

Novel flat data center network architecture based on optical switches with fast flow control

Citation for published version (APA):

Yan, F., Miao, W., Dorren, H., & Calabretta, N. (2016). Novel flat data center network architecture based on optical switches with fast flow control. *IEEE Photonics Journal*, 8(2), 1-10. [0601310].
<https://doi.org/10.1109/JPHOT.2015.2513209>

DOI:

[10.1109/JPHOT.2015.2513209](https://doi.org/10.1109/JPHOT.2015.2513209)

Document status and date:

Published: 20/04/2016

Document Version:

Accepted manuscript including changes made at the peer-review stage

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Novel Flat Data Center Network Architecture Based on Optical Switches with Fast Flow Control

Fulong Yan^(1,2), Wang Miao⁽¹⁾, Harm Dorren⁽¹⁾, Nicola Calabretta⁽¹⁾

⁽¹⁾COBRA Research Institute, Eindhoven University of Technology, Eindhoven 5600MB, the Netherlands

⁽²⁾Beihang University, Beijing 100191, China

Abstract: We propose a novel flat and scalable data center network (DCN) architecture based on fast (nanoseconds) distributed buffer-less optical switches and efficient optical flow control. The proposed DCN architecture scales as the square of the port count of the optical switches. In order to investigate the performance of the proposed architecture, the system operation of an electronic Top-of-the-Rack (ToR) and the optical switch is fully described and all functional sub-systems are modeled. The performance in terms of DCN scalability, average latency, packet loss, normalized throughput of the network, and electronic buffer size of the ToR are numerically assessed for a medium-size data center supporting 5,760 servers and a large size data center connecting 100,000 servers. Considering a traffic pattern with high inter-cluster (40%) traffic distribution and buffer size of 40KB, the results report an end-to-end latency of less than 8.1 μ s (including retransmission), a packet loss $< 10^{-7}$ under load of 0.4 for the large size data center. Moreover, we provide a preliminary experimental validation of the DCN by using a 4×4 optical switch prototype showing dynamic switching at 40 Gb/s and error free operation with less than 1.5 dB penalty for the longest path.

Index Terms: Optical packet switch, data center networks, optical label processor, optical flow control, OMNeT++ simulation.

1. Introduction

Driven by the cloud computing paradigm and data-rich applications, data centers experience a steady annual increase of over 30% in the amount of accommodated traffic [1]. Emerging data center applications and workloads are creating new communication patterns with more than 75% of the total data center traffic processed within data centers (server to server and rack to rack) [2]. This huge increase of intra data center traffic requires architectural and technological innovations to the underlying interconnect networks in order to enable the scalable growth both in communicating endpoints and traffic volume, while decreasing the costs and the energy consumption.

Nowadays, for the widely deployed architectures based on electrical switches, interconnecting a large amount of servers with commercial switches will either require multiple-layer solution or utilize customized core switches, which introduce extra latency and huge cost. Accompanying with the fast development of the electrical switching technology, 40 Gb/s per port Ethernet switches have already been deployed in the data centers, and 100 Gb/s and beyond is under development. The increase of the data rate demands DCN with higher capacity to prevent the issue of bandwidth bottlenecks and latency [3] limiting east-west communications between groups of servers located in different racks and clusters. Building data center with scaling out number of servers (>10000) and scaling up data rate (>10 Gb/s) requires electrical switches with high radix to avoid hierarchical multi-layers architectures with bandwidth bottleneck, high end-to-end latency as well as poor cost-efficiency. However, the limited I/O bandwidth of the switch application specific integrated circuit (ASIC) caused by the limited ball grid array (BGA) density will prevent the implementation of high radix switch at high data rates [4]. Multi-stage switching architectures could built up electronic switches with high radix at the expense of a large number of optical interconnects that lead to a high costs, power consumption, and latency.

On the contrary, optical switching technologies can transparently switch high data rate/format signals, reducing the required number of switch ports and avoiding costly O/E/O in hierarchical multiple-stage architecture. Optical circuit switch (OCS) with large port-count is commercially available. However, the milliseconds configuration time introduces large latency and prevents efficient statistical multiplexing of data packets. Helios [5] and c-Through [6] utilize MEMS switches for optical circuit switching and electrical switches for small data packet switching. On the other hand, fast optical switch with sub-microseconds reconfiguration time has been so far demonstrated only for a moderate port count [7]. Simple multi-layer structure employing low-radix optical packet switch (OPS) can fulfill the interconnection requirements, which again is limited by the issues of hierarchical topology including bandwidth bottleneck as well as large latency and lack of optical buffers [8-12]. Recently, a scalable OPS with modular structure has been investigated and experimental verified in a simplified star topology DCN architecture [13]. However, despite the potential scalability of a single OPS to provide a large port count, a DCN capable to interconnect thousands of racks in a flat fashion requires an innovative DCN architecture with distributed optical switches.

