

Equalizer State Caching for Fast Data Recovery in Optically-Switched Data Center Networks

Zhouyi Hu, *Member, IEEE*, Zichuan Zhou, *Student Member, IEEE*, Chun-Kit Chan, *Senior Member, IEEE*, and Zhixin Liu, *Senior Member, IEEE*

Abstract—Optical switching offers the potential to significantly scale the capacity of data center networks (DCN) with a simultaneous reduction in switching time and power consumption. Previous research has shown that end-to-end switching time, which is the sum of the switch configuration time and the clock and data recovery (CDR) locking time, should be kept within a few nanoseconds for high network throughput. This challenge of low switching time has motivated research into fast optical switches, ultra-fast clock and amplitude recovery techniques. Concurrently, the data rate between server-to-server and server-to-switch interconnect is increasing drastically from the current 100 Gb/s (4×25 Gb/s) to 400 Gb/s and beyond, motivating the use of high order formats such as 50-GBaud four-level pulse-amplitude modulation (PAM-4) for signalling. Since PAM-4 is more sensitive to noise and distortion, digital equalizers are generally needed to compensate for impairments such as transceiver frequency roll-off, dispersion and optical filtering, adding additional time for equalizer adaptation and power consumption that are undesired for fast optical switching systems. Here we propose and investigate an equalizer state caching technique that reduces equalizer adaptation time and computation power consumption for fast optical switching systems, underpinning optically-switched DCNs using high baud rate and impairment-sensitive formats. Through a proof-of-concept experiment, we study the performance of the proposed equalizer state caching scheme in a three-node optical switching system using 56 GBaud PAM-4. Our experimental results show that the proposed scheme can tolerate up to ± 0.8 -nm (± 100 -GHz) instantaneous wavelength change with an adaptation delay of only 0.36 ns. Practical considerations such as clock phase misalignment, temperature-induced wavelength drift, and equalizer precision are also studied.

Index Terms—Clock and data recovery, data center networks, digital equalization, fast optical switching, PAM-4.

I. INTRODUCTION

OPTICAL switching is gaining significant interest in data center networks (DCNs) because of its potential to overcome the bandwidth and port density limits in conventional electronic switches [1]. To ensure high network throughput for small packets (e.g., 64 to 576 bytes) dominated DCN [1, 2], fast

optical switches (e.g., configuration in 1 to 2 nanoseconds) and sub-nanosecond clock and data recovery (CDR) are required [1]. These building blocks are being actively studied, demonstrating sub-nanosecond switching time in an integrated optical switch [3] and sub-nanosecond CDR using clock synchronization and clock phase caching techniques [1].

The relentless growth of machine-to-machine data traffic has also stimulated the rapid increase of the interconnection data rate. It is expected that the point-to-point data rate will increase from the current 100 Gb/s, which consists of four transceivers operating with 25 Gb/s on-off-keying, to 400 Gb/s and beyond, requiring high baud rate and high order modulation format [4]. For example, the IEEE 802.3bs task force recently adopted four-level pulse-amplitude modulation (PAM-4) as the appropriate modulation format [5]. Many demonstrations beyond 50 GBaud PAM-4 signalling have also been reported [6-11]. However, high baud rates and high order modulation formats have reduced tolerance to noise and signal distortions. Consequently, adaptive digital equalizers such as feed-forward equalizers (FFE) and decision feedback equalizers (DFEs) are generally required to mitigate the transmission impairments, minimizing the pre-forward error correction (FEC) bit error rate (BER) so that low complexity FEC codes can be used for low power consumption and low processing latency [12, 13].

Using a receiver-side digital equalizer introduces additional complexity and equalizer adaptation time that did not exist in previously demonstrated on-off-keying (OOK) based systems [1, 3, 14, 15]. For example, standard FFEs/DFEs using least means square (LMS) algorithms require 1000s of symbols to train their tap coefficients, resulting in 100s of nanoseconds data recovery time [1]. Furthermore, the equalizer state (e.g., tap coefficients in FFEs/DFEs) must be reacquired when switching between different transceiver pairs to compensate for variation of transmission impairments due to, for instance, wavelength drift of the lasers and multiplexers/demultiplexers, as well as different fiber lengths and transceiver bandwidths. Therefore, achieving low-adaptation-time digital equalization for short optical packets is a challenge that must be addressed.

The receiver-side equalization requirements for 25-Gb/s

This work was supported in part by in part by EPSRC programme EP/R035342/1, EP/R041792/1 and Royal Society PIFR1\180001. (Corresponding author: Zhouyi Hu.)

Zichuan Zhou and Zhixin Liu are with the Department of Electronic and Electrical Engineering, University College London, London WC1E 6BT, U.K (e-mail: zichuan.zhou.14@ucl.ac.uk; zhixin.liu@ucl.ac.uk).

Zhouyi Hu and Chun-Kit Chan are with the Department of Information Engineering, The Chinese University of Hong Kong, Hong Kong (e-mail: hz016@ie.cuhk.edu.hk; ckchan@ie.cuhk.edu.hk).

Zhouyi Hu is now with Aston Institute of Photonic Technologies, Aston University, Birmingham, B4 7ET, United Kingdom (e-mail: z.hu6@aston.ac.uk).

