Flex-grid All-optical Interconnect Supporting Transparent Multi-hop Connection in Data Centers

Yuanyuan Hong[1], **Xuezhi Hong**[2], **Sailing He**[1, 3], **and Jiajia Chen**[3]

[1]Center for Optical and Electromagnetic Research
Zhejiang University, Hangzhou 310058, China
[2]ZJU-SCNU Joint Research Center of Photonics
South China Normal University, Guangzhou 510006, China
[3]School of Information and Communication Technology
KTH Royal Institute of Technology, Kista 16443, Sweden

Abstract— In this paper a novel flex-grid all-optical interconnect scheme is proposed for data center, where a transparent multi-hop connection can be realized by using optical bypass rather than going through optical-electric-optical conversions, enhancing energy efficiency for inter-rack communications. Moreover, elastic spectrum allocation is adopted in the proposed scheme to increase the resource utilization and scalability. Transmission performance evaluation has been carried out to verify the feasibility of the proposed scheme.

1. INTRODUCTION

With the rapid growth of new applications, such as cloud service, scientific computing, data analysis and large-scale social network service, the capacity demand for data center network is increasing explosively. Besides the network capacity, another challenging issue in the design and implementation of the data center is the power consumption. Therefore, interconnection scheme which can provide high capacity and low power consumption is of great importance to the data center.

Aiming at addressing the two challenges aforementioned, many optical interconnection schemes have been proposed, including the hybrid electronic/optical approaches, e.g., Helios [1] and c-Through [2], as well as purely optical switching based solutions, e.g., OSA [3]. Compared to the hybrid electronic/optical schemes, the purely optical switching paradigms are transparent to the optical signals (such as data rates, modulation formats) and can offer elastic spectrum allocation, leading to high spectrum efficiency in data center network [4]. The interconnection network in a large-scale data center typically consists of several tiers, e.g., edge, aggregation and core tiers. The aggregation/core tier is responsible for traffic flows among different racks as well as the ones from/to the Internet. Optical circuit switching schemes such as OSA can provide substantially higher bandwidth than electrical packet switching, and hence is considered in the core/aggregation tier. To realize non-blocking connection between any two racks and keep the optical switch matrix in a reasonable size, the multi-hop routing is employed [3], where and the traffic passes several other racks until arriving at the destined rack. However,the existing optical circuit switching based schemes need to use O-E-O conversion at each hop for multi-hop connections [3, 5], which causes relatively low power efficiency and increases the complexity (and cost) of the interconnects at aggregation/core tier. In this regard, we enhance the optical circuit switching based interconnect architecture by offering transparent multi-hop connection. Moreover, the proposed scheme can easily adopt advanced modulation formats and flex-grid technology, which greatly enhances resource efficiency scalability [6]. An extensive assessment has been carried out for our proposed scheme, where the transmission performance of using different modulation formats and the baud rate has been validated by simulation. Besides, the impact of wavelength selective switch (WSS) transfer function on transmission performance has also been investigated.

2. FLEX-GRID ALL-OPTICAL INTERCONNECT OFFERING TRANSPARENT MULTI-HOP CONNECTIONS

The schematic diagram of the proposed scheme is shown in Fig. 1. In a top of the rack (ToR) switch which handles traffics from servers inside the rack, M bandwidth variable optical transceivers (BVTs) are employed to generate and detect optical signals with both varied baud rate and central frequency towards the core tier switch. Here M is the number of flows that one rack can communicate with the other racks simultaneously (i.e., the number of BVTs at the ToR). The ToR switches of different racks are connected via an optical switch matrix (OSM) which serves as the core tier