In this work, we investigate a novel flat DCN architecture that employs distributed buffer-less optical switches with optical flow-control to implement scalable parallel all-optical intra-/inter-cluster networks. The novel all-optical DCN architecture enables large-scale interconnectivity by using feasible optical switches with moderate number of ports. While the nanoseconds operation of the optical switches allows efficient statistical multiplexing of high data rate optical packets, the implementation of fast optical flow control mitigates the lack of optical buffers by retransmitting the contended packets.

Numerical simulation based on OMNeT++ is employed to assess the scalability and the performance of the novel flat DCN in terms of latency, packet loss, and normalized throughput under different inter/intra-cluster traffic distributions and patterns. Preliminary experimental validation of the DCN architecture based on 4×4 switch prototypes indicates successful error-free dynamic switching operation for 40 Gb/s data packets with less than 1.5 dB power penalty.

The paper is organized as follows. In Section II, the novel DCN architecture and the system operation of the ToR and fast optical switch with optical flow control are described. Section III reports the simulations layout and analyzes the DCN performance and scalability under different traffic distributions. In Section IV, the proof-of-concept of the DCN architecture has been experimentally validated. Finally, Section V concludes the paper by summing up the most important results.

2. System Operation

The proposed flat DCN architecture based on two parallel inter- and intra-cluster networks is shown in Fig. 1. It consists of N clusters, and each cluster groups M racks by using an intra-cluster optical switch (IS). Each rack contains k servers interconnected by an electronic ToR switch. The ToR switch is equipped with two wavelength division multiplexing (WDM) bi-directional optical links. One optical link is used to interconnect the ToR to the IS for intra-cluster communication. The second optical link interconnects the ToR to the inter-cluster optical switch (ES). The i -th ES interconnects the i -th ToR of each cluster, with $i = 1, \dots, N$. It can be seen from Fig.1 that a single-hop communication is needed between racks of the same cluster, while at most two hops is needed to interconnect racks of different clusters. It is worth to notice that the proposed DCN architecture allows ToRs to be interconnected via different path connections, increasing network fault-tolerance. As shown in Fig. 1, the number of interconnected ToRs (and servers) scales as $N \times M$. By using a feasible 64×64 port IS and ES, up to 4,096 ToRs and then 163,840 servers (in case each ToR groups 40 servers) can be interconnected.

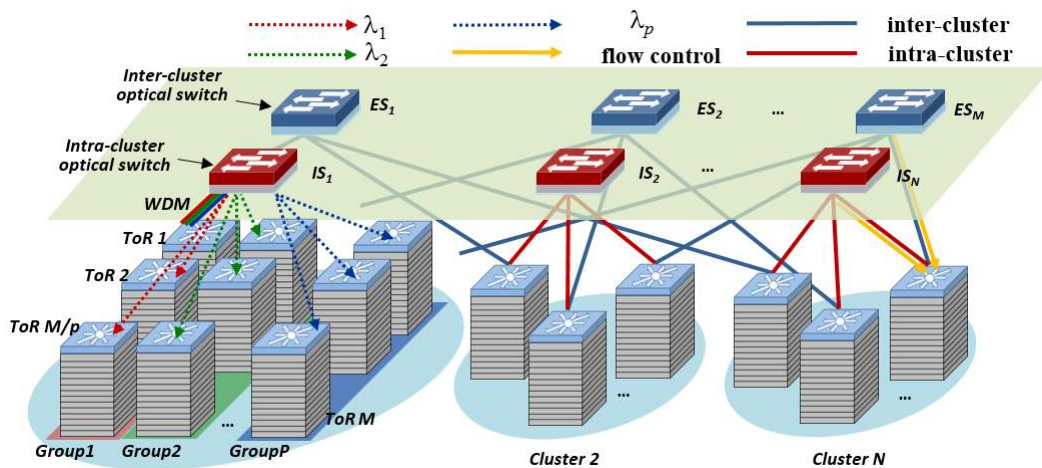


Fig. 1. Novel flat DCN architecture based on optical switch with fast flow control

The functional blocks of the ToR are illustrated in Fig. 2. The ToR aggregates the traffic coming from the k servers. Assume that each server produces a traffic at data rate b_{serv} . Therefore, ToR aggregates a total traffic of $k \times b_{serv}$. Part of this traffic is exchanged between intra-rack servers (intra-rack traffic), and the rest of traffic is directed to servers in the same cluster (intra-cluster traffic) and servers located in different clusters (inter-cluster traffic).

