European Conference on Optical Communication (ECOC) 2022 © Optica Publishing Group 2022

12.8 Tb/s SDM Optical Interconnect for a Spine-Leaf Datacenter Network with Spatial Channel Connectivity

Ruben S. Luis[1], Benjamin J. Puttnam[1], Georg Rademacher[1], Satoshi Shinada[1], Tetsuya Hayashi[2], Tetsuya Nakanishi[2], Yuki Saito[2], Tetsu Morishima[2], Hideaki Furukawa[1]

(1) National Institute of Information and Comm. Technology, Koganei, Japan, rluis@nict.go.jp
(2) Optical Communications Laboratory, Sumitomo Electric Industries Ltd., Yokohama, Japan,

Abstract *We experimentally demonstrate 12.8 Tb/s optical-interconnects using an 8-core or two 4-core multicore fibers with 64×200 Gb/s PAM-4 lanes implementing SDM spine-leaf datacenter network topologies. We evaluate a conventional topology with 12.8 Tb/s interconnects and the use of low-loss optical cross-connects for spatial channel connectivity. ©2022 The Author(s)*

Introduction

Spatial-division multiplexing (SDM) optical interconnects (OICs) have been widely proposed to support the exponential growth of intra-datacenter network traffic[1],[2] as well as the introduction of high-density co-packaged network interfaces[3]. In particular, the use of uncoupled core multi-core fibers (MCFs) has been targeted, given the possibility of all optical spatial demultiplexing, negligible inter-core crosstalk and low inter-core skew[4]. Fig. 1 shows a few of the SDM-OIC throughputs demonstrated in recent experiments[5]–[10]. The highest throughput reported thus far being 3.8 Tb/s using a 4-λLAN-WDM 4-SDM system with 85 GBaud 8-ary pulse amplitude modulation (PAM) signals[9]. However, the current traffic growth forecasts predict a need for interfaces with more than 10 Tb/s within a few years[2], which is yet to be achieved.

Here, we propose and demonstrate a new SDM-OIC with 12.8 Tb/s throughput (8-λLAN-WDM 8-SDM) implemented with an 8-core MCF with standard cladding diameter and 8 WDM 200 Gb/s lanes per core using 112 GBaud PAM4 signals. We further demonstrate the use of low loss optical cross-connects (OXCs) to introduce spatial channel granularity for the implementation of a high-capacity spine-leaf network. Fig. 2 shows the considered network scenarios A (12.8 Tb/s spine-leaf) and B (spine-leaf with flexible OXC network architecture). Scenario A provides the highest reported SDM-OIC throughput using intensity-modulated signals, improving the previous record by a factor of more than 3. Scenario B allows for multiples of 1.6 Tb/s throughput (8×200 Gb/s WDM lanes) interconnections between spine and leaf switches with a single SDM-OIC connected to each spine switch to serve up to 8 leaf switches. This is a dramatic reduction of the number of required SDM-OICs with respect to standard OICs whilst maintaining the predictable 2-hop loss, skew and latency characteristic of the spine-leaf topology[4]. The network architecture is ensured by a low-loss MEMS-based OXC, which can flexibly establish spatial channel connections with minimal performance impact. This approach exploits the benefits of SDM systems in terms of capacity, resource sharing, integration potential and switching. These results show the potential of high-capacity SDM-OICs and SDM networks to support the expected growth and scalability of future intra-datacenter networks.

Experimental Demonstration

Fig. 3-a) and -b) show the experimental setups used for scenarios A and B, respectively. The SDM/WDM transmitter produced 8 spatial channels, each carrying 8 O-band carriers with 112 GBaud PAM4 signals for a total of 64×200 Gb/s SDM/WDM lanes, assuming hard-decision forward-error correction (HD-FEC) with a 12% overhead[11]. One of the spatial channels, referred to as test spatial channel, was produced by modulating the 8 lightwaves from 7 fixed wavelength distributed feedback lasers (DFBs), selected from a set of 8 possible lasers, and one tunable external

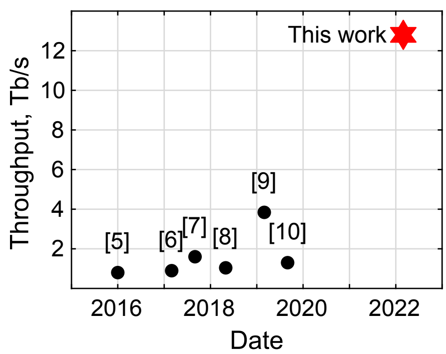

Fig. 1: Recent throughput achievements with SDM-OIC.