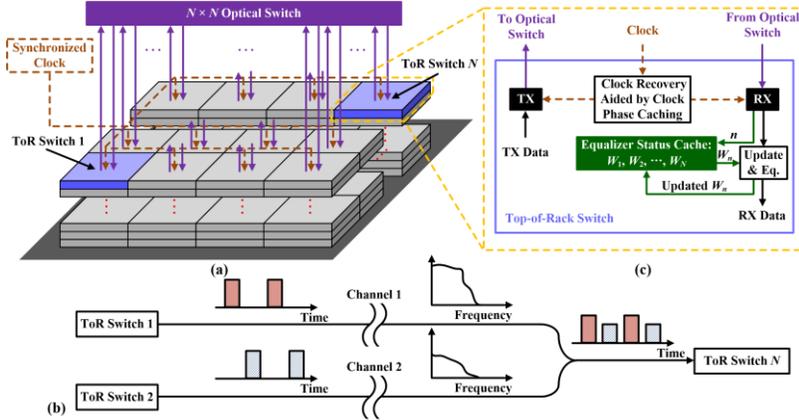


Fig. 1. (a) Architecture of an optically-switched DCN. (b) Illustration of the impairments through different transmission links. (c) Concept of the proposed equalizer state caching method (ToR: Top-of-Rack; TX: transmitter; RX: receiver; Eq.: equalization).

PAM-4 burst-mode signalling were discussed in [16], which indicate that the equalizers can work in burst mode with about 200 symbols for training. However, this is still too slow to be employed in optically switched DCN. The challenges and design considerations of burst-mode equalizers are discussed in [17], showing less than 2-dB power penalty with about 200 training symbols using the gear-shifted LMS algorithm. The recursive least squares (RLS) algorithm has been proposed to achieve coefficient convergence over 10 times faster than the standard LMS algorithm. Unfortunately, it also significantly increases the computation complexity, imposing a significant challenge for potential real-time implementation [17, 18]. In [19], the authors designed a 25-GBaud PAM-4 burst-mode receiver with analog gain and offset control with 250 nm SiGe BiCMOS. However, the preamble required for reconfiguration needed to take several tens of nanoseconds. Fast equalization with sub-nanosecond adaptation time has not yet been achieved.

In this paper, we propose and demonstrate equalizer state caching for low-adaptation-delay digital equalization for short packets (256 bytes) using 56-GBaud PAM-4. In this method, the optimal equalizer state (or more specifically, tap coefficients of the equalizer) that corresponds to certain transceiver pairs is stored in respective transceivers and is regularly updated, ensuring that the equalizer state is close to the optimum for the communicating transceivers. This is analogous to the concept of clock phase and amplitude caching that are previously reported in [1, 20, 21], Equalizer state caching achieves low overhead, low adaptation delay and low DSP complexity for optically switched DCN.

The remainder of this paper is organized as follows: Section II introduces the architecture of an optically-switched DCN and the concept of our proposed equalizer state caching method. Section III shows our three-node prototype and our proof-of-concept experiment that verifies the feasibility of our proposed method. Section IV shows our results and performance analysis. The robustness of the methods is studied with respect

to different operational conditions including wavelength drift, clock phase mismatch, and equalizer coefficient precision. Finally, Section V concludes the paper.

II. OPERATIONAL PRINCIPLE

A. Architecture of optically-switched data center networks

Fig. 1(a) shows an example architecture of an optically-switched DCN comprising N racks, where each rack consists of up to 64 servers that connect to the top-of-rack (ToR) switch [1]. The ToRs are connected to a non-blocking $N \times N$ optical switch to facilitate switching in a synchronized manner. A distributed synchronized clock is disseminated to all ToRs and the transceivers through the control plane to enable clock synchronized networking (brown dashed lines). Note that this example network architecture is the same as in [1], where clock synchronization and clock phase caching facilitate sub-nanosecond clock recovery.

Although this architecture is suitable for any fast switches (e.g., fast tunable lasers [20], semiconductor optical amplifier-based switch [22], Mach-Zehnder arrays [23]), we focus this paper on tunable laser-based switching [24-26], where the $N \times N$ switch fabrics are passive elements such as array waveguide grating router (AWGR). Since every ToR switch needs to exchange packets with all other switches, the impairments through different transmission links are different due to wavelength drift, dispersion at different wavelengths, end-to-end loss and different fiber lengths, as conceptually shown in Fig. 1(b).

As reported in [27], the operating temperature range in a practical data center environment could vary from 18 to 27 °C, which results in a change of signal propagation delay (or time-of-flight) and a change of clock phase in clock synchronized DCN. In addition to the change of the time-of-flight, temperature variation also results in a change of operation wavelength of the lasers and the passband of the passive AWGR core, and consequently leading to a change of the

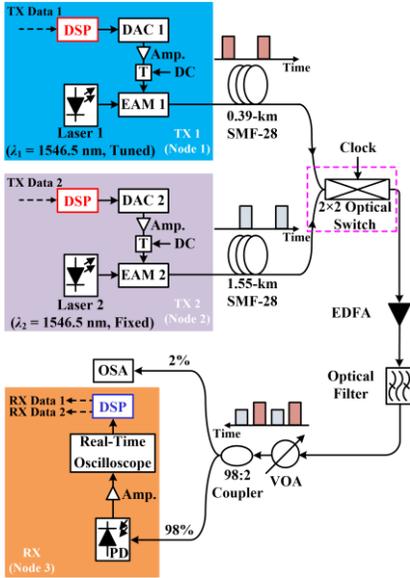


Fig. 2. Proof-of-concept experiment for verifying the proposed 56-Gbaud PAM-4 burst transceiver in optically-switched DCN.

channel responses. For example, some wavelengths may be located at the center of an AWGR passband while some may be offset from the center. Typical values of the center frequency drift due to temperature variation are 1 GHz/°C for silica AWGRs and 10 GHz/°C for silicon-on-insulator AWGRs [4, 28]. In a controlled environment in data centers, this offset could be 10s of GHz as the temperature varies. Although active wavelength locking can be used to lock the laser wavelength to the center of the passband, additional feedback control would be needed, adding cost and power consumption [29]. Although the change of channel response is small and slow, it is necessary to enable adaptive equalization to ensure stable long-term operation of optically-switched DCN.