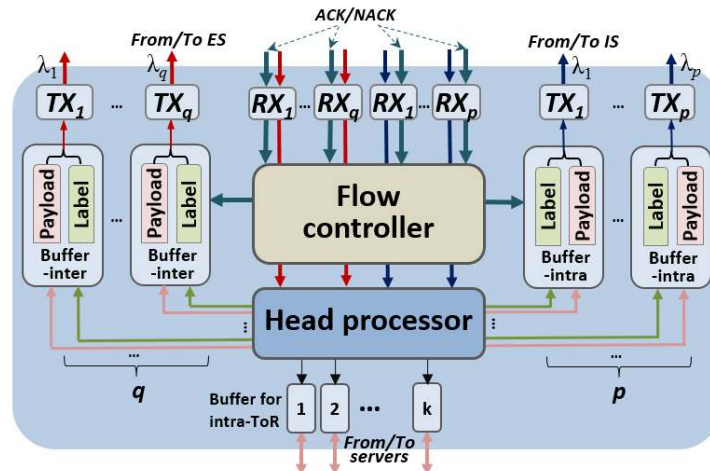


Fig. 2. ToR schematic diagram equipped with WDM transceivers and fast flow control

When the packets sent by the server arrive to the ToR, the packet header (destination) will be checked by the header processor. In case of intra-rack traffic, the ToR directly processes and forwards the traffic to the server destination. Intra-rack contention is solved by using k intra-rack buffer queues. In case of intra- or inter-cluster traffic, the traffic is directed to the intra-cluster interface and inter-cluster interface, respectively. The intra-cluster interface consists of p WDM transceivers (TX in the figure) with dedicated electronic buffers that interconnects the ToR to the IS optical switch with optical flow controlled links. According to the packet header destination, each buffer stores the traffic destined to a group of intra-cluster ToRs. The WDM transceiver forwards the stored traffic destined to a group of M/p ToRs, and in general the buffer of the transceiver at λ_i stores and forwards the traffic destined to the intra-cluster ToR $(i-1) \times M/p + 1, \dots, \text{ToR } i \times M/p$ (Group i), with $i=1, \dots, p$. As an example shown in Fig. 1, with regards to ToR 1, the buffer of the transceiver at λ_1 stores and forwards the traffic destined to the intra-cluster ToR 2, ToR 3, ..., ToR M/p (Group 1). Moreover, the allocation of the wavelengths should meet the requirement that for the ToRs in the same group, the wavelengths used to address the same ToR destination should not overlap with each other. A wavelength mapping rule for the ToRs in Group i interconnected by the IS is shown in Table I, in which $F = M/p$. The copy of the stored packet, combined with an optical label generated by the ToR, is forwarded to the IS. The optical label, processed on-the-fly by the IS optical switch, determines the forwarding of the packet to the ToR destination. Due to the buffer-less operation of the optical switch, if a contention occurs at the IS optical switch, the optical flow control provides a NACK to retransmit the stored packet, otherwise an ACK to release the packet from the buffer [13], similarly as demonstrated in [14, 15].

TABLE I
ToR GROUPING AND WAVELENGTH MAPPING RULES

ToR of Group i	Destination ToR (at the i -th RX link)					Group NO.
ToR ₁	λ_1	-	ToR ₂ ^{i}	...	ToR _F ^{i}	1
	λ_2	ToR _{F+1} ^{i}	ToR _{F+2} ^{i}	...	ToR _{2F} ^{i}	2
					
	λ_p	ToR _{(p-1) \times F + 1} ^{i}	ToR _{(p-1) \times F + 2} ^{i}	...	ToR _M ^{i}	p
ToR ₂	λ_1	ToR _{(p-1) \times F + 1} ^{i}	ToR _{(p-1) \times F + 2} ^{i}	...	ToR _M ^{i}	p
	λ_2	ToR ₁ ^{i}	-	...	ToR _F ^{i}	1
					
	λ_p	ToR _{(p-2) \times F + 1} ^{i}	ToR _{(p-2) \times F + 2} ^{i}	...	ToR _{(p-1) \times F} ^{i}	$p-1$
.....						
ToR _{M/p} (ToR _F)	λ_1	ToR _{F+1} ^{i}	ToR _{F+2} ^{i}	...	ToR _{2F} ^{i}	2
	λ_2	ToR _{2F+1} ^{i}	ToR _{2F+2} ^{i}	...	ToR _{3F} ^{i}	3
					
