layer 2-switch #2 (L2SW #2). In contrast, GbE link #2 (GbE #2) runs through L2SW #1, PXC #3, PXC #4, and L2SW #2. The two GbE links are bundled via the IEEE 802.3ad Link Aggregation Protocol at L2SW #1 and L2SW #2. Domain #1 is controlled by NRM #1, and Domain #2 is under the control of NRM #2. SAGE [1] comprising SAGE-Tx and SAGE-Rx systems manages visualization and high-definition video streams, enables viewing of images on ultra-high-resolution displays such as the EVL 55-tiled LambdaVision display wall, and functions as an interface between SAGE and the photonic networks. In this experiment, the visualizing processes for rendering frames among SAGE-Tx and SAGE-Rx systems are synchronized, and they require two synchronized parallel GbE lambdas..

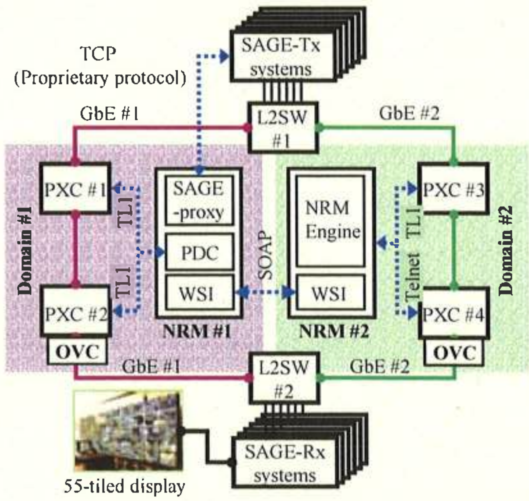

Figure 1. Experimental configuration

We developed two different NRM prototypes. NRM #1 developed by EVL consists of a SAGE-proxy, a Photonic Domain Controller (PDC) and a WSI. It receives end-to-end connection requests generated by SAGE, converts them into (multi-)domain lambda reservation requests, and sends these requests to the PDC. The PDC provides and manages lambdas in its local domain (Domain #1). It also collaborates with other NRMs through the WSI for inter-domain lambda provisioning. For local domain requests, the PDC finds available resources from the database, registers them in the reservation table, and configures local PXCs via TL1 at the scheduled time. For a foreign domain (in this case, Domain #2) requests, it converts the requests into the appropriate format and sends them to the destination NRM (in this case, NRM #2). On the other hand, NRM #2 developed by NTT consists of the NRM-Engine and a WSI. The NRM-Engine is invoked by a reservation request received from the PDC through the WSI. The NRM-Engine computes an available route according to the reservation request. If it finds an appropriate route, it registers the calculated route in the database and returns the results to the PDC through the WSI. NRM #2 periodically checks the characteristics of each of the allocated GbE links kept in the database, and controls PXCs via TL1 or Telnet when the time comes to activate or deactivate the GbE link. Moreover, NRM #1 and NRM #2 are equipped with the WSI to exchange information regarding network topology and reserved lambdas. Based on WSI, NRMs can compute the latency to be added to the OVC capable interfaces.

The reservation message diagram is shown in Fig. 2. Before starting the SAGE streaming application, SAGE sends a reservation request with parameters such as the destination and source site IDs to NRM #1. NRM #1 computes the characteristics of the GbE links that the streaming application needs. Then NRM #1 calculates the appropriate routes and checks the availability. If NRM #1 decides that the GbE link over Domain #2 is needed as well, as in the above example, it sends one more reservation request to NRM #2. NRM #2 executes calculations for the available route and lambda allocation processes in response to the reservation request. At the same time, the NRMs compute the latency for de-skewing based on the exchanged network topology information and set the latency for each OVC capable interface when the start time arrives. When the route satisfying the requirement of the streaming application is found and available, NRM #2 returns an Acknowledgement (Ack) message to NRM #1. NRM #1 and NRM #2 then configure the PXCs and activate the GbE links before the start time arrives. Finally, NRM #1 sends an Ack message to the SAGE application when it confirms that the GbE links have been activated. Only at that time, the SAGE application starts streaming.