B. Equalizer state caching

The main contribution of this work is shown in the block diagrams on the right (Fig. 1(c), highlighted in green), in which equalizer state caching is described. The signal processing flow of the equalizer state caching is described as follows: The burst-mode receiver is in `idle` mode until the start of a burst. The burst-mode receiver is then reconfigured for the reception of the payload from corresponding network nodes. As reported in [16, 19, 20, 30], the operation of reconfiguration via preambles may take several tens of ns. Meanwhile, fiber links in a DCN are relatively static. Therefore, the utilization of state caches with a slow update rate becomes an attractive solution to reduce the reconfiguration time [1, 3, 20]. Sub-nanosecond CDR in optically-switched DCN can be implemented using clock phase caching [1]. When the receiver is able to cache equalizer coefficients, fast burst-mode channel equalization can also be realized [3, 20].

As shown in Fig.1 ©, clock is first recovered using clock phase caching when a packet arrives. Since each network node in an N -rack DCN stores all equalizer state caches of the other racks, i.e., $\{W_1, W_2, \dots, W_N\}$, the corresponding tap coefficients of a transceiver pair, for example, W_n for the packets from the n -th ToR, are pre-trained and stored in transceivers for signal detection digital signal processing (DSP). The cached equalizer states can be updated via preambles or other blind adaptive algorithms to track channel impairment changes. After burst-mode reconfiguration, an equalizer leveraging the equalizer coefficients for packets from the n -th node is used to mitigate signal impairments.

III. PROOF-OF-CONCEPT EXPERIMENT

A. Experimental setup

Fig. 2 shows our proof-of-concept experimental setup. Our three-node prototype contained two independent transmitters that sent optical packets to the same receiver node. The two sets of 56-Gbaud PAM-4 data packets were generated offline and re-sampled to 92-GSa/s to generate the electrical waveforms using two digital-to-analog converters (DAC) with 8 bits resolution and approximately 5.5 bits effective number of bits (ENOB). The amplified electrical signals, combined with a DC voltage through bias tees, drove two electro-absorption modulators (EAMs, about 40 GHz bandwidth) to generate optical bursts. The wavelengths (λ_1 and λ_2) of the two transmitters (Tx 1 and Tx 2) were both set to 1546.5 nm initially. Laser 1 can be thermally tuned to investigate the impact of wavelength drift. The extinction ratios of the modulated signals were about 12 dB and 10 dB, respectively. After modulation, the optical powers of two optical packets were -4.4 dBm and -5.2 dBm for Tx1 and Tx2, respectively. The modulated optical bursts were launched into two standard single-mode fiber (SMF-28) links with lengths of 0.39 km (0.6-dB link loss) and 1.55 km (1-dB link loss) before being switched by a 2x2 LiNbO₃ Mach-Zehnder (MZ) optical switch (about 20-dB extinction ratio) to enable fast optical switching. Due to the relatively high insertion loss of the MZ switch (6 dB), an erbium-doped fiber amplifier (EDFA) was used to amplify the multiplexed optical packets to 16 dBm before passing through a 2-nm-bandwidth optical filter, which emulated passive switching elements (such as AWGRs). The optical signal-to-noise ratio (OSNR) of the amplified packets was about 45 dB. At the receiver side, a variable optical attenuator (VOA) was used to vary the received optical power for receiver sensitivity measurement. A 50-GHz photodiode (PD) followed by a 45-GHz bandwidth RF amplifier was used to amplify the received signal. The analog signals were captured by a 160-GSa/s analog-to-digital converter (8 bits resolution with about 5.5 bits ENOB) before being processed by subsequent offline DSP, implemented in MATLAB using double-precision floating-point arithmetic. The impact of the bit precision of equalizer coefficient caching on its performance will be discussed later in Section IV E.

Commented [LZ1]: What do you mean 'in idle'? Not clear what idle mode means here.

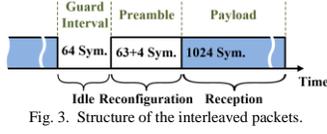


Fig. 3. Structure of the interleaved packets.

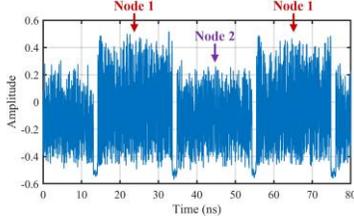


Fig. 4. Received time-domain waveform.

B. Packet structure

Fig. 3 depicts the optical packet structure. Each packet consisted of 67 preamble symbols and a 1024 symbol (i.e., 256 byte) payload modulated with PAM-4. The preamble contained a Gold sequence of 63 symbols for frame synchronization and packet identification. The first four symbols were used as a label to indicate the start of a burst. An additional training sequence could be inserted into the packets if needed (see Section IV for discussion). The total length of a packet was 1091 (=63+4+1024) symbols, corresponding to 19.5 ns when operating at 56 GBd (the preamble occupies about 1.2 ns). A 64 symbol (about 1.1 ns) guard interval was scheduled between consecutive packets.

Fig. 4 shows the received time-interleaved packets from both nodes. Due to different transmitter output power and optical path loss, the received packets had different amplitudes. Conventionally, burst mode automatic gain control (AGC) circuit is used to scale the signal amplitude for signal detection and processing. Using equalizer state caching, the amplitude scaling factor is included in the equalizer coefficients. Thus, costly burst mode automatic gain and offset control are no longer needed. Note that if only one tap coefficient is used, equalizer state caching becomes ‘amplitude caching’ [21].