	λ_p	ToR ₁ ^{i}	ToR ₂ ^{i}	...	-	1

The inter-cluster interface structure and operation is similar to the intra-cluster interface. The inter-cluster interface consists of q parallel WDM transceivers with dedicated buffers to connect the ToR to the ES optical switch. The number of p and q WDM transceivers depends on the required capacity to guarantee a target oversubscription. For instance if the ToR oversubscription should be equal to 1, the capacity offered by the WDM transceivers should be equal to $k \times b_{serv}$ (the total data traffic generated by the servers). For instance for $k = 40$ servers with 10 Gb/s interface, to guarantee an oversubscription equal to 1, the $p+q$ WDM transceivers should provide a capacity of 400 Gb/s. Multiple 25 Gb/s or 50 Gb/s p and q WDM transceivers can be tailored according to the expected intra- and inter cluster data traffic. For intra-cluster connection, the packet undergoes a single IS hop to the ToR destination and then the server within the rack. In case of inter-cluster connection, it might be that the packet has to undergo two OPS hops (from ToR to ES \rightarrow ToR \rightarrow IS to ToR destination or from ToR to IS \rightarrow ToR \rightarrow ES to ToR destination) before reaching the final destination (see example in Fig. 1). In this case the packet has to be detected and processed by the intermediate ToR, adding extra latency.

The buffer-less optical switch of IS based on broadcast&select architecture is schematically shown in Fig. 3 (for ES, just substitute q for p and N for M). It processes in parallel the multiple WDM input packets by using autonomously controlled $1 \times F$ optical switches. Note that F is equal to M/p for IS, and F equals N/q for ES since each of the p (or q) WDM transceivers groups the traffic destined to M/p intra-cluster ToRs (or N/q inter-clusters ToRs). This allows to scale the port count of the switch to $M \times M$ (or $N \times N$) by using $p \times M$ (or $q \times N$) parallel smaller $1 \times F$ switches with moderate broadcast splitting losses and therefore lower OSNR degradation.

At the optical switch node, the optical label is extracted and processed by the switch controller, while the optical payload is transparently switched by the $1 \times F$ switch. According to the optical label, the switch controller enables the fast (nanosecond) SOA-based gates of the $1 \times F$ switch. Multicast operation is also supported by enabling multiple SOA-based gates. The use of SOA gates has twofold advantages: fast nanosecond switching time, and optical amplification of the signal to compensate the splitting losses of the broadcast&select architecture. Moreover, the proposed technologies allow photonic integration of the switch in a single chip [16]. The control complexity and the configuration time are largely determined by the label

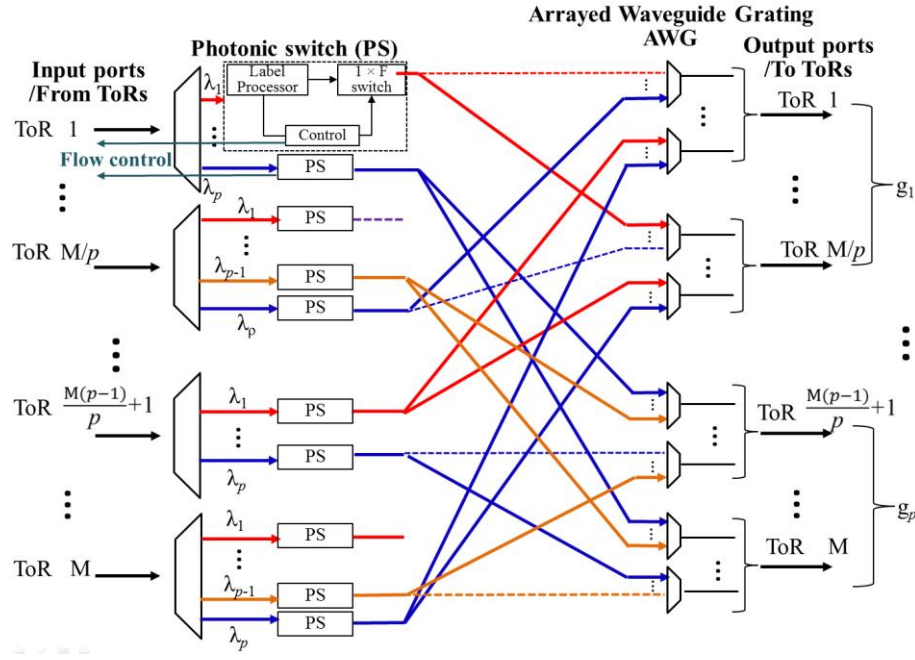


Fig. 3. Schematic of the optical switch (IS) architecture.

processing time. The in-band RF tone labeling technique and the distributed control have resulted in port-count independent nanoseconds reconfiguration time allowing operation on large as well as small flows exploiting statistical multiplexing [17, 18]. Benefitted from the modular structure and parallel processing of each channel, the possible contentions among the F input ports can be resolved in a distributed manner which significantly minimizes the processing latency. The packet with higher priority is forwarded to one or several of the F possible output ports and the others are blocked and the corresponding flow control signals are generated and sent back to the ToRs. According to the received ACK (or NACK), the flow controller at the ToR releases (or retransmits) the packets stored in the buffers.

