A Software-Defined Optical Gateway for Converged Inter/Intra Data Center Networks

Payman Samadi, Hang Guan, Ke Wen, Keren Bergman
Department of Electrical Engineering, Columbia University, 500 West 120th Street, New York, New York 10027
E-mail: ps2805@columbia.edu

Abstract—We present a software-defined optical gateway for converged inter/intra data center networks. The optical gateway enables transparent rack-to-rack connectivity across data centers through WDM channels. Sub-second reconfiguration time and wavelength defragmentation are experimentally demonstrated.

I. Introduction:

The continuous growth in data generation and cloud-based services has increased the deployment of mid-sized data centers. Such increase has escaladed the traffic on the inter data center transport networks that are generally over-provisioned [1, 2]. Furthermore, the intra data center networks have many underutilized links, especially at the core and edge tiers [3]. One solution would be a converged inter/intra data center architecture that can partially connect data centers on-demand. This solution will improve the link utilization of both inter and intra data center networks and provides scalability in distance and performance reliability of applications.

Recently, researchers have proposed Optical Circuit Switching (OCS) in a hybrid architecture to offload elephant flows from the electronic packet switched networks of data centers [4, 5]. However, complexity in recognition of these flows leads to underutilization of the OCS network. Moreover, the outbound (north-south) traffic in data center networks is routed from the core switches by fixed number of WDM channels. Thus, in the light outbound traffic intervals, the wavelength capacity of the transport network is wasted. In [6], we proposed a converged architecture to enable rack-to-rack connectivity across data centers to improve the optical link utilization of both inter and intra data center networks, and also reliability and scalability of data centers. The design potentially provides an on-demand optical circuit between remote racks by assigning WDM channels. To implement this architecture, an optical gateway is required to add and drop WDM channels among several hundreds of racks of different data centers bi-directionally. Fast reconfiguration time and software programmability are also required to integrate with the Software-Defined Networking (SDN) data center control planes.

In this work, we present a software-defined optical gateway that is the enabler for the converged inter/intra data center architecture. The optical gateway connects racks/pods across data centers on-demand. It consists of

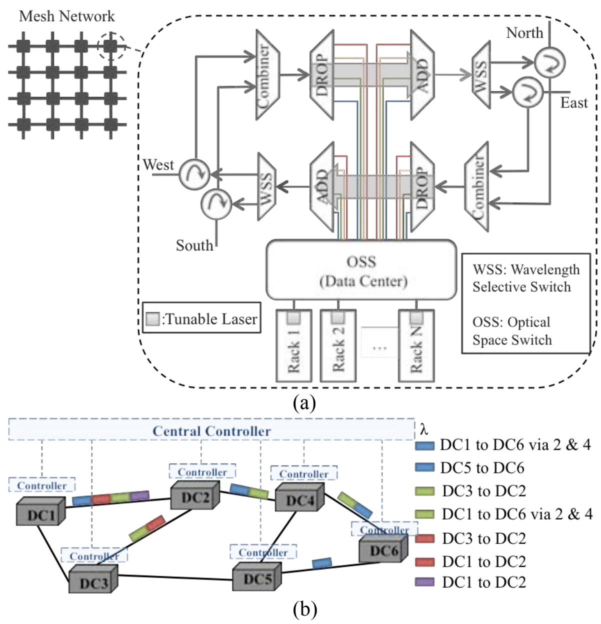

Figure 1: (a) 2×2 bidirectional software-defined optical gateway in a mesh network. The OSS provides the switching substrate for both inter and intra connectivity, (b) An example of wavelength assignment in the transport network across data centers by the distributed controller.

a hardware that connects Top-of-Rack (ToR) or aggregate switches of the data center network and a software that integrates with the data center control plane. The hardware architecture consists of an Optical Space Switch (OSS), WDM de/multiplexers and 1×2 Wavelength Selective Switches (WSS). The routing strategy in the transport network is the shortest path and wavelengths are assigned from a pool of available wavelengths on-demand. Experimental results demonstrate add/remove of new connections and wavelength defragmentation in sub-seconds.

II. Architecture:

Figure 1(a) demonstrates the architecture of the optical gateway in a mesh network for the sake of generality. The architecture consists of an OSS as a switching substrate for the intra data center network at either the edge or aggregation tiers, two WDM multiplexers and two de-multiplexers to add and drop WDM channels bi-directionally, and two 1×2 WSS to provide connectivity to adjacent optical gateways in the transport network. The WDM channel power

978-1-4799-8179-3/15/$31.00 ©2015 IEEE

adjustments are performed by the WSS based on the transport network Erbium Doped Fiber Amplifiers' (EDFA) gain profile. The optical transmitters in the data center are tunable lasers.