C. Burst-mode DSP with equalizer state caching

Fig. 5(a) and Fig. 5(b) show the DSP used at the transmitter side and the receiver side with the equalizer state cache highlighted in dark green. At the transmitter sides, 1024 symbol Gray-coded PAM-4 payloads and 67 preamble symbols were generated and re-sampled to 92 GSa/s to form root-raised-cosine (RRC, a roll-off factor of 0.08) shaped packets at 56 GBd. The signals were then clipped using the hard clipping method [31] with a clipping ratio of 2.2 and were subsequently loaded to the DACs to generate waveforms.

At the receiver side, the digitized packets were filtered by a matched RRC filter. The filtered signals were re-sampled to 2.5 Sa/symbol for subsequent burst-mode detection, including start-of-burst detection and packet identification, which were realized by cross-correlating the received signal with the known

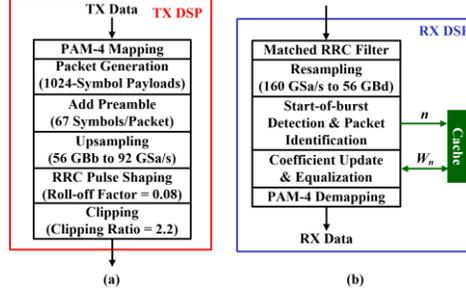


Fig. 5. Block diagrams of (a) the transmitter-side DSP, and (b) the receiver-side DSP.

Gold sequences. After identifying the transmitter ($n = 1$ or 2 in this work) of the packet, the equalizer read in the corresponding equalizer coefficients that were pre-trained and stored in the memory, facilitating fast equalization. As the packet was being processed, the equalizer coefficients were updated blindly using a decision-directed LMS algorithm [32]. As we focus on the performance of the equalizer state caching scheme, we used a 31-tap FFE, which can be effectively implemented in a field-programmable gate array (FPGA) platform. Any amplitude change (for example due to temperature variation or wavelength change) is accounted for by the equalizer coefficients. The BER performance of the system was studied using offline DSP, calculated by processing 480 packets, i.e., 0.49 million PAM-4 symbols or 0.98 million bits.

D. Emulation of wavelength drift dependent impairment

When using uncooled lasers, which are advantageous to their decreased power consumption and cost versus cooled lasers, temperature variation introduces laser wavelength variation of approximately 0.1 nm/°C, resulting in a wavelength offset from the center of a wavelength multiplexer/demultiplexer. This leads to different link loss, dispersion and signal spectra which require equalizer state changes to maximize link performance. To study the impact of laser wavelength variation, we manually tuned the operational wavelength of the lasers by controlling the current of their thermoelectric cooler (TEC). By tuning the laser temperature from 10 °C to 34 °C with a 4 °C interval, we created wavelength misalignment from the center of the optical filter from -1.2 nm to +1.2 nm.

In Fig. 6(a), the solid blue line shows the optical spectra when laser 2 operates at 22 °C and laser 1 at 34 °C, respectively. The dotted brown line shows the power profile of the 2-nm bandwidth optical filter. By heating laser 1 to 34 °C, the optical filter introduces an additional 4 dB of loss and the edge filtering causes the asymmetric optical spectrum.

Fig. 6(b) shows the tap coefficient of an optimized equalizer for different wavelength offset values. The different equalizer states result from different optical power, dispersion and optical filtering impairments. This indicates that the cached equalizer states need to be updated regularly to compensate for different impairments due to temperature variation.

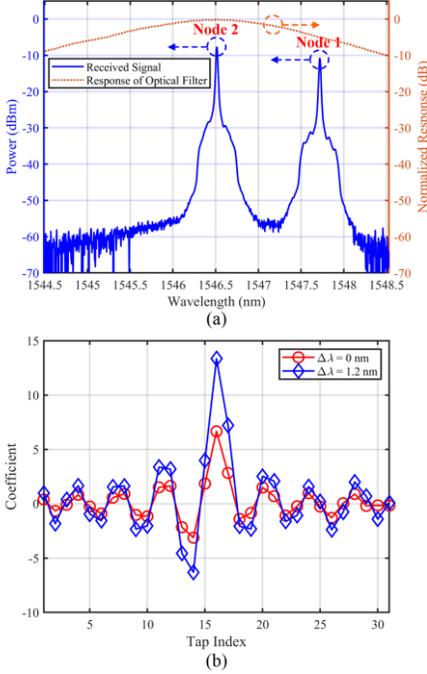


Fig. 6 (a) Illustration of wavelength drift during burst-mode transmission. (b) Equalizer coefficients of Node 1 before and after wavelength drift.

IV. RESULTS AND DISCUSSIONS

A. BER performance

Fig. 7 (a) and (b) show the measured BER results of the detected packets from Node 1 and Node 2, respectively, when both laser diodes are operating at 1546.5 nm. The optical signal power in the x-axis is the average power of the optical packets, which are calculated from the measured optical power and the duty cycle of the packets. For example, for Node 1 only transmission, the optical power is scaled up by 3.26 dB from the measured average power, correcting the fact that no signal is transmitted for 53% of the time. The FFE contains 31 taps and is trained using 10,000 symbols (i.e., about 10 packets) with LMS algorithm to obtain the optimum state (coefficients). In signal detection, the equalizer is always adaptively updated using the decision-direct method. The cached state sets the initial condition of the equalizer, which is crucial to minimizing the equalizer convergence time in optically switched data center networks.