A comparison between the proposed system and the DCN architectures building upon electrical switches (e.g. Fat-tree and Leaf-Spine) has been numerically investigated in [19]. It is reported that the proposed flat architecture can provide higher bi-section bandwidth and lower latency benefitted from the transparency of optical switching technology without changing the infrastructure or adding more switching ports. By eliminating the massive O/E/O conversions, the power consumption and the cost have been significantly reduced. Moreover, the parallel intra-/inter-cluster switching networks allow for suitability of higher data-rate and square scalability of the optical switch radix, making large-scale and high-density DCN achievable with moderate radix optical switches.

3. Simulation Results

OMNeT++ [20] has been employed as the simulation tool to fully investigate the performances of the proposed DCN under packet transmission. To investigate the scalability, DCNs consisting of 144 ToRs interconnected by 12×12 radix IS and ES and 2500 ToRs interconnected by 50×50 radix IS and ES have been considered. Each ToR connects 40 servers with 10 Gb/s link, resulting in DCNs of 5,760 (medium size) and 100,000 (large size) servers. The packet size is set to 1000 Bytes at the ToR output that each slot is 800 ns. The total data rate of the optical links to the IS and ES is 200 Gb/s (4×50 Gb/s WDM channels). Therefore, the designed oversubscription of the system is 1:1. The round trip time (RTT) for ToRs with direct connection is 560 ns, including the delay to process the labels and control the switches (60 ns), and the optical link ($2 \times 50 = 100$ m) latency (500 ns). And the latency for direct connection ToRs (single OPS hop via an OPS) is 560 ns, while the latency for ToRs without direct connection (two OPS hops via OPS \rightarrow intermediate ToR \rightarrow OPS) would be 1200 (560 + 80 + 560) ns.

TABLE II
TRAFFIC PATTERNS AND DISTRIBUTIONS

Traffic patterns	Heavy inter-cluster traffic (case A)	Light inter-cluster traffic (case B)
Inter-cluster	40%	25%
Intra-cluster	20%	25%
Intra-ToR	40%	50%

The operation of the system is synchronous and the simulations run for 1×10^8 time slots. At each time slot, the server generates a packet according to a binomial distribution with a fixed load, thus there is a certain probability to have a new packet at each server. All servers are programmed to operate independently with the same load. Traffic is classified into 3 categories (inter-cluster, intra-cluster and intra-ToR). In each category, packets destinations are chosen randomly between all possible destinations according to a uniform distribution. Considering the facts that most of the traffic is exchanged within the ToR and cluster, two traffic distributions with different ratios of 3 categories are studied. As shown in Table II, case A is the large inter-cluster traffic while case B is light inter-cluster traffic. Different buffer size for 200 Gb/s link ports are considered (40 KB and 60 KB). Since the inter-cluster traffic will pass both the intra-cluster link and the inter-cluster link, the real oversubscription of case A and case B is 1:1 and 3:4 respectively.

In the section below, the experimental results on the performances of the DCN are presented and discussed. Figure 4 illustrates the average server-to-server latency including the latency to switch the packet to the server inside the rack as a function of the load. $8.1 \mu\text{s}$ server-to-server latency has been obtained at load of 0.4 for both medium and large size DC and different buffer sizes. It is clearly shown that the average latency of large size DC is greater than medium size DC, and the latency with buffer size 60 KB is greater than buffer size 40 KB. For larger load, the effect of the buffer on the system performance of latency is visible. The average latency differences between buffer size 40 KB and 60KB become large as the load increases. For load above 0.4, latency increases rapidly as it may take several slots for packet being successfully transmitted, before to saturate at high load. When the load is close to 1, the average latency of both medium size and large size DC under case B increases rapidly and exceeds case A for the reason that the packet loss of case A is higher than case B, and the packet loss ratio can not be neglected, thus there are more packets going through the inter-cluster in case B, which contributes to the larger latency.