Figure 1(b) shows an inter data center network equipped with the optical gateway at each node. Distributed SDN control planes − while managing their own intra data center network − send and receive requests for the inter data center connectivity from/to a central controller of the transport network. The central controller finds an optimal combination of routing and wavelength assignment for connection requests using Integer Linear Programming (ILP) as shown below:

Let F_{ij}^{sdw} be a binary variable and have value 1 when wavelength w of link ij is used for traffic from the source s to the destination d. The optimization goal is to minimize the highest indexed wavelength (w_{max}) used to accommodate a connection request from $\hat{s}$ to $\hat{d}$.

$$\min \; w_{max}$$

$$\text{subject to} \quad w_{max} \geq w \, F_{ij}^{\hat{s}\hat{d}w}, \; \forall i, j, w$$

$$\sum_i F_{ij}^{\hat{s}\hat{d}w} - \sum_k F_{jk}^{\hat{s}\hat{d}w} = \begin{cases} -1, & j=\hat{s} \\ 1, & j=\hat{d} \\ 0, & otherwise \end{cases}$$

$$F_{ij}^{\hat{s}\hat{d}w} = 0 \text{ or } 1, \; \forall i, j, w$$

$$\sum_s \sum_d F_{ij}^{sdw} \leq 1, \; \forall i, j, w$$

III. Experimental Evaluations:

We built the optical gateway using a Calient S320 MEMS optical switch as the OSS, two Nistica 1×2 WSSs and four 6-channel WDM de/multiplexers (100 GHz ITU-Grid DWDM C36-C42). The control plane software is developed in python and in-house APIs are used to control the OSS and WSS, and factory-built API for the tunable laser. In order to improve the measurement accuracy, in all experiments the new connections were added and removed in 500 ms intervals. In the measurements, we captured the optical signal on an oscilloscope by an optical to electrical conversion using a photo-detector.

In the first experiment, we evaluate the switching time of the optical gateway to provide a new connection. Four WDM channels (C36-C39) enter from the north and exit to the south (see Figure 1(a)) while C40 is added. Figure 2(a) demonstrates the switching time that is 359 ms on average. It consists of the OSS and WSS switching delays, and the control plane software execution overheads.

In the second experiment, we evaluate the performance of the optical gateway for wavelength defragmentation. WDM Channels C36, C38 and C39 enter from north and exit to the south (see Figure 1(a)). The spectrum at the south port of the optical gateway is demonstrated in Figure 2(b). The C41 is a tunable laser inside the data center, e.g. Rack 1 (see Figure 1(a)) that

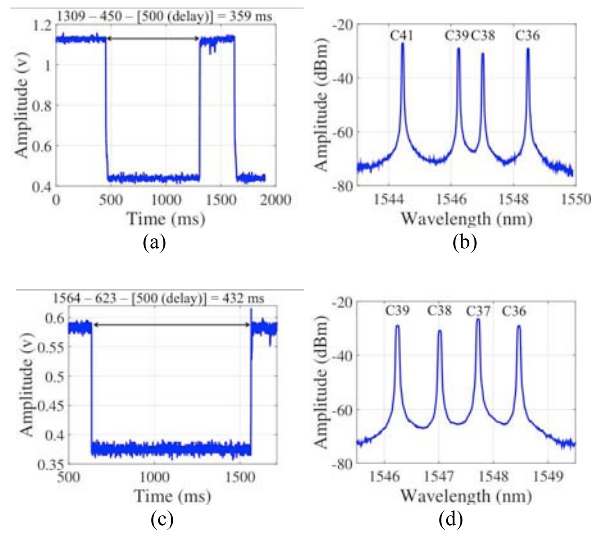

Figure 2: (a) Switching time to add a new connection, (b) Wavelengths before and after defragmentation, (c) Switching time for wavelength defragmentation, and (d) Wavelengths after defragmentation.

is switched to C37 for more efficient usage of the spectrum. In this experiment, first the laser is tuned from C41 to C37, then the OSS makes the connections for adding this channel, and finally the WSS removes C41 and adds C37 to its output port. Figure 2(c) shows the switching time that is 432 ms and Figure 2(d) shows the spectrum after defragmentation at the south port (see Figure 1(a)) of the optical gateway.

IV. Conclusion:

In this work, we presented a bi-directional software-defined optical gateway – a key enabler for remote rack-to-rack connectivity and cross data center scalability. We showed sub-second reconfiguration time to i) make a new WDM connection and ii) perform wavelength defragmentation using tunable lasers.

Acknowledgements:

This research was supported in part by CIAN NSF ERC under grant EEC-0812072 and NSF CNS-1423105. We would also like to thank Calient Technologies for their generous donations to our testbed.

References:

[1] D. Kilper, *et al.*, "Optical Networks Come of Age," OSA Optics and Photonics News 50-57, Sep. 2014.

[2] S. Jain, *et al.*, "B4: Experience with a globally-deployed software defined WAN", ", In Proc. ACM SIGCOMM, 2013.

[3] T. Benson, *et al.*, 2010. Understanding data center traffic characteristics. ", In Proc. ACM SIGCOMM, 2010.

[4] N. Farrington, *et al.*, "Helios: A hybrid electrical/optical switch architecture for modular data centers", In Proc. ACM SIGCOMM, 2010.

[5] G. Wang, *et al.*, "c-through: Part-time optics in data centers", In Proc. ACM SIGCOMM, 2010.

[6] P. Samadi, *et al.*, "Virtual machine migration over optical circuit switching network in a converged inter/intra data center architecture", In press OSA OFC, 2015.