In both Fig.7(a) and Fig.7(b), the red asterisk and brown square markers represent the BERs of the interleaved packets, transmitted over back-to-back (BtB) and the fiber link (0.39 km for Node 1, 1.55 km for Node 2), respectively. Both nodes achieved a BER below the hard-decision forward error

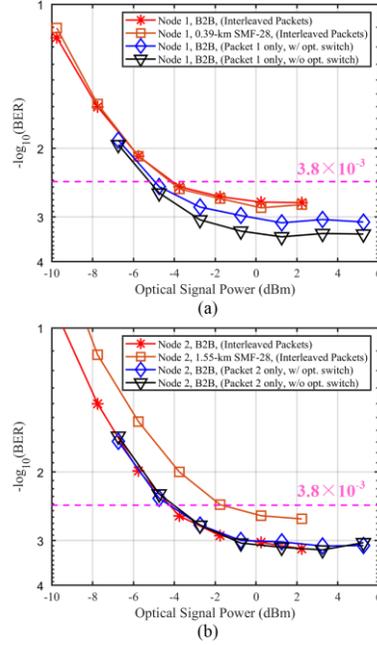


Fig. 7 Transmission performance of the packets from (a) Node 1, and (b) Node 2, when both laser diodes are working at the initial condition (1546.5nm). Optical signal power was calculated by the received optical power and the duty cycle of packets.

correction (HD-FEC) threshold of 3.8×10^{-3} [33] after transmission. The error floor as the power increased to higher than -2 dBm is primarily due to the limited bandwidth and ENOB of the transceiver DAC.

Due to the crosstalk and the amplitude noise of the driving signal to the MZM switch, a slight increase of the noise floor was observed for Node 1 signals after passing through the switch (Fig.7(a), diamond and triangle markers). The switch induced signal degradation was negligible for Node 2 signal due to lower crosstalk. As shown in Fig. 7(a), when both nodes are transmitting, a power penalty of approximately 1 dB was observed for Node 1 at both BtB and after 0.39 km SMF-28. Since the BER curves at BtB and after transmission are the same, we believe the 1-dB power penalty was due to the crosstalk from Node 2. However, the BtB BER results for the interleaved packets for Node 2 (Fig. 7(b)) show similar performance as non-interleaved packets, suggesting that Node 2 was not limited by the crosstalk from Node 1. Therefore, we can conclude that the 2.2 dB power penalty after 1.55-km SMF-28 for Node 2 was due to dispersion.

The crosstalk to Node 1 results from the limited extinction ratio (ER) of the EAMs used for Node 2 (about 10 dB) and the optical switch (about 20 dB). When the transmitters have the same wavelength, the crosstalk led to beat interference that degraded BER performance. In contrast with dispersion

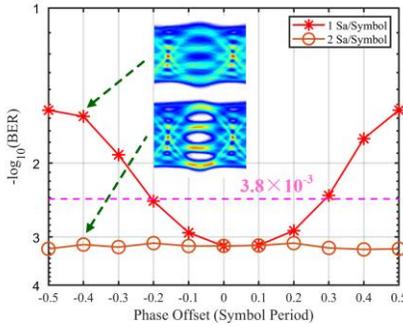


Fig. 8. Impact of phase mismatch on BER (Node 2, BtB, optical signal power=2.3 dBm); Insets: eye diagrams of recovered PAM-4 symbols with different sampling schemes, where phase offsets are both -0.4 symbol period.

induced inter-symbol interference, crosstalk cannot be easily mitigated using a linear equalizer. Thus, we emphasize that high extinction ratio transmitters and low crosstalk switches should be considered for optically switched data center networks.

B. Impact of timing misalignment

It has been shown that the offset of the sampling clock phase has a big impact on BER and CDR locking time in threshold detection based burst mode receivers [1]. To investigate the impact of sampling clock phase misalignment in equalizer-cashed burst mode receivers, we introduce an offset of the re-sampling clock phase in the DSP and show the BER of Node-2 packets in Fig. 8. Without losing generality, only the interleaved signal with an optical signal power of 2.3 dBm was evaluated at BtB transmission. The red asterisk and brown circle markers show the BER curves for the cases of 1 Sa/Symbol and 2 Sa/Symbol, respectively, where the numbers of equalizer taps are set to 15 and 31, respectively, for maintaining the same period of inter-symbol interference elimination. When sampling at the optimum clock phase, both detection schemes show the same BER performance. The BER for 1-Sa/Symbol detection starts to degrade after introducing timing misalignment and it tolerates up to about half a symbol before reaching the BER of 3.8×10^{-3} . For 2-Sa/symbol equalization, the BER is insensitive to clock phase misalignment. This implies that the CDR against clock drift could be simplified if a 2-Sa/symbol DSP equalizer is implemented in the burst mode receiver.

C. Comparison of convergence speed

To compare the equalizer convergence with and without equalizer caching, we plotted the normalized mean-squared error when using the whole packet as training symbols for Node 1 (Fig. 9(a)) and Node 2 (Fig.9(b)) when they both operated at 1546.5 nm. The normalized mean square error was defined as the ratio between error power and signal power in decibels, which was calculated using 480 packets. For the uncached equalizer, the equalizer initial condition is set to all zero except the center tap is 1. As shown in Fig. 9, with uncached states, the Node 1 equalizer requires approximately 300 symbols at the

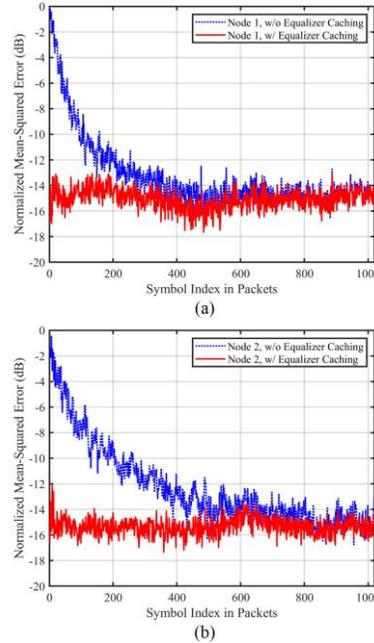


Fig. 9. Normalized mean square error in dB using the whole packet for equalizer training. (a) Node 1 after 0.39-km SMF-28 and (b) Node 2 after 1.55-km SMF-28. Signals are captured after transmission, both lasers are operating at 1546.5 nm.

beginning of each packet for training the equalizers, which corresponding to about 5.4 ns equalizer adaptation time. The number of required training symbols increases to about 700 symbols (about 12.5 ns adaptation time) for Node 2 due to the dispersion. With the cached equalizer states, equalizers can start with close-to-optimum states, and therefore, requiring minimum number of training symbols at the beginning of each packet, enabling low overhead and high throughput data transmission in optically switched data center networks.