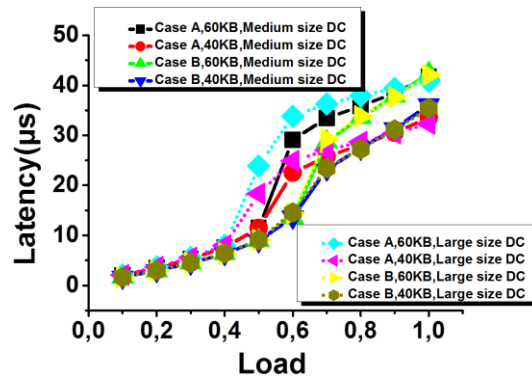


Fig. 4. Server to server average latency versus load

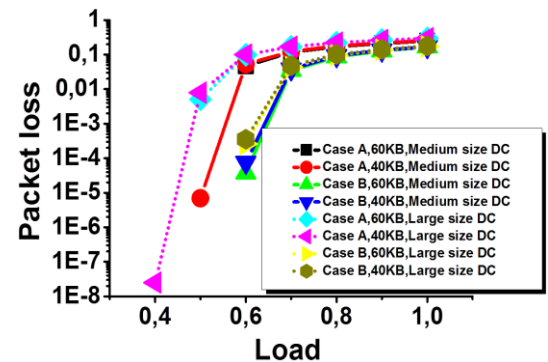


Fig. 5. Packet loss versus load

Fig. 5 shows the packet loss of the system. A packet loss less than 10^{-5} for load of 0.5 has been achieved for medium size DC under case A, while the packet loss is less than 10^{-7} for large size DC at load of 0.4 under case A. In agreement with the curve of the average latency, the packet loss increases with the load. For load exceeding 0.5 (case A, large size DC) and 0.6 (case A, medium size DC), the packets loss is unavoidable as the links and the buffers are fully occupied, and larger buffers do not help to decrease packet loss but provide extra latency. Despite the influence of the DC size, the packet loss of case A is greater than case B simply for case A has more inter-cluster traffic and the buffer will be occupied quicker compared with case B.

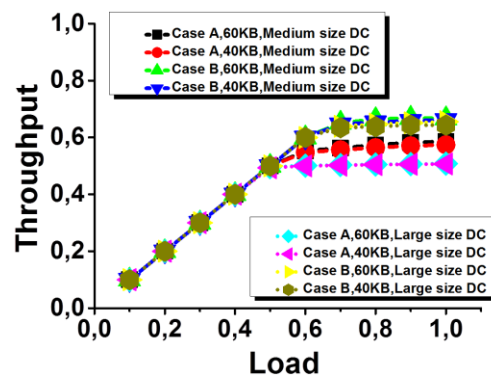


Fig. 6. Normalized throughput versus load

Fig. 6 reports the normalized network throughput of both DC sizes under different conditions. As shown in Fig. 6, the normalized throughput increases as the load increase. However, for case A, large size DC starts to saturate at load of 0.5 and medium size DC saturates at load of 0.6. And both size DCs saturate at load of 0.7 under case B. The reason behind is that under case A, there are more inter-cluster traffic which needs pass two hops, therefore it takes up more link bandwidth compared with intra-cluster traffic.

4. Experimental Results

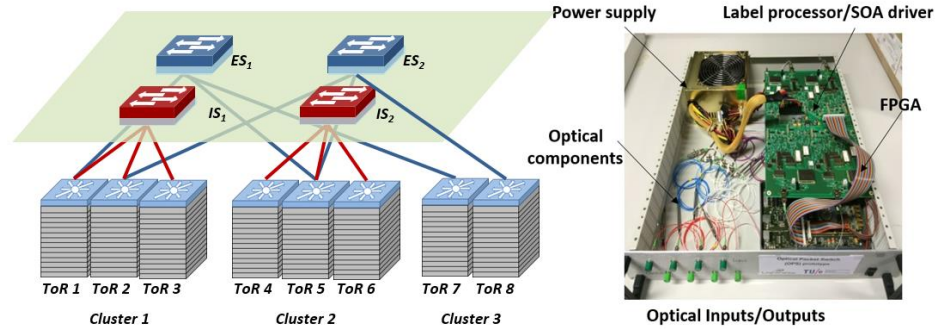


Fig. 7. (a) Experiment set-up (b) 4 × 4 optical switch prototype

The experimental set-up to validate the novel DCN consists of 3 clusters each formed by 3 ToRs as shown in Fig. 7(a). For inter- and intra-cluster connection we employed 4 × 4 optical switch prototypes shown in Fig. 7(b). The ToRs generate 40 Gb/s NRZ-OOK packets with 540 ns duration and 60 ns guard time at $\lambda_1 = 1552.55$ nm, $\lambda_2 = 1554.17$ nm and $\lambda_3 = 1555.78$ nm for each optical switch input port. In the ToR, an FPGA acting as a local controller stores the look-up table with the packet destination. A 4-bit (2 for inter- and 2 for intra-) in-band RF tone label according to the port destination will be generated for each packet [21]. The label information will be stored in a FIFO queue with a size of 16 packets and removed from the queue in response to a positive ACK. Otherwise the label and payload are retransmitted [22]. When the packet is received at the ToR, the optical label will be first processed to check if it belongs to the current ToR, otherwise it will be optically bridged to the next switch node and a copy will be stored. The locally generated traffic and the packets to be bridged are scheduled by the 1 × 2 switch and the bridged packets have been given the higher priority when no retransmission is under process [23]. The blocked traffic will wait for the next available time slot.

