Fast Optical Circuit Switch for Intra-Datacenter Networking

Koh UEDA, Student Member, Yojiro MORI, Hiroshi HASEGAWA, Hiroyuki MATSUURA†, Kiyo ISHII††, Haruhiko KUWATSUKA††, Shu NAMIKI††, Toshio WATANABE††*, Members, and Ken-ichi SATO†††, Fellow

SUMMARY This paper presents a fast and large-scale optical circuit-switch architecture for intra-datacenter applications that uses a combination of space switches and wavelength-routing switches are utilized. A 1,440 × 1,440 optical switch is designed with a fast-tunable laser, 8 × 8 delivery- and-coupling switch, and a 180 × 180 wavelength-routing switch. We test the bit-error-ratio characteristics of all ports of the wavelength-routing switch using 180-wavelength 10-Gbps signals in the full C-band. The worst switching time, 498 microseconds, is confirmed and all bit-error ratios are acceptable.

key words: switching, circuit, array waveguide devices

1. Introduction

With the advent of cloud-computing and big-data services, intra-datacenter traffic is growing at 24% a year [1]. Recent advances in machine-to-machine communication spurred by artificial-intelligence based applications will further strengthen this trend. A large part of intra-datacenter traffic originates from elephant flows, which are produced by bandwidth-intensive applications such as video streaming, storage backup, and virtual machine migration. The acronym, mice flows, derives from latency-sensitive applications including web searches [2]. With the increase in intra-datacenter traffic, the power consumption of electrical Ethernet and packet switches is becoming a crucial issue. When the number of server racks explodes and high radix electrical switches are unavailable, necessary switching level increases, which results in excessive delay.

One feasible solution is the opto-electronic hybrid switching network [3], [4]. To realize such networks, optical circuit/flow switches that eliminate costly optical-to-electrical and electrical-to-optical conversions and offload most of the elephant flows from electronic switches are being investigated [4]–[9]. The key attributes for developing cost-effective optical switches are their port count and switching speed. High port count optical switches enable the construction of flatter networks which contribute to reduce latency and control burden in terms of circuit/flow switching [10], [11]. In present intra-datacenter networks, most traffic is East-West streams that stimulate the transition to flatter switching tier networks based on the fat-tree architecture. The control latency of the optical switch must be short, e.g. 100 μs, so that most traffic can be unloaded from the electrical domain [12].

Various optical switching technologies, the micro-electro-mechanical system (MEMS) and the semiconductor optical amplifier (SOA) switching system are being considered for intra-datacenter interconnections [4]–[6]. However, MEMS-based switches need optical-path adjustment at the fabrication stage, the complexity of which increases superlinearly in terms of the port count. In addition, the architecture creates a single point of failure. Furthermore, mechanical switches are slow, e.g. several milliseconds [13]. On the other hand, the SOA-based switch offers compact integrated devices with nanosecond-order switching speeds [5], [6], but its high power consumption limits the available port count.

The wavelength-routing (WR) switch based on an N×N cyclic arrayed-waveguide grating (AWG) can create low-power switching systems [14], since the AWG is a passive device. However, enlarging the cyclic-AWG scale triggers frequency deviation of the passband from the designated frequencies, i.e. ITU-T grid [14] and as a result the attainable port count of the WR switch is limited. To relax the impact of the frequency deviation, the uniform loss and cyclic-frequency (ULCF) AWG configuration was developed [7] and a prototype 64 × 64 AWG router has been realized by combining a 64 × 128 cyclic AWG and 64 1 × 2 optical couplers, where each pair of AWG output ports are bridged as one router output port so that the passband deviation is reduced. However, the attainable port count is much smaller than that considered necessary for intra-datacenter networking. Coordinated wavelength tuning to the passband of the AWG has been also considered [15]; however, it needs complicated mechanisms to control the laser diodes and increases the cost of tunable lasers.

To resolve these difficulties, we previously proposed an optical switch architecture that utilizes a combination of M × M wavelength-independent delivery-and-coupling (DC) switches and WR switches based on N × N cyclic AWGs constructed with multi-stage cyclic AWGs [8]. The combination can substantially enlarge the switch scale since...
the total port count becomes $MN \times MN$. We fabricated a 270 $\times$ 270 optical-switch prototype using $3 \times 3$ DC switches and 90 $\times$ 90 WR switches, and experimentally confirmed its effectiveness [16]. Further expansion of the switch scale needs enlargement of DC switch and/or WR switch scale. However, incrementing the port-count of a DC switch inevitably increases the optical coupler loss while for a WR switch it increases the passband frequency deviation resulting in excessive filtering loss. Applying erbium-doped fiber amplifiers (EDFAs) can compensate such losses; however, the solution necessitates a substantial number of costly EDFAs since there are no effective EDFA insertion points where multiple wavelength signals are aggregated.