D. Impact of a change of channel response

Next, we investigated the effectiveness and robustness of the equalizer state caching scheme when the channel responses vary. As discussed in the experimental setup, we emulated the impact of temperature variation by tuning laser 1 from 1545.3 nm to 1547.7 nm by increasing the temperature.

To study the impact of wavelength drift on the equalizer's convergence speed, we used the whole packet as training symbols and showed the change of the normalized mean-squared error over time, for different wavelength offsets, in Fig. 10(a). All the curves are obtained using the same initial equalizer state (the cached state), which is pre-trained when Node 1 is operating at 1546.5 nm ($\Delta\lambda=0$ nm). The results for negative wavelength offsets (i.e., $\Delta\lambda = -0.4, -0.8$ and -1.2 nm) are similar to that of their positive counterpart. Thus, for clarity,

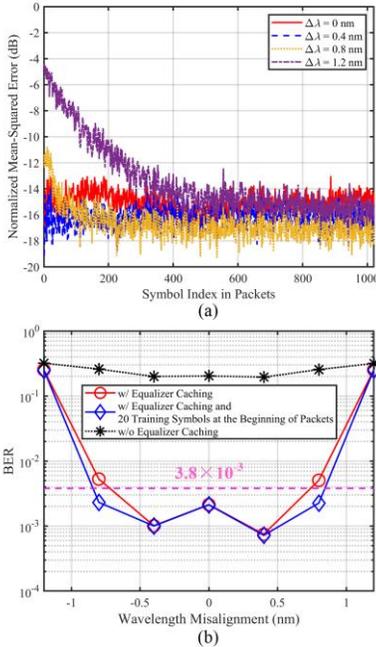


Fig. 10. Impact of wavelength drift on (a) the convergence of equalizers with coefficient caching and (b) the BER with different numbers of training symbols (Node 1, after 0.39-km SMF-28 transmission, optical signal power=2.26 dBm).

we only plotted the results for the positive wavelength offsets. As shown in Fig. 10(a), introducing 0.4-nm wavelength offset leads to lower error (thus better performance) due to the reduced crosstalk from Node 2. The cached state can be directly applied to the equalizer. At 0.8-nm wavelength offset, we start to see an increased normalized mean-squared error to about -12 dB at the beginning of the packet. However, it quickly drops to below -14 dB after using about 30 symbols for training. A significant increase of the normalized mean-squared error was observed at a wavelength offset of 1.2 nm, requiring approximately 500 symbols for training. Failure of convergence quickly resulted in a significant increase of BER. As shown in Fig. 10(b), the BER increases drastically when wavelength offset increased to ± 1.2 nm.

Due to the increase of BER at the beginning of the packets, the average BER values for a ± 0.8 nm wavelength offset go above HD-FEC threshold (red circle marker in Fig. 10(b)). This can be mitigated by adding a small number of training symbols at the beginning of each packet, e.g., 20 symbols, to assist the convergence of the equalizer. Finally, we show the BER performance using amplitude caching [21] and blind adaptation in asterisk markers. The BER results are above 0.1 regardless of the wavelength offset, indicating an unusable system. Thus, we conclude that equalizer state caching is necessary for optically switched networks using high baudrate and high order

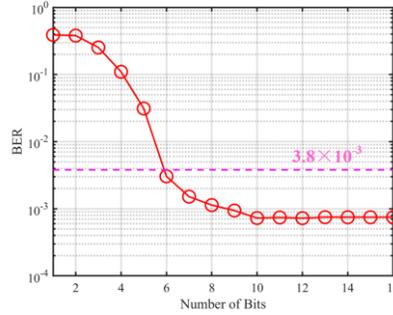


Fig. 11. Impact of the precision of equalizer coefficient caching on its performance (Node 1, after 0.39-km SMF transmission, optical signal power=2.3 dBm, $\Delta\lambda=0.4$ nm).

modulation formats.

The above experimental results allow us to estimate the temperature variation tolerance and the required update frequency in worst-case scenarios. Assuming a wavelength offset tolerance of ± 0.8 nm and a laser wavelength-temperature change rate of 0.1 nm/ $^{\circ}\text{C}$, the maximum tolerance of temperature variation in our system is ± 8 $^{\circ}\text{C}$. Considering the worst-case rate of temperature change of 0.11 $^{\circ}\text{C}/\text{s}$ [1], the equalizer state should be updated every 72 seconds.

In future optically switched networks with thousands of ports, dense wavelength division multiplexing (DWDM) with optical bandwidth of 50/100 GHz might be needed [4]. Consequently, the filtering introduced penalty would be more severe, requiring either tighter temperature control or more frequent update of the equalizer states.

