Clock Synchronized Transmission of 51.2 GBd Optical Packets for Optically Switched Data Center Interconnects

Zichuan Zhou, Kari Clark, Callum Deakin, Petros Laccotripes and Zhixin Liu
Optical Networks Group, Department of Electronic and Electrical Engineering,
University College London, London, UK WC1E 7JE
Author e-mail address: zczlzz0@ucl.ac.uk, zhixin.liu@ucl.ac.uk

Abstract: We demonstrate clock-synchronized transmission of 128-byte optical packets at 51.2GBd. We study the impact of reference clock phase noise on the clock phase misalignment and system scalability for high baud rate DC interconnects. © 2021 The Authors

1. Introduction

The drastic increase of data traffic in cloud data centers has stimulated research on optical switching for future data center (DC) interconnects, offering potential benefits of large port count, large bandwidth, and lower power consumption [1]. It has been shown that clock synchronization and fast (e.g. <1 ns locking time) clock and data recovery (CDR) are essential for achieving high throughput (e.g. >90%) in optically switched DC networks [2]. One of the major technical challenges in such systems is the timing jitter, which affects the signal bit error ratio (BER) as well as the amount of network overhead for clock synchronization. For a given transmitter to receiver pair, the total timing jitter arises from two sources: a) random temperature variation that changes the light propagation delay in optical fibers [3], and b) jitter of the clock sources that the transmitter and receiver are both synchronized to. The temperature induced timing jitter is relatively slow and, therefore, can be compensated by using clock phase caching [2] or significantly reduced by using low thermal sensitivity fibers [4,5]. The clock source induced jitter, however, is primarily high frequency and hard to compensate. Although the impact of the source clock jitter on system performance was insignificant in low baud rate interconnects (e.g. <25 GBd), it becomes prominent as the baud rate for DC interconnects increases to 50 GBd and even >100 GBd [6,7]. It is therefore important to study the jitter and timing tolerance of clock synchronized networks for high baud rate signaling in optically switched DC networks.

In this paper, we experimentally investigate the impact of source clock jitter on a real-time 51.2 GBd on-off-keyed (OOK) packet transmission system. Based on our real-time clock-synchronized prototype system, we report the tolerance of clock timing misalignment for 51.2 GBd signals. To show the impact of source clock jitter, we experimentally measure the system's tolerance to clock timing misalignment using four different clock sources with different phase noise. In addition, we show that the phase noise of the source clock also influences the distance scalability of the DC networks, by characterizing the change of the timing tolerance window for different fiber lengths. All investigations were performed using a target BER of $<10^{-10}$, which is preferred for minimizing latency associated with forward error correction (FEC) in DC interconnects.

2. Experimental Set-up

Fig.1 shows the experimental set-up. The clock distribution section is shown with green background. Four inexpensive commercial 100-MHz oscillators with different phase noise were used as source clocks: Wenzel Sprinter, Texas Instruments LMK61PD0A2, Crystek CPRO33-100, and Agilent E4432B. The source clock seeded a phase-lock-loop (PLL1, TI LMX2595) to generate a 400 MHz digital clock signal and was modulated on to a 1550-nm continuous wave (CW) using $LiNbO_3$ Mach-Zehnder modulator (MZM 1) biased at the quadrature point. After modulation, the optical clock signal was subsequently split and sent to the transmitter (Tx) through a 2-m patchcord and to the receiver (Rx) through a 0.26/0.52/1/1.55/2.07 km standard single mode fiber (SMF-28). The clock distribution link was amplified and attenuated by a variable optical attenuator (VOA1) to emulate the loss induced by a light splitting approach to clock distribution. The power of the optical clock signals is the same (-5 dBm) for both Tx and Rx. The 400-MHz optical clock was detected by a PIN photodiode followed by a transimpedance amplifier, which increased the RF power to -7.5 dBm.

At the Tx side (orange background), the detected clock signal was fed to PLL2, which generated a 400 MHz reference clock to drive the Tx-side field programmable gate arrays (FPGA, Xilinx VCU108) and a 12.8 GHz clock, which was doubled to 25.6 GHz to drive the 2:1 RF multiplexer. The two FPGA transmitters each output 25.6 GBd OOK packets, which were electronically interleaved by the RF multiplexer to create 51.2 GBd packets. The generated 51.2 GBd packets each contained a 1536-bit (30 ns) payload consisting of three consecutive PRBS-9 sequences each followed by a final 0 bit, with a 96 bit pattern embedded at the beginning of the third sequence for frame alignment,

with a 2.5 ns gap between consecutive packets. The electronic packet stream modulated an 18-dBm CW laser through a 40-GHz RF amplifier and a 40 GHz MZM, generating optical packets as shown in the inset in Fig. 1.