In this paper, we propose a large-scale optical-switch architecture for intra-datacenter networks that makes use of cost-effective wavelength-aggregated amplification and ultra-dense wavelength routing. The proposed architecture comprises a DC-switch part, an aggregation-amplification part, and a WR-switch part. Thanks to the wavelength aggregation, one E DFA can simultaneously amplify multiple wavelength signals and hence the per-port EDFA cost is drastically reduced, a key benefit of the proposed architecture. Furthermore, ultra-dense wavelength routing is realized by introducing two-stage wavelength routing that combines an interleaver with steep skirt characteristics and a pair of AWGs with relaxed passband-center-frequency deviation. As a result, we can realize a cost-effective and large-scale optical circuit switch. We newly fabricate a pair of 1 $\times$ 90 AWGs on a monolithic planar-lightwave-circuit (PLC) chip, and develop control system for fast-tunable lasers [17]. The passbands conform to the ITU-T 25-GHz grid and the wavelength-tuning time is less than 436 $\mu$s over the full C-band. We construct part of a 1,440 $\times$ 1,440 optical switch by combining an 8 $\times$ 8 DC-switch part, a 180 $\times$ 1 wavelength-aggregated amplification part, and a 1 $\times$ 180 WR-switch part. Its good transmission characteristics and fast switching time are verified by transmission experiments. To the best of our knowledge, this is the first proof-of-concept demonstration of such a large-scale fast optical switch.

The organization of this paper is as follows: Section 2 details the proposed optical switch architecture. Section 3 introduces an optical circuit switch prototype fabricated with PLC technologies. The evaluations detailed in Sect. 4 confirm the effectiveness of the proposed switch architecture, where performance in both static and dynamic states is evaluated. Finally, this paper is concluded in Sect. 5. Note that a preliminary edition of this paper was presented in Opto Electronics and Communications Conference/International Conference on Photonics in Switching (OECC/PS) 2016.

2. Proposed Optical-Switch Architecture

Figure 1 shows the basic concept of our proposed optical-circuit switch; it combines $M \times M$ DC switches and $N \times N$ WR switches. In accordance with a combination of carrier wavelength of the wavelength-tunable transmitter and connection of the $M \times M$ DC switch, the signal can be transported to an arbitrary output port. It offers enlarged switch scale since the total switch scale is the product of $M$ and $N$. Furthermore, our scheme is tolerant against failures because a failed switching part can be replaced in a module-by-module manner.

Figure 2 shows the previously proposed $MN \times MN$ optical switch architecture [8] in which $M \times N$ WR switches constructed by two-stage cyclic AWGs are bridged by $N \times M$ DC switches, each of which comprises $M1 \times M$ Mach-Zehnder-interferometer (MZI) switches and $M \times M$ optical couplers. To expand the switch scale, $M$ and/or $N$ must be enlarged; however, the intrinsic loss of the DC switch and/or excess filtering loss of the WR switch would increase. Moreover, the filter bandwidth narrows as the signal passes through non-ideal filters twice and this limits the available bandwidth of the WR switch [18]. From these reasons, the available switch scale is rather limited. Introducing EDFAs can ease this problem, but a costly EDFA is needed at each port (see Fig. 2) since there is no signal aggregation point.

Figure 3 depicts a newly proposed high-port-count optical-switch configuration; it utilizes $N \times M$ DC switches, $M \times N$ optical couplers, $M$ EDFAs, $M$ interleavers, and $M$ pairs of $1 \times N/2$ non-cyclic AWGs, where $M$ is the degree of the DC switch, $N$ represents the port count of the coupler and the number of wavelengths, and the product of $M$ and $N$ equals overall switch scale. The opera-
Table 1  The numbers of necessary components for an \( MN \times MN \) optical switch. The numbers in parentheses are those evaluated here.

<table>
<thead>
<tr>
<th></th>
<th>( M \times M ) DC switch</th>
<th>( N \times 1 ) coupler</th>
<th>EDFA</th>
<th>Interleaver</th>
<th>Pair of ( 1 \times N/2 ) AWGs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number</td>
<td>( N (180) )</td>
<td>( M (8) )</td>
<td>( M (8) )</td>
<td>( M (8) )</td>
<td>( M (8) )</td>
</tr>
<tr>
<td>Per-port cost</td>
<td>( 1/M (1/8) )</td>
<td>( 1/N (1/180) )</td>
<td>( 1/N (1/180) )</td>
<td>( 1/N (1/180) )</td>
<td>( 1/N (1/180) )</td>
</tr>
</tbody>
</table>

Fig. 3  Newly proposed optical-switch architecture.