E. Impact of digital resolution

To understand the requirement of ADC dynamic range, we studied the required resolution for the cached equalizer states, which directly correlates to the required memory size for storing cached coefficients. To make a fair comparison, we always scale the amplitudes of received signals between -0.5 and 0.5. The values of the equalizer coefficients in this work are thus all limited to within a range of [-15, 15]. Fig. 11 shows the BER of Node 1's packets at $\Delta\lambda=0.4$ nm using different resolution (shown as the number of bits) for the coefficients (calibrated under the initial condition of $\Delta\lambda=0$ nm). The results indicate that more than 8 bits resolution is needed to ensure good equalizer performance.

V. SUMMARY

We investigated the performance of equalizer state caching for low-adaptation-delay optically-switched data center interconnects. We showed that equalizer state caching can tolerate up to ± 0.8 -nm (± 100 -GHz) instantaneous wavelength change with only 20 training symbols, i.e., an adaptation delay of 0.36 ns ($=20/56$ Gbaud), enabling low BER optical burst detection with low overhead. Based on a proof-of-concept experimental setup using 56-Gbaud PAM-4 signalling, we showed that equalizer state caching is insensitive to clock phase

misalignment when operating at 2 Sa/symbol and can adaptively compensate for the variation of channel response due to slow wavelength drift. These results indicate that equalizer state caching should be considered for optically-switched systems to mitigate the impact of slowly varying optical interconnect impairments.

ACKNOWLEDGEMENT

The authors thank Dr. Kari A. Clark for discussion.

REFERENCES

- [1] K. Clark *et al.*, "Synchronous subnanosecond clock and data recovery for optically switched data centres using clock phase caching," *Nat Electron* vol. 3, pp. 426–433, Jul. 2020.
- [2] Q. Zhang *et al.*, "High-resolution measurement of data center microbursts," in *Proc. 2017 Internet Measurement Conference (IMC'17)* pp. 78–85 (ACM, 2017).
- [3] H. Ballani *et al.*, "Sirius: A flat datacenter network with nanosecond optical switching," in *Proceedings of the Annual conference of the ACM Special Interest Group on Data Communication on the applications, technologies, architectures, and protocols for computer communication*, pp. 782–797, 2020.
- [4] B. Buscaino *et al.*, "Multi-Tb/s-per-fiber coherent co-packaged optical interfaces for data center switches," *J. Lightw. Technol.*, vol. 37, no. 13, pp. 3401–3412, 2019.
- [5] J. D'Ambrosia, "IEEE P802.3bs baseline summary," 2015. [Online]. Available: <http://www.ieee802.org/3/bs/index.html>
- [6] A. Dochhan *et al.*, "Transceivers for inter-data center connections," in *Proc. IEEE Phot. Conf. (IPC)*, Orlando, FL, Oct. 2017.
- [7] M. H. Eiselt *et al.*, "Direct detection solutions for 100G and beyond," in *Proc. OFC/NFOEC 2017*, Los Angeles CA, March 2017, paper Tu3L3.
- [8] N. Eiselt *et al.*, "Experimental Demonstration of 112-Gbit/s PAM-4 over up to 80 km SSMF at 1550 nm for Inter-DCI Applications," in *Proc. ECOC 2016*, Duesseldorf, Germany, paper M2.D1, 2016.
- [9] M. Guo *et al.*, "112-Gb/s PAM4 with Joint Pre-and Post-Equalization for Data Center Interconnects," in *Proc. Asia Communications and Photonics Conference (ACPC)*, 2019, paper T2G.2.
- [10] J. K. Perin *et al.*, "Data center links beyond 100 Gbit/s per wavelength," *Opt. Fiber Technol.*, vol. 44, pp. 69–85, Aug. 2018.
- [11] O. Ozolins *et al.*, "100 GHz Externally Modulated Laser for Optical Interconnects," *J. Lightw. Technol.*, vol. 35, no. 6, pp. 1174–1179, Mar. 2017.
- [12] F. Chang *et al.*, "Forward error correction for 100 G transport networks," *IEEE Commun. Mag.*, vol. 10, no. 3, pp. S48–S55, Mar. 2010.
- [13] B. Teipen *et al.*, "Forward error correction trade-offs in reduced-latency optical fiber transmission systems," in *Proc. ECOC 2012*, P4.07, 2012.
- [14] K. Clark *et al.*, "Low Thermal Sensitivity Hollow Core Fiber for Optically-Switched Data Centers," *J. Lightw. Technol.*, vol. 38, no. 9, pp. 2703–2709, May. 2020.
- [15] R.S. Sohanpal *et al.*, "Clock and Data Recovery-Free Data Communications Enabled by Multi-Core Fiber with Low Thermal Sensitivity of Skew," *J. Lightw. Technol.*, vol. 38, no. 7, pp. 1636–1643, Apr. 2020.
- [16] M. D. Santa *et al.*, "25Gb/s PAM4 adaptive receiver equalization requirements for burst-mode transmission systems," in *Proc. Eur. Conf. Exhib. Opt. Commun.*, Dusseldorf, Germany, Sep. 2016, pp. 1–3.
- [17] G. Coudyzer *et al.*, "Study of burst-mode adaptive equalization for >25G PON applications," *J. Opt. Commun. Netw.*, vol. 12, no. 1, pp. A104–A112, Jan. 2020.
- [18] P. Ossieur *et al.*, "Burst-mode electronic equalization for 10-Gb/s passive optical networks," *IEEE Photon. Technol. Lett.*, vol. 20, no. 20, pp. 1706–1708, Oct. 2008.
- [19] G. Coudyzer *et al.*, "A 25Gbaud PAM-4 Linear Burst-Mode Receiver with Analog Gain-and Offset Control in 0.25 μm SiGe: C BiCMOS," *IEEE Journal of Solid-State Circuits*, vol. 55, no. 8, pp. 2206–2218, Aug. 2020.
- [20] K. Shi *et al.*, "System demonstration of nanosecond wavelength switching with burst-mode PAM4 transceiver," in *Proc. Eur. Conf. Opt. Commun.*, 2019, pp. 1–4.
- [21] T. Gerard *et al.*, "Fast and Uniform Optically-Switched Data Centre Networks Enabled by Amplitude Caching" in *Proc. Optical Fiber Communication Conference (OFC) 2021*.
- [22] Q. Cheng *et al.*, "Low-energy, high-performance lossless 8×8 SOA switch," in *Proc. Optical Fiber Communication Conference (OFC) 2015*, paper. Th4E.6, 2015.
- [23] C. P. Chen *et al.*, "Programmable dynamically-controlled silicon photonic switch fabric," *J. Lightwave Technol.*, vol. 34, no. 12, pp. 2952–2958, June, 2016.
- [24] A. C. Funnell, K. Shi, P. Costa, P. Watts, H. Ballani, and B. C. Thomsen, "Hybrid wavelength switched-TDMA high port count all-optical data centre switch," *J. Lightw. Technol.*, vol. 35, no. 20, pp. 4438–4444, Oct. 2017.
- [25] T. Gerard *et al.*, "Packet Timescale Wavelength Switching Enabled by Regression Optimisation," *IEEE Phot. Technol. Lett.*, vol. 32, no. 8, pp. 477–480, April. 2020.
- [26] T. Verolet *et al.*, "Ultra-Fast Tunable Laser Enabling 4 ns Coherent Slot Switching Beyond 100 Gbit/s," in *Proc. Eur. Conf. Opt. Commun.*, 2020, Paper. We1B-4.
- [27] ASHRAE, "Data center networking equipment—Issues and best practices," ASHRAE TC9.9, White Paper, 2011.
- [28] L. Pavesi and D. J. Lockwood, "Silicon photonics III," in *Topics in Applied Physics*, vol. 122. Berlin, Germany: Springer, 2016.
- [29] J. Zhu *et al.*, "First demonstration of a WDM-PON system using novel full C-band tunable SFP+ transceiver modules [Invited]," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 7, no. 1, pp. A28–A36, Jan. 2015.
- [30] A. Forencich *et al.*, "A Dynamically-Reconfigurable Burst-Mode Link Using a Nanosecond Photonic Switch," *J. Lightw. Technol.*, vol. 38, no.6, pp. 1330–1340, Mar. 2020.
- [31] X. Li *et al.*, "Effects of clipping and filtering on the performance of OFDM," in *Vehicular Technology Conference*, IEEE 47th, 1997, vol. 3, 1634–1638.
- [32] P. Duhamel *et al.*, "Classical Adaptive Algorithms (LMS, RLS, CMA, Decision Directed) seen as recursive structure," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1993, Canada, pp.496-499.
- [33] ITU-T Recommendation G.975.1, 2004, Appendix I.9.