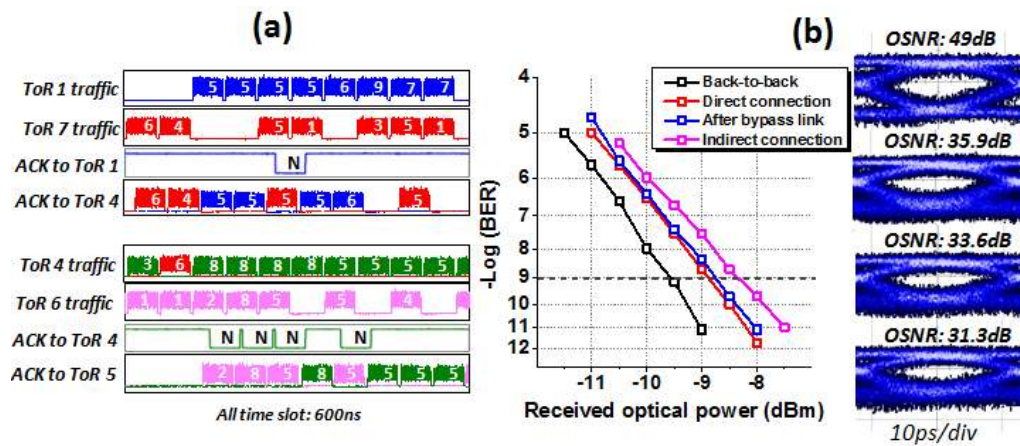


Fig. 8. (a) Packets traces for dynamic switching (b) BER curves and eye diagrams

Fig. 8(a) shows the recorded packets and ACK signals to validate the intra- and inter-cluster dynamic switching operation. As a selected case, intra-cluster (direct path) traffic from ToR 4 destined ToR 6 (ToR j denotes the j th ToR in the system, as shown in Fig. 7(a)), and inter-cluster (indirect path via ES₁ and ToR 4) ToR 1 and ToR 7 traffic with destination ToR 5 are investigated. The traffic generated by the ToRs is reported in Fig. 8(a). Each packet has been labeled with the destination ToR j ($j = 1, \dots, 9$). At the switch node ES₁, the packets from the ToR 1 and the ToR 7 (higher priority) with destination ToR 5 are transmitted to ToR 4. In case of contention, the traffic from ToR 1 will be blocked, a negative ACK is sent to ToR 1 and the blocked packet is retransmitted in the next time slot. At the ToR 4, packets with destination ToR 5 and ToR 6 will be optically forwarded to the IS₂ and stored in case retransmission is needed. At IS₂, packets coming from ToR 4 and ToR 6 (higher priority) and destined ToR 5 might have contention. Optical flow control provides, in case of contention, the ACK signals to ToR 4 as shown in Fig. 8(a), where the packets received by the ToR 5 are also reported. The BER performance and the eye diagrams for the 40 Gb/s payload are shown in Fig. 8(b). Error free operation with less than 1.5dB power penalty and 31.3dB OSNR for the indirect connection (longest path) has been obtained.

5. Conclusions

We propose a novel flat DCN architecture based on distributed buffer-less optical switches with fast flow control. The system performance of the proposed architecture was numerically assessed with OMNeT++ and experimentally verified by utilizing 4 × 4 OPS prototypes. The numerical results show that the proposed DCN architecture allows for 8.1 μ s end-to-end average latency and less than 10⁻¹⁰ packet loss for traffic load of 0.4 when it inter-connected 100,000 servers. Besides, the experiment

on the test-bed employing 4×4 OPS shows dynamic switching and error free operation with less than 1.5 dB penalty. Those results indicate that a scalable DCN can be built by using feasible optical switches with moderate switch radix, which provides a promising solution for the DCN architecture.

Acknowledgements

The authors would like to thank the FP7 LIGHTNESS project (n° 318606) and NSFC 61271196 for supporting this work.