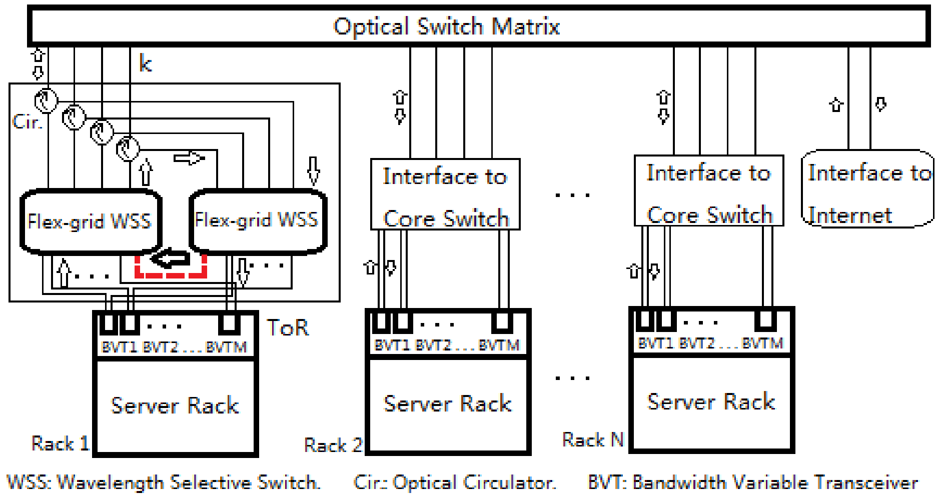


Figure 1: Proposed flex-grid all-optical interconnect supporting transparent multi-hop connections in data centers.

switch handling the communications among racks as well as the interfaces to the outside of the data center. Between the ToR switch (edge tier) and OSM (core tier) two $(M+1) \times K$ flex-grid WSSs are employed. One is for the upstream direction (from edge tier to core tier) and the other is for the downstream direction. In the upstream, the traffic sent by M transceivers at ToR is combined into K groups and then transmitted to the OSM. In the downstream the traffic from the OSM are delivered to K input ports of WSS passing through an array of optical circulators. If the signals are not destined to this rack, they are conveyed to the input port of WSS for the upstream transmission (dashed red line in Fig. 1) and multiplexed with the signals generated by the connected rack. The multiplexed signals are rearranged into k groups for the next hop. The path computation typically is done by the controller where spectrum allocation and routing algorithm are employed. In such a way, the all-optical transparent connection can be established by bypassing the racks rather than sending back to the ToR with optical-electric-optical (O-E-O) conversion as in the multi-hop scheme (e.g., in [3]. Therefore, the all-optical connections can eliminate the power consumption caused by O-E-O conversion and reduce system complexity by avoiding extra transceivers at ToR for relay.

3. PERFORMANCE EVALUATION

In this section, performance evaluation has been carried out by using a professional simulation software VPI transmission Maker 9.1 [7] for communication system performance. We measure symbol error rate (SER) as a function of the received optical power and study on the impact of modulation formats, the number of hops, baud rate as well as WSS transfer function on transmission performance. In the simulation setup, we choose K as 4 which means each rack can communicate with other 4 racks through OSM simultaneously. The choice of K is based on the balance between the number of hops and the size of the OSM [3]. By changing the channel bandwidth and modulation format, the capacity of each WDM channel in the simulation can range from 10 Gb/s to 320 Gb/s, which is designed to meet the various capacity demands in the core tier for different applications [8].

3.1. Flex-grid All-optical Multi-hop Transmission Performance

As shown in Figs. 2 and 3, we investigate the transmission performance of the proposed scheme under different numbers of hops, modulation formats and baud rates by simulation. The Nyquist signaling is employed to realize a high spectrum efficiency system [9]. With Nyquist signaling the transmitted symbols are shaped by the raised-cosine pulses to demonstrate zero inter-symbol interference (ISI) property between consecutive symbols. Raised-cosine pulse shaping of Nyquist signaling brings higher spectrum efficiency at the expense of higher system complexity.For all the modulation formats listed in Fig. 2, the required received optical power increases for a certain level of symbol error ratio (SER) when the number of hops changes from 1 to 3. This can be attributed to the cascaded WSS filtering effect in the all-optical multi-hop connections.Larger inter-symbol interference (ISI) is induced with a larger number of hops. Table 1 and Table 2 show the optical

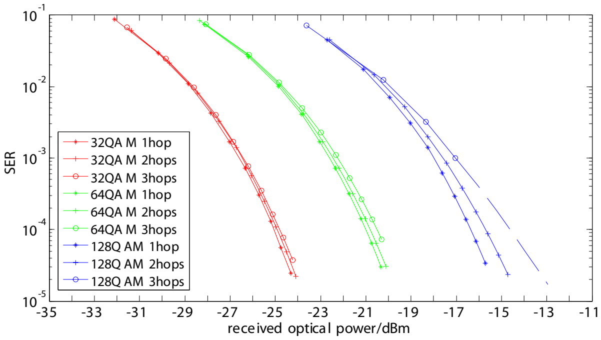


Figure 2: Results of SER versus received optical power in 1/2/3 hops and 32QAM/64QAM/128QAM cases with 40 Gbaud/s.