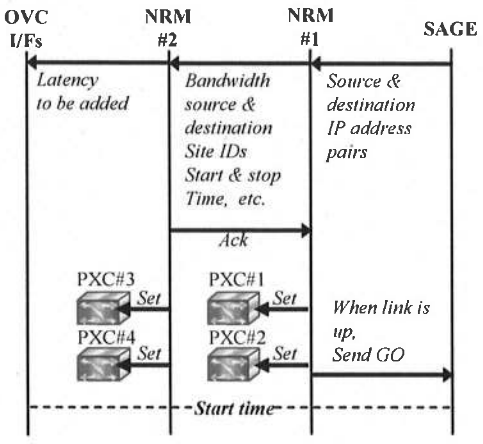

Figure 2. Messaging Diagram

Parallel visualization over OVC-enabled lambdas

In this experiment, we measured the SAGE rendering capacity against the relative latency between the GbE links.

When we initiated a TCP-based streaming application on SAGE, SAGE sent reservation requests to NRM #1 with the node-IDs (IP addresses) of the SAGE-Tx and SAGE-Rx systems as parameters. At that time, NRM #1 estimated that the required transmission capacity between SAGE-Tx and SAGE-Rx systems for rendering the streaming application on SAGE would be approximately 1.6 Gbps. Then, NRM #1 determined that two GbE connections, GbE #1 and GbE #2, would be required. NRM #1 allocated GbE #1 and sent a reservation request to NRM #2 to reserve GbE #2. NRM #2 allocated GbE #2 as requested. Following the above allocation process, NRM #1 and NRM #2 configured the PXCs. After NRM #1 confirmed that GbE #1 and GbE #2 were established, NRM #1 sent an Ack message to SAGE. Then, SAGE started the streaming application, and the streaming application was successfully rendered on the six-tiled display. When the latencies of GbE #1 and GbE #2 were zero, the rendering capacity was approximately 1.6 Gbps. During the experiments, the rendering capacity decreased with the increase in the round trip time between the SAGE-Tx and SAGE-Rx systems, and with the slowdown of the synchronous processing between the SAGE-Tx and SAGE-Rx systems caused by the increased round trip time.

To demonstrate the OVC functionality, we intentionally introduced additional latency into GbE #1. The rendering capacity of the SAGE-Rx system decreased with the relative latency between GbE #1 and GbE #2. A part of a frame comprising packets streamed on GbE #2 reached the SAGE-Rx system prior to the other part of the frame comprising packets streamed on GbE #1. To render a frame, the SAGE-Rx system buffered the part of the frame comprising the packets streamed on GbE #1 until the other part of the frame comprising the packets streamed on GbE #2 arrived at the SAGE-Rx system. As a result of the waiting process, the rendering performance of the SAGE-Rx systems decreased to approximately 480 Mbps at the relative latency of 10 milliseconds, which is anticipated when we reserve multiple-lambdas over different domains. In contrast, when the relative latency was adjusted to zero by OVC through the collaboration of NRM #1 and NRM #2, the rendering capacity recovered to 1.2 Gbps, which was an expected value with a 10-ms latency for both GbEs.

Conclusion

The SAGE parallel visualization application streamed smoothly on parallel lambdas over multiple domains, configured by OVC and two interworked NRMs. OVC realized by the NRM coordination successfully adjusted the relative latency within a millisecond. The experimental results showed the complete recovery of the SAGE rendering performance deterioration induced by the relative latency deviation.

References

1 L. Renambot *et al.*, Proceedings of WACE 2004
2 M. Tomizawa *et al.*, Tech. Digest of OFC2005 OThG6
3 A. Hirano *et al.*, FGCS, vol. 22, Issue 8, pp. 876-883, 2006
4 A. Takefusa *et al.*, OGF19, 2007
5 StarLight http://www.startap.net/starlight/
6 I-WIRE http://www.iwire.org/