Newly proposed optical-switch architecture.

Fig. 4  Power diagram of a channel in the proposed optical-switch architecture.

The power required by the receiver and that restricted by quantum noise; generally, the former limitation is more stringent than the latter one. By adopting optical amplifiers, the optical power can be higher than the receiver-power requirement. Thus, introducing the EDFA allows a larger DC-switch scale. Moreover, we can pre-compensate the loss of the WR switch part which includes an interleaver and an AWG.

Table 1 summarizes the necessary number and the per-port cost contribution of each component for an \( MN \times MN \) switch. Each value in parentheses corresponds to the case of \( M = 8 \) and \( N = 180 \). The costly EDFA and interleaver are shared by \( N \) ports. Accordingly, the proposed architecture yields cost-effective optical switches that will suit cost-sensitive datacenter applications.

3. Prototype Fabrication

To verify the technical feasibility of the proposed optical-switch architecture, we monolithically fabricated interleaved \( 1 \times 100 \) AWGs with the PLC technology. The prototype employed an athermal structure to reduce power consumption. A pair of AWGs were jointly implemented on a single PLC chip of 36.4×44.0 mm\(^2\) (Fig. 5(a)), and were compactly contained in a module box of 120×70×7 mm\(^3\) (Fig. 5(b)). Each AWG can route up to 100 channels aligned on the 50-GHz grid, i.e., 196.225–191.275 THz for odd channels and 196.250–191.300 THz for even channels. The maximum frequency deviation from the grid is under 3.5 GHz as shown in Fig. 6, and it can be accepted in 50-GHz-grid systems; however, the signal quality is seriously degraded when the passband-frequency interval is 25 GHz, because guard bands between wavelength channels are much smaller. On the other hand, passband-frequency deviation of the interleaver we used is less than 0.8 GHz, which is accurate enough to de-interleave channels aligned on the 25-GHz grid.
Figure 5 illustrates passbands of the WR switch part, i.e., a combination of the interleaver and the AWG. Passbands for the center wavelengths of 1546.717 nm and 1546.917 nm are shown as examples. We can confirm that the passbands of the odd and even channels are interleaved with the 25-GHz offset. The extinction ratio of 25 dB was achieved, and hence the inter-channel crosstalk was well suppressed. As a result, our scheme enables ultra-dense wavelength routing with high port counts.

4. Experiments

To evaluate the transmission characteristics and switching time of the proposed switch architecture, we constructed part of a 1,440 × 1,440 optical switch by combining a fast-tunable laser, 8 × 8 DC switch, 180 × 1 coupler, EDFA, 1 × 2 interleaver, and pair of 1 × 90 AWGs. Here, a 256 × 1 coupler and 1 × 100 AWG were used as a 180 × 1 coupler and 1 × 90 AWG, respectively. We measured the bit-error-ratio (BER) characteristics in both static and dynamic wavelength states. Figure 8 shows the experimental configuration. The wavelength under test was generated by the previously developed fast-tunable laser [17].

Regarding the wavelength switching time, we measured 32,220 (180 × 179) combinations and the average and worst values were 348 µs and 436 µs, respectively, so shutter time was set to 498 µs including ~60 µs margin. The laser output was modulated at 10 Gbps by an intensity modulator. The wavelength signal was then input to an 8 × 8 DC switch with input power $P_{in}$. The insertion loss of the DC switch was 11.6 dB including 1/8 coupler intrinsic loss of 9 dB. The DC switch used an electro-optic effect switch and its switching time was around 200 ns, which is much faster than that of tunable lasers. As crosstalk sources, 180-wavelength signals on the 25-GHz ITU-T grid in the full C-band were generated using commercially available continuous-wave (CW) sources and another intensity modulator, where a wavelength-selective switch (WSS) based on liquid crystal on silicon (LCOS) eliminated the same wavelength as the target one and equalized the other-wavelength...
signal powers simultaneously. The target wavelength signal and the other wavelength signals in the full C-band were then aggregated by a 256 x 1 coupler in place of the 180 x 1 coupler with 25.7-dB loss. After all signals were amplified by an EDFA, an interleaver with 1.5-dB loss de-interleaved the signals into odd channels and even channels and a pair of 1 x 90 AWGs routed the 180-wavelength signals according to their wavelengths. Finally, the number of bit errors was counted with a BER tester having burst-mode operation.