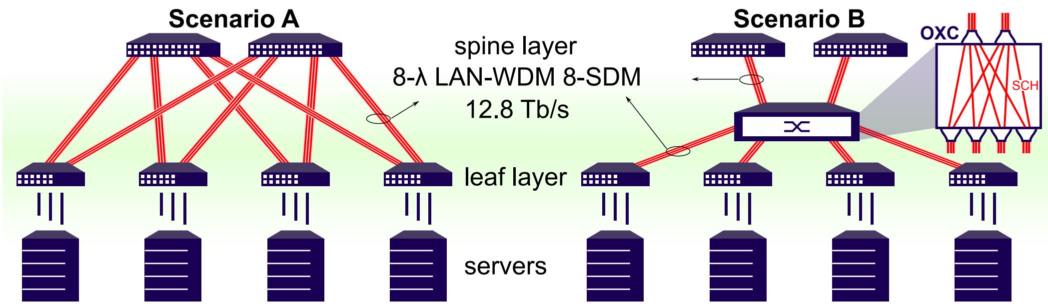

Fig. 2: Considered SDM datacenter network scenarios: Scenario A - spine-leaf topology and; Scenario B - spine-leaf with OXC for spatial channel network connectivity.

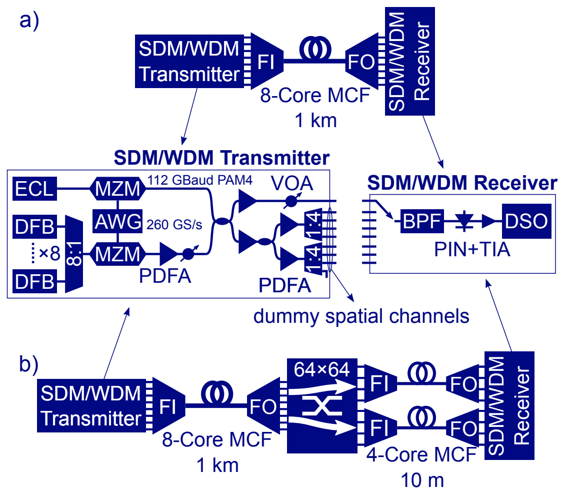

Fig. 3: Experimental setup for scenarios A (a) and B (b). The insets show (A) the transmission spectrum, the profiles of the (B) 8-core and (C) 4-core MCFs.

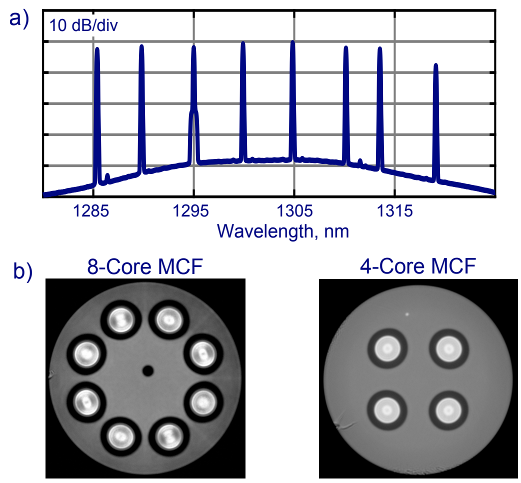

Fig. 4: Spectrum of the WDM signal transmitted on the test spatial channel (a) and profiles of the 8- and 4-core MCFs (b).

cavity laser (ECL), used as a test channel. This channel was independently modulated in a Mach-Zehnder modulator driven by a 260 GS/s arbitrary waveform generator with PAM4 root-raised cosine signals with a 0.01 roll off. A linear pre-equalization filter was used to compensate for the system frequency limitations in the digital domain. The remaining channels were jointly modulated in a single MZM with the same format. Test and dummy channels were combined and amplified with praseodymium doped fiber amplifiers (PDFAs) to compensate the loss of the power splitters and combiners, forming a test spatial channel, with the spectrum shown in Fig. 4-a). A sample of this signal was amplified and split to form 7 replicas as dummy spatial channels. It should be noted that the PDFAs in the transmitter were required to compensate for the losses of the power splitters but would not be required in a full implementation.

For scenario A, we used a 1 km, 8-core, homogeneous single-mode MCF designed to operate in the O-band[5]. The 8-cores were positioned in a ring configuration within a 125 µm diameter cladding, as shown in Fig. 4-b). Each core had a trench assisted index profile to mitigate crosstalk and transmission loss ranging from 0.34 dB/km to 0.41 dB/km in the O-band. Fused-fiber-based fan in-fan out (FIFO) devices were used to couple signals into and out of the fiber. The combined FIFO and fiber loss was between 2.5 dB and 4.5 dB and the crosstalk was lower than -35 dB. This was mainly introduced by the FIFO devices and had no discernible performance impact.