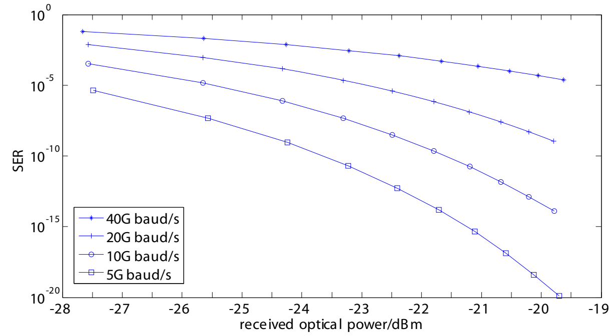


Figure 3: Results of SER versus received optical power for baud rate of 5 Gbaud/s, 10 Gbaud/s, 20 Gbaud/s, and 40 Gbaud/s corresponding to 30 Gbps, 60 Gbps, 120 Gbps and 240 Gbps when using 64QAM.

Table 1: Penalty of received optical power (dB) at SER = 10^{-4} caused by using higher modulation format with the same number of hops.

	1 hop	2 hops	3 hops
32QAM vs. 64QAM/dB	4.10	4.15	4.30
64QAM vs. 128QAM/dB	4.65	5.05	5.85

Table 2: Penalty of received optical power (dB) at SER = 10^{-4} caused by increasing the number of hops for the same modulation format.

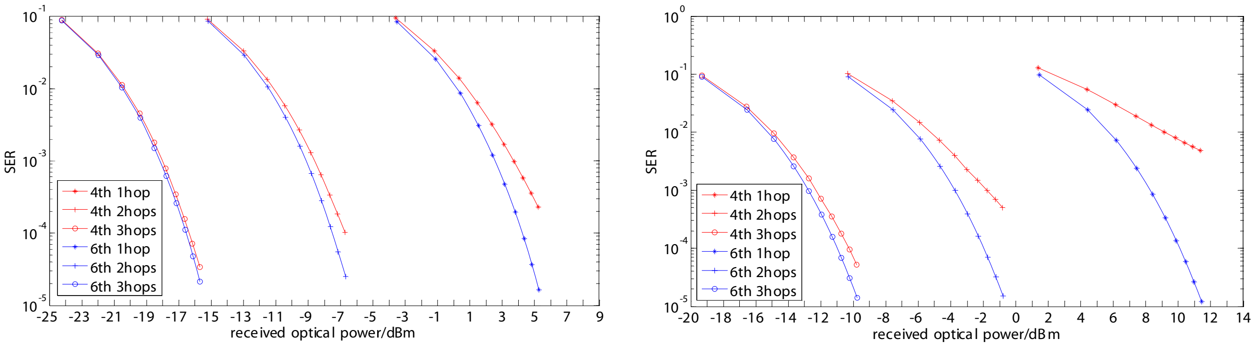
	32QAM	64QAM	128QAM
1 hop vs. 2 hops/dB	0.15	0.20	0.60
2 hops vs. 3 hops/dB	0.15	0.30	1.10

power penalty at SER = 10^{-4} caused by employing higher modulation formats and increasing number of hops, respectively, in both of which a larger optical power penalty is observed for higher modulation format when the number of hops increases. This can be explained by the fact that the signals with higher modulation format and passing more WSSs are more sensitive to ISI.

Flex-grid spectrum allocation can be employed in the proposed scheme. The connections having a spectral slot of 12.5 GHz, 25 GHz, 37.5 GHz, and 62.5 GHz corresponds to a baud rate of 5 Gbaud/s, 10 Gbaud/s, 20 Gbaud/s, and 40 Gbaud/s, respectively. Quadrature amplitude modulation (QAM) is used here. With different modulation format ranging from 4QAM to 256QAM and flexible channel bandwidth, various capacities are achieved (from 10 Gb/s to 320 Gb/s). The SER as a function of received optical power with different baud rate signals are presented in Fig. 3. It can be observed that the channel with higher baud rate has worse SER performance, where the influence of ISI is more severe.