First, we measured the BER characteristics of 180-wavelength channels in the static wavelength state. We set input power $P_{in}$ to 0 dBm; the input power was the minimum level that achieved BERs below the threshold of forward-error correction (FEC) using 7% overhead. Figure 9 shows the BER characteristics measured as a function of the WR-switch port number. The BER fluctuations are observed due to inequality of gain and noise figure of EDFA and that of port loss and extinction ratio of AWG; however, we confirm that BERs under the FEC threshold were obtained in all wavelength (180) channels when $P_{in}$ was 0 dBm; this input power can easily be attained with commercially available transmitters.

Next, we measured BER transitions induced by switching, where each BER was calculated using a 10^{-12}-length bit-sequence window. Input power $P_{in}$ was set to 0 dBm. Figure 10 plots measured dynamic BER transitions, where laser wavelength was changed between an edge and its adjacent channels (i.e. $\lambda_1$ and $\lambda_2$), edge and center channels (i.e. $\lambda_1$ and $\lambda_91$), and both edge channels (i.e. $\lambda_1$ and $\lambda_180$). During switching, BER was around 0.5 (i.e. $\log_{10}$(BER) $\sim 0.3$) since signal power was cut by a shutter to suppress crosstalk. In all cases, switching time was 498 $\mu$s as designed and BERs below the FEC limit were confirmed.

5. Conclusions

We proposed a novel optical-switch architecture that offers high port counts for intra-datacenter interconnection. The switch comprises DC switches and WR switches, each of which exploits a combination of an interleaver and AWGs to realize dense-wavelength routing cost-effectively. Based on the proposed architecture, we demonstrated part of a 1,440 x 1,440 optical switch by combining an 8 x 8 DC-switch part, 180 x 1 aggregation-amplification part, and 1 x 180 WR-switch part. Overall transmission performance was evaluated both in static and dynamic wavelength states. Switching time of less than 498 $\mu$s was attained thanks to the use of fast-tunable lasers. Our proposed switch offers high scalability in terms of hardware cost. The switching time is expected to be further reduced with subsequent research.

Acknowledgments

This work was supported by the Project for Developing Innovation Systems of the Ministry of Education, Culture, Sports, Science and Technology (MEXT) and New Energy and Industrial Technology Development Organization (NEDO), Japan.

References

Koh Ueda received the B.S. degree from Nagoya University in 2014. He is currently pursuing the Ph.D. degree in Department of Electrical Engineering and Computer Science at Nagoya University. His research interest includes optical network, especially in datacenter network. He is a student member of IEICE.

Yojiro Mori received the Ph.D. degree in engineering from the University of Tokyo, Japan, in 2013. He is currently an Assistant Professor at Nagoya University. Before joining the university, he was a Research Fellow of the Japan Society for the Promotion of Science from 2011 to 2012. In 2013, he joined the Department of Electrical Engineering and Computer Science at Nagoya University. His current research interests include digital coherent technologies and optoelectronic devices for photonic networks.

Hiroyuki Matsuura received the B.S. degree and M.S. degree in Electrical and Electronic Engineering, from Tokyo Institute of Technology, Tokyo, Japan, in 1995, 1997, and 2000, respectively. He is currently an associate professor at Nagoya University, Japan. From 2000 to 2005, he was an assistant professor in the Department of Communications and Integrated Systems, Tokyo Institute of Technology. His current research interests include development of network architectures and design/control network algorithms for future photonic networks. Dr. Hasegawa is a senior member of IEICE and a member of IEEE.

Hiroshi Hasegawa received his B.E., M.E., and D.E. degrees, all in electrical and electronic engineering, from Tokyo Institute of Technology, Tokyo, Japan, in 1995, 1997, and 2000, respectively. He is currently an associate professor at Nagoya University, Japan. From 2000 to 2005, he was an assistant professor in the Department of Communications and Integrated Systems, Tokyo Institute of Technology. His current research interests include development of network architectures and design/control network algorithms for future photonic networks. Dr. Hasegawa is a senior member of IEICE and a member of IEEE.