At the receiver side, we manually selected the test spatial channel after transmission and used a tunable band pass filter (BPF) to select the test wavelength. This replaced the WDM demultiplexer that would be present in a full implementation. The selected signal was detected by a single-ended 50 GHz p-i-n photodetector followed by a 66 GHz transimpedance amplifier (TIA). The electrical output was digitized by a 110 GHz real-time digital sampling oscilloscope operating at 256 GS/s for offline processing. The digital signal processor (DSP) consisted of a resampling stage to 2 samples per symbol, followed by a re-timing stage and a 17-tap feed-forward equalizer. The performance of each spatial and wavelength channel was evaluated by direct error counting on five 10 µs traces with a total of 11.2×10^{6} bits. The throughput per channel was 224 Gb/s before HD-FEC. The selected code had a bit-error rate

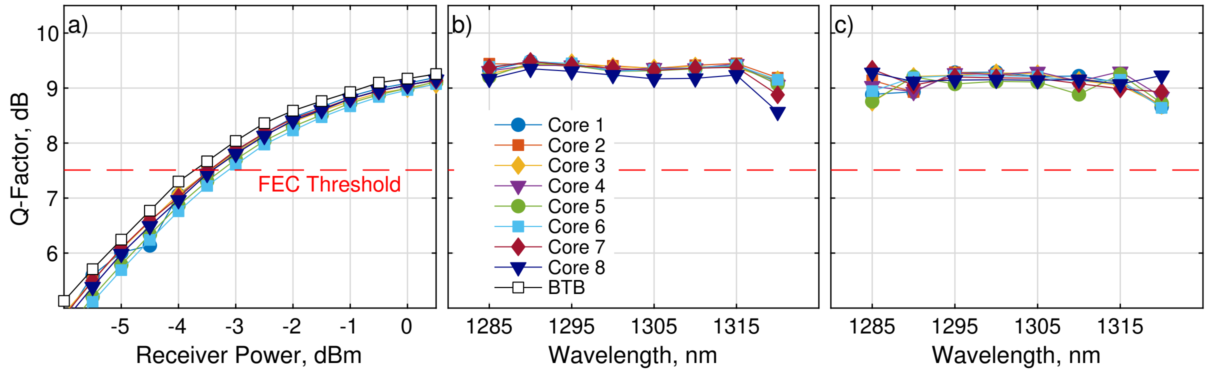

Fig. 5: Network performance. a) Q-factor dependence on the receiver power in back-to-back and scenario A; b) and c) Wavelength dependence of the Q-factors in scenario A and B, respectively.

(BER) threshold of 8.8×10^{-3}, which corresponded to a Q-factor of 7.52 dB[11]. After HD-FEC, the combined net throughput of the 64 lanes of the transmitted signal was 12.8 Tb/s.

Fig. 3-b) shows the experimental setup used for scenario B. In this case, the signals were transmitted through the previously described OIC and directed to an OXC. The latter was implemented using a bidirectional 64×64 3D-MEMS-based optical switch. This link was equivalent to a connection from the spine to the OXC, as shown in Fig. 2. The OXC was a low loss device but included monitoring ports and VOAs for power control. For this reason, the average loss of the connections was approximately 1 dB. The signals at the OXC input were directed to 8 outputs, to be connected to the leaf. In the absence of a second 8-core MCF, we used two 4-core fibers[12] in parallel to emulate the connection from the OXC to the leaves shown in Fig. 2. The 4-core fibers were 10 m long and were connected to free-space FIFO devices. Similarly to the 8-core fiber, the 4-core fibers also had standard cladding diameter. The cores were arranged in a square disposition, as shown in Fig. 4-b). The total loss of this connection was approximately 2 dB due to the FIFO devices and the crosstalk was below -60 dB. We used the same transmitter and receiver setup as for scenario A.