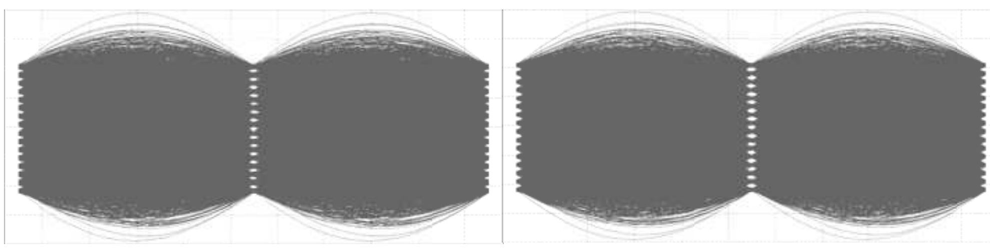
3.2. Impact of WSS Transfer Function

In our scheme, the all-optical multi-hop connection highly relies on the WSS to realize optical bypass. The imperfection of the WSS passband causes ISI and thus lead to performance degradation to the system. Therefore, in this section we further study the impact of WSS passband profile on signal transmission performance. Today's commercial WSSs can have different transfer functions for passband [10]. The WSS considered here is assumed to have a super-Gaussian shape transfer function as $|H(f)|^2 = \exp(-(f - f_0/\Delta f)^{2n})$, wherein f is the optical frequency, f_0 is the center frequency of channel, Δf is the bandwidth and n *is the order of filter* [11]. We have run a set of simulation with WSSs having the same 3-dB bandwidth Δf but different orders of filter. The transfer function with the order of 4 (i.e., $n = 4$) is typical for early generation free-space single dimension MEMS mirror array based WSS, while the order of 6 (i.e., $n = 6$) represents the state-of-the-art WSS [12]. To show the impact of WSS passband profile on the all-optical multi-hop interconnection, we measure the SER as a function of received optical power with either a 4th or



(a) (b)

Figure 4: SER versus received optical power in 40 Gbaud/ssystem with a 4th and 6th super-Gaussian WSS under (a) 128QAM, and (b) 256QAM.



(a) (b)

Figure 5: Eye diagram of 40 Gbaud/s 256QAM signal after 3 hops with a (a) 4th (SER $= 1.38e-2$) or (b) with 6th (SER $= 9.54e-4$) super-Gaussian WSS.

a 6th super-Gaussian WSS under various numbers of hops and modulation formats, as shown in Fig. 4.

According to Fig. 4, the system with 6th super-Gaussian WSS obviously outperforms the one with 4th super-Gaussian WSS in terms of SER under the same numbers of hops and modulation format. The reason is that the passband is more flat with a higher order super-Gaussian WSS [10] which causes less ISI to the Nyquist signal. The filtering induced ISI can be observed in the diagrams shown in Fig. 5. The performance gain by using the 6th super-Gaussian WSS increases with a larger number of hops. It is because by the cascading filtering effect of WSS for multiple hops. We also note that the performance of higher modulation format is more sensitive to the impact of the filter order. Therefore, in the cases with higher modulation formats and larger number of hops, WSS with the higher filter order (i.e., with higher passband flatness) is more preferable for the all-optical multi-hop connection.

4. CONCLUSION

A flex-grid all-optical interconnect scheme supporting transparent multi-hop connection for inter-rack communication is proposed. The transmission performance has been shown in terms of SER for various numbers of hops, modulation formats,baud rates and WSS transfer function. These simulation results verify that supporting up to 3 hops without any O-E-O conversion our scheme still could achieve acceptable SER performance by using WSSs. It is observed that the higher modulation format and larger number of hops will have negative impact on transmission performance, which could be an important factor to be considered in design of routing strategy. Moreover, the results also imply that a high filter order in WSS is recommended for system, particularly for the case with a relatively large number of hops and high modulation format.