Kiyoshi Ishii is with Optical Network Technologies Group of AIST, Japan. She completed the doctor’s program in 2011 at Nagoya University, Japan. Her current research interests include optical network systems, energy efficient networks, photonic interconnects, and network resource optimization.
Haruhiko Kuwatsuka received the B.S. and Ph.D. degrees in Applied Physics from Tokyo University in 1985 and 1993, respectively. During 1985–2008, he was engaged in the research and development of optical semiconductor devices for optical communication systems. From 2008, he works on the research of network photonics devices in National Institute of Advanced Industrial Science and Technology (AIST). He is now Research Manager in Electronics and Photonics Research Institute in AIST.

Shu Namiki received M.S. and Dr. Sci. degrees in applied physics from Waseda University, Tokyo, Japan, in 1988 and 1998, respectively. From 1988 to 2005, he was at Furukawa Electric Co., Ltd., where he developed award-winning high-power pump lasers, and patented multi-wavelength-pumped-fiber Raman amplifiers. From 1994 to 1997, he was a Visiting Scientist at the Massachusetts Institute of Technology, Cambridge, where he studied mode-locked fiber lasers and ultra-short pulses in fiber. In 2005, he moved to the National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan, where he is currently the Director of Data Photonics Project Unit, and is also Chair of Executive Committee of a national project called “Vertically Integrated Center for Technologies of Optical Routing toward Ideal Energy Savings (VICTORIES)” in collaboration with ten telecom-related companies. He has authored or coauthored more than 450 conference presentations, papers, book chapters, articles, and patents. His current research interests include software defined dynamic optical path networking and their enabling devices such as nonlinear fiber-optics and silicon photonics. He was an Associate Editor and Advisory Editor of the journal Optics Express and the Co-Editor-in-Chief of the IEICE Transactions on Communications. He was on the technical committee for OFC, ECOC, CLEO, OECC, OAA, etc. and served as Program Co-Chair in OFC 2015, and will serve as General Co-Chair in OFC 2017. Dr. Namiki is a Fellow of the Optical Society of America and IEEE. He is a member of the Institute of Electronics, Information, and Communication Engineers, the Japan Society of Applied Physics.

Toshio Watanabe received the B.E., M.E., and Dr.Eng. degrees in organic and polymeric materials engineering from Tokyo Institute of Technology in 1990, 1992, and 2001, respectively. From 1992 to 2015, he was with NTT Corporation, where he engaged in research and development of polymeric optical waveguides, silica-based planar lightwave circuit (PLC) optical switches, photonic network systems, and metro-access WDM systems. He is currently an associate professor at Kagoshima University. He received the Microoptics Conference (MOC) Paper Award in 2011. Dr. Watanabe is a member of IEICE, IEEE, the Japan Society of Applied Physics, and the Physical Society of Japan.

Ken-ichi Sato (M’87–SM’95–F’99) received his B.S., M.S., and Ph.D. degrees in electronics engineering from the University of Tokyo, in 1976, 1978, and 1986, respectively. He is currently a Professor at the Graduate School of Engineering, Nagoya University, and he is an NTT R&D Fellow. Before joining the university in April 2004, he was an Executive Manager of the Photonic Transport Network Laboratory at NTT. He has been a leading researcher in the field of telecommunications; his most significant achievements lie in two of the very important transport network technology developments. One is ATM (Asynchronous Transfer Mode) network technology, which includes the invention of the Virtual Path concept. The other is photonic network technology, which includes the invention of the optical path concept and various networking and system technologies. His R&D activities cover transport network architectures, network design, photonic network systems including optical cross-connect/ADM and photonic IP routers, and optical transmission technologies. He has authored/co-authored more than 450 research publications in international journals and conferences. He holds 40 granted patents and more than 100 pending patents. His contributions to asynchronous transfer mode (ATM) and optical network technology development extend to coediting the IEEE Journal on Selected Areas in Communications (four special issues) and the Journal of Lightwave Technology (two special issues); organizing several workshops and conference technical sessions; serving on numerous committees of international conferences including OFC 2016 General Chair and OFC 2014 Program Chair; authoring a book, Advances in Transport Network Technologies (Artech House); and coauthoring 14 other books. Prof. Sato is the President and a Fellow of the Institute of Electronics, Information and Communication Engineers (IEICE) of Japan and a Fellow of the IEEE. He received the Young Engineer Award in 1984, the Excellent Paper Award in 1991, the Achievement Award in 2000, and the Distinguished Achievement and Contributions Award in 2011 from the IEICE of Japan, and the Best Paper Awards in 2007 and 2008 from the IEICE Communications Society. He was also the recipient of the Distinguished Achievement Award of the Ministry of Education, Science and Culture in 2002, and the Medal of Honor with Purple Ribbon from Japan’s Cabinet Office in 2014.