Fig. 5-a) shows the estimated Q-factor dependence on the power at the photodetector input in back-to-back and after transmission in scenario A for each core. The sensitivity after transmission was below -3 dBm with a maximum transmission penalty of 0.7 dB with respect to back-to-back. Fig. 5-b) and -c) show the wavelength dependence of the Q-factor after transmission through scenario A and B, respectively, with the maximum launch power. In the case of scenario A, we observed some wavelength dependent penalty due to limitations of the BPF at the receiver side, particularly at 1320 nm wavelength, suggesting that use of a CWDM demultiplexer could achieve a better performance. Despite these limitations, the performance was well above the required FEC threshold with an average Q-factor of 9.3 dB. In the case of the performance of scenario B, the additional loss of the OXC, MCF and FIFOs was well within the power budget afforded by the system. As such, the wavelength dependence of the Q-factor after transmission through scenario B shown in Fig. 5-c) depicts little degradation with respect to scenario A. The average Q-factor was 9 dB, with a 0.3 dB penalty with respect to scenario A. This resulted from the OXC whilst still allowing improved flexibility of the spatial channel network architecture without impact on throughput.

Conclusions

We proposed and demonstrated a record throughput optical interconnect with 12.8 Tb/s (8-λLAN-WDM 8-SDM) operating in an SDM spine-leaf architecture. Our approach exploited both the high-capacity and the flexibility of SDM systems and networks adopting spatial channel switching. We used an 8-core multicore fiber with standard cladding diameter, which was compatible with standard fiber plant and cabling to accommodate 64×200 Gb/s lanes implemented using 112 GBaud PAM4 signals. We evaluated two network scenarios assuming a conventional spine leaf topology with 12.8 Tb/s interconnects; and a spine leaf topology implemented with an optical cross-connect to provide a flexible network architecture with spatial channel connectivity. These results show that the combination of high capacity SDM systems and spatial switching can provide the flexibility and throughput requirements of future datacenters.

References

[1] B. J. Puttnam *et al.*, "Space-division multiplexing for optical fiber communications", *Optica*, vol. 8, no. 9, pp. 1186–1203, Sep. 2021.

[2] P. J. Winzer and D. T. Neilson, "From scaling disparities to integrated parallelism: A decathlon for a decade", *J. of Lightw. Technol.*, vol. 35, no. 5, pp. 1099–1115, Mar. 2017.

[3] H. Dong and J. D. Downie, "Applications for spatial divison multiplexing fiber and associated cost per bit", in *2020 IEEE Photon. Soc. Summer Topicals Meeting Series*, Jul. 2020, pp. 1–2.

[4] L. Zhang *et al.*, "Enabling technologies for optical data center networks: Spatial division multiplexing", *J. of Lightw. Technol.*, vol. 38, no. 1, pp. 18–30, Jan. 2020.

[5] T. Hayashi *et al.*, "125-µm-Cladding eight-core multicore fiber realizing ultra-high-density cable suitable for o-band short-reach optical interconnects", *J. of Lightw. Technol.*, vol. 34, no. 1, pp. 85–92, Jan. 2016.

[6] S. Beppu *et al.*, "56-Gbaud PAM4 Transmission over 2-km 125-µm-cladding 4-core multicore fibre for data centre communications", in *2017 European Conference on Optical Communication (ECOC)*, Sep. 2017, pp. 1–3.

[7] P. De Heyn *et al.*, "Ultra-dense 16x56Gb/s NRZ GeSi EAM-PD arrays coupled to multicore fiber for short-reach 896Gb/s optical links", in *Optical Fiber Communication Conference*, OSA, 2017, Th1B.7.

[8] O. Ozolins *et al.*, "7x149 Gbit/s PAM4 Transmission over 1 km multicore fiber for short-reach optical interconnects", in *2018 Conference on Lasers and Electro-Optics (CLEO)*, May 2018, SM4C.4.

[9] A. Masuda *et al.*, "255-Gbps PAM-8 Transmission under 20-GHz bandwidth limitation using NL-MLSE based on Volterra filter", in *Optical Fiber Commun. Conf. (OFC) 2019*, San Diego, California, 2019, W4I.6.

[10] R. S. Luis *et al.*, "Dynamic crosstalk and skew on a 1.3 Tb/s full-duplex O-band short reach transmission using an 8-core fiber", in *45th European Conference on Optical Communication (ECOC 2019)*, Sep. 2019, Tu.1.A.3.

[11] G. Tzimpragos *et al.*, "A survey on FEC codes for 100 G and beyond optical networks", *Commun. Surveys Tutorials*, vol. 18, no. 1, pp. 209–221, 2016. DOI: 10.1109/COMST.2014.2361754.

[12] T. Matsui *et al.*, "Design of 125 µm cladding multi-core fiber with full-band compatibility to conventional single-mode fiber", in *2015 European Conference on Optical Communication (ECOC)*, Sep. 2015, We.1.4.5.