References

- [1] Cisco global cloud index: Forecast and Methodology, CISCO, (2013).
- [2] T. Benson et al., "Network Traffic Characteristics of Data centers in the Wild," Proc. ACM, New York (2010)
- [3] Alizadeh, Mohammad, et al. "Less is more: trading a little bandwidth for ultra-low latency in the data center." *Proceedings of the 9th USENIX conference on Networked Systems Design and Implementation*. USENIX Association, 2012.
- [4] N. Binkert et al., "The role of optics in future high radix switch design." Proc. IEEE ISCA, California (2011).
- [5] N. Farrington, G. Porter, S. Radhakrishnan, H. H. Bazzaz, V. Subramanya, Y. Fainman, G. Papen, A. Vahdat, Helios: a hybrid electrical/optical switch architecture for modular data centers, ACM SIGCOMM Computer Communication Review 41 (4) (2011) 339-350.
- [6] G. Wang, D. G. Andersen, M. Kaminsky, K. Papagiannaki, T. S. Ng, M. Kozuch, M. Ryan, c-Through: Part-time optics in data centers, in: ACM SIGCOMM Computer Communication Review, Vol. 40, 2010, pp. 327-338.
- [7] C. Kachris et al., Optical Interconnects for Future Data Center Networks, Springer (2013).
- [8] O. Liboiron-Ladouceur, P. G. Raponi, N. Andriolli, I. Cerutti, M. S. Hai, P. Castoldi, A scalable SpaceTime multi-plane optical interconnection network using energy-efficient enabling technologies [invited], Optical Communications and Networking, IEEE/OSA Journal of 3 (8) (2011) A1-A11.
- [9] R. Proietti, Z. Cao, C. Nitta, Y. Li, S. Yoo, A scalable, low-latency, high-throughput, optical interconnect architecture based on arrayed waveguide grating routers, Lightwave Technology, Journal of 33 (4) (2015) 911-920. doi:10.1109/JLT. 2015.2395352.
- [10] P. N. Ji, D. Qian, K. Kanonakis, C. Kachris, I. Tomkos, Design and evaluation of a flexible-bandwidth OFDM-based intra-data center interconnect, Selected Topics in Quantum Electronics, IEEE Journal of 19 (2) (2013) 3700310-3700310.
- [11] G. Porter, R. Strong, N. Farrington, A. Forencich, P. Chen-Sun, T. Rosing, Y. Fainman, G. Papen, A. Vahdat, Integrating microsecond circuit switching into the data center, Vol. 43, ACM, 2013.
- [12] C. Kachris et al., "Optical Interconnects for Future Data Center Networks", Springer (2013).
- [13] Miao, Wang, et al. "Novel flat datacenter network architecture based on scalable and flow-controlled optical switch system." Optics express 22.3 (2014): 2465-2472.
- [14] Shacham, Assaf, and Keren Bergman. "An experimental validation of a wavelength-striped, packet switched, optical interconnection network." Journal of Lightwave Technology 27.7 (2009): 841-850.
- [15] Proietti, Roberto, et al. "All-optical physical layer NACK in AWGR-based optical interconnects." Photonics Technology Letters, IEEE 24.5 (2012): 410-412.
- [16] N. Calabretta, K. Williams, and H. Dorren, "Monolithically Integrated WDM Cross-Connect Switch for Nanoseconds Wavelength, Space, and Time Switching," in Proc. ECOC 2015.
- [17] H. J. S. Dorren, S. Di Lucente, J. Luo, O. Raz, and N. Calabretta, "Scaling Photonic Packet Switches to a Large Number of Ports," Journal of Optical Communications and Networking, vol. 4, no. 9, pp. A82-A89, 2012
- [18] N. Calabretta, 'FPGA based label processor for low-latency and large port count optical packet switches,' Journal of Lightwave Technology, vol. 30, n. 19, pp. 3173-3181, 2012.
- [19] Yan, Fulong, et al. "On the cost, latency, and bandwidth of LIGHTNESS data center network architecture." Photonics in Switching (PS), 2015 International Conference on. IEEE, 2015.
- [20] OMNeT++ Network Simulation, <http://www.omnetpp.org>.
- [21] W. Miao et al., "Novel flat datacenter network architecture based on scalable and flow-controlled optical switch system," Opt. Express, vol. 22, pp. 2465 (2014).
- [22] J. Gripp, J. E. Simsarian, J. D. LeGrange, P. Bernasconi, and D. T. Neilson, "Photonic Terabit Routers: The IRIS Project," in Optical Fiber Communication Conference, San Diego, 2010.
- [23] W. Miao et al., "Performance Assessment of Optical Packet Switching System with Burst-Mode Receivers for Intra-Data Center Networks", European Conference on Optical Communications, (2014).