ACKNOWLEDGMENT

The work described in this paper was carried out with the support by the National High Technology Research and Development Program (863) of China (No. 2012AA012201), the project titled "Enabling Scalable and Sustainable Data Center Networks" funded by Swedish Foundation of Strategic Research, the project titled "Towards Flexible and Energy-Efficient Data centre Networks" funded by Swedish Research Council, and the China Postdoctoral Science Foundation (No. 2013M531868).

REFERENCES

1. Farrington, N., G. Porter, S. Radhakrishnan, H. H. Bazzaz, V. Sub-Ramanya, Y. Fainman, G. Papen, and A. Vahdat, "Helios: A hybrid electrical/optical switch architecture for modular data centers," *Proceeding of ACM Special Interest Group on Data Communication*, 339–350, New Delhi, India, Aug. 2010.
2. Wang, G., D. G. Andersen, M. Kaminsky, K. Papagiannaki, T. S. Eugene Ng, M. A. Kozuch, and M. Ryan, "C-through: Part-time optics in data centers," *Proceeding of ACM Special Interest Group on Data Communication*, 327–338, New Delhi, India, Aug. 2010.
3. Chen, K., A. Singla, A. Singh, K. Ramachandran, L. Xu, Y. Zhang, X. Wen, and Y. Chen, "OSA: An optical switching architecture for data center networks with unprecedented flexibility," *T. Netw.*, Vol. 22, No. 2, 498–511, 2014.
4. Wen, K., X. Cai, Y. Yin, D. J. Geisler, R. Proietti, R. P. Scott, N. K. Fontaine, and S. J. B. Yoo, "Adaptive spectrum control andmanagement in elastic optical networks," *J. Sel. Area Comm.*, Vol. 31, No. 1, 39–48, 2013.
5. Liu, L., W. Peng, R. Casellas, T. Tsuritani, I. Morita, R. Martínez, R. Muñoz, and S. J. B. Yoo, "Design and performance evaluation of an open flow-based control plane for software defined elastic optical networks with direct detection optical OFDM (DDO-OFDM) transmission," *Opt. Express*, Vol. 22, No. 1, 30–40, 2014.
6. Christodoulopoulos, K., I. Tomkos, and E. A. Varvarigos, "Elastic bandwidth allocation in flexible OFDM-based optical networks," *J. Lightwave Technol.*, Vol. 29, No. 9, 1354–1366, 2011.
7. Garcia, M., H. Coll, D. Bri, and J. Lloret, "Software tools and simulators in the education of engineering of telecommunications," *Proceeding of 5th WSEAS/IASME International Conference on Engineering Education and Research*, 337–342, Heraklion, Greece, Jul. 2008.
8. Mahimkar, A., A. Chiu, R. Doverspike, M. D. Feuer, P. Magill, E. Mavrogiorgis, J. Pastor, S. L. Woodward, and J. Yates, "Bandwidth on demand for inter-data center communication," *Proceedings of the 10th ACM Workshop on Hot Topics in Networks*, Cambridge, MA, USA, Nov. 24, 2011.
9. Shi, J., J. Zhang, L. Tao, Y. Wang, Y. Wang, and N. Chi, "Characteristics of filtering effect for a polarization division multiplexed-quadrature phase shift keying signal with Nyquist and super-Nyquist spectrum efficiency employing various algorithms," *Opt. Eng.*, Vol. 53, No. 7, 076109–076109, 2014.
10. "A performance comparison of WSS switch engine technologies," JDSU white paper, http://www.jdsu.com/productliterature/wsscomp_wp_cms_ae.pdf.
11. Pulikkaseril, C., L. A. Stewart, M. A. F. Roelens, G. W. Baxter, S. Poole, and S. Frisken, "Spectral modeling of channel band shapes in wavelength selective switches," *Opt. Express*, Vol. 19, No. 9, 8458–8470, 2011.
12. Marom, D. M., D. T. Neilson, D. S. Greywall, C. Pai, N. R. Basavanhally, V. A. Aksyuk, D. O. López, F. Pardo, M. E. Simon, Y. Low, P. Kolodner, and C. A. Bolle, "Wavelength-selective $1 \times k$ switches using free-space optics and MEMS micromirrors: Theory, design, and implementation," *J. Lightwave Technol.*, Vol. 23, No. 4, 1620–1630, 2005.