Zhouyi Hu (S'17–M'20) received the B.S. degree in Optoelectronic Information Engineering from Huazhong University of Science and Technology in 2016 and the Ph.D. degree in Information Engineering from The Chinese University of Hong Kong in 2020. From July 2019 to April 2020, he was also with University College London as a visiting researcher. Upon graduation, he worked as a research associate at The Chinese University of Hong Kong from October 2020 to November 2020. In December 2020, He joined AiPT, Aston University as a research associate.

His research interests include physical layer security in PONs, optical wireless communications, advanced modulation formats, and DSP for optical systems.

Zichuan Zhou (S'18) received the B.Eng degree in Electronic and Electrical Engineering from University College London in 2018 and the MRes. degree in Integrated Photonic and Electronic Systems from University of Cambridge in 2019. He is currently pursuing the Ph.D. degree in Electronic and Electrical Engineering at University College London, UK.

His research interests include data center interconnection, integrated photonics and optical transceiver.

Calvin CK Chan (Senior member, IEEE; Fellow, OSA) received his B.Eng., M.Phil. and Ph.D. degrees from the Chinese University of Hong Kong, all in Information Engineering. Upon graduation, he joined the Department of Electronic Engineering at the City University of Hong Kong as a Research Assistant Professor. Later, he joined the Photonic

Networks Department at Bell Laboratories, Lucent Technologies in Crawford Hill Holmdel NJ, as a Member of Technical Staff. He then served as Senior Optical System Engineer at Jedai Broadband Networks, Inc. in NJ, USA before he joined Department of Information Engineering at the Chinese University of Hong Kong and now serves as Professor.

Prof. Chan has published more than 320 technical papers in refereed international journals and conferences. He holds two issued US patents, one edited book and 2 book chapters. He has served as the co-chair and TPC members for a number of international conferences, including OFC, OECC, CLEO/Pacific Rim, ACP, etc. He was an associate editor of OSA Journal of Optical Networking, IEEE/OSA Journal of Optical Communications and Networking and Frontiers in Optoelectronics. His main research interests include WDM optical metro/access networks, high-speed optical and digital signal processing, optical performance monitoring, optical network planning and optimization. He is an OSA Fellow.

Zhixin Liu (Senior member, IEEE) received the B.Eng. degree in information engineering and the B.B.A. degree in business administration from Tianjin University, Tianjin, China, in 2006, the M.S. degree in electrical engineering from Shanghai Jiao Tong University, Shanghai, China, in 2009, and the Ph.D. degree in information engineering from the Chinese University of Hong Kong, Hong Kong, in 2012.

He joined the Optoelectronics Research Centre, University of Southampton, Southampton, U.K., in 2013. In 2016, he started his academic career at the Department of Electronic and Electrical Engineering at University College London. His research interests include transceivers for transmitters for short and medium reach optical communication systems, coherent optical signal processing, and high fidelity signal generation and digitization. He is a Senior Member of the Optical Society of America.