To emulate the rack-to-rack interconnects, the optical packets was transmitted through 130 m SMF-28 before detection by a 50 GHz bandwidth photodiode at the Rx side (purple background). A variable optical attenuator (VOA2) was used to attenuate the optical power for BER measurement. The detected signal was amplified by a 65-GHz RF amplifier before entering 1:2 RF demultiplexer driven by a 25.6 GHz clock generated by PLL 3. The demultiplexer converted the 51.2 GBd packet to two streams of 25.6 GBd packets, which were fed into the Rx FPGA for real-time BER calculation. Following the same configuration as the Tx, the Rx FPGA was driven by a 400 MHz reference clock generated by PLL 3, which was synchronized to PLL 2 by optical clock distribution. The BER for 51.2 GBd transmission was measured by averaging the BER of the two 25.6 GBd packet streams (from the first two 512-bit payloads of the 51.2 GBd packets). To claim a BER of under 10^{-10}, we kept the BER calculator running for 4 seconds, which corresponded to a total of 1.311×10^{11} bits.

To study the timing window for clock phase offset, we firstly enclosed the fiber spools in a thermal insulation chamber to minimize the error caused by change of light propagation delay due to the temperature variation, ensuring that the clock phase remained stable over long periods of time without requiring active clock phase tracking. Then we used a calibrated 40-GHz RF phase shifter (with 0.175ps resolution) to shift the 25.6 GHz clock that drove the demultiplexer shifter. The impact of clock phase noise decorrelation was studied by changing the fiber length in the clock transmission path.

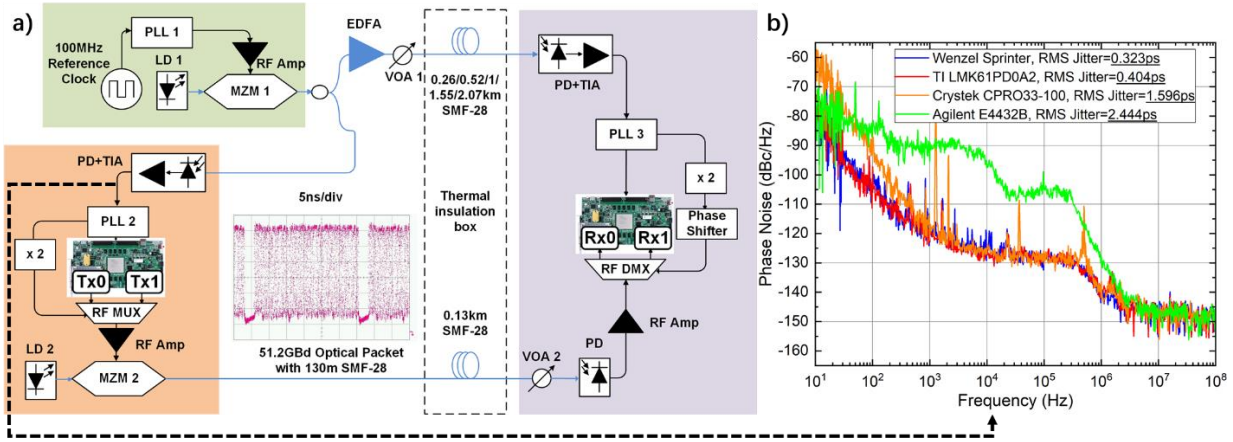

Fig. 1. a) Experimental Set-up for 51.2 GBd OOK packet clock synchronized data center interconnection. PLL: Phase-lock-loop, TIA: transimpedance amplifier, VOA: variable optical attenuator, PD: photodetector, MZM: Mach-Zehnder Modulator, EDFA: Erbium-doped fiber amplifier. b) Phase noise measurement of different oscillators used to generate the 400 MHz synchronization clock

3. Experimental Results

Fig. 1b shows the measured phase noise of the detected 400 MHz clock signals and the jitter values (integrated from 10 Hz to 100 MHz). Although the 100-MHz Wenzel Sprinter oscillator has lower phase noise at high frequency, it exhibits the same performance as the TI LMK61PD0A2 (an integrated circuit oscillator) after conversion to 400 MHz, indicating that the high frequency (>2 kHz) phase noise is dominated by PLL 1. Similar high frequency phase noise performance is achieved using the Crystek oscillator (orange curve), however, it has worse noise between 10 and 2000 Hz due to higher source oscillator noise. Agilent E4428B (green curve) has the worst phase noise and the highest jitter (about 2.44 ps). Using Wenzel Springer as source oscillator, we measured the performance of the 51.2GBd signal after optimizing the Rx-side clock phase. As shown in Fig.2a, the optical power for the BER of 10^{-10} at back-to-back and after 130-m SMF-28 are the same (-2.6 dBm), indicating no transmission impairment at the optimal clock phase. The low receiver sensitivity is due to the high input power (0 dBm) required by the RF demultiplexer.

Fig. 2b shows the BER as we offset the Rx-side sampling clock by tuning the RF phase shifter. Different markers show the tolerance window for timing or clock phase offset when using different source oscillators. The received optical power is adjusted to 4.5 dBm for all source oscillators. Due to the high baud rate and the limited transceiver bandwidth, the timing tolerance window is only 8.0 ps (corresponding to about 0.4 of a symbol) when using the Wenzel Sprinter as the source oscillator. With higher clock phase noise from using the TI LMK61PD0A2, Crystek CPRO, or Agilent E4432B as source oscillators, the timing tolerance window decreases to 7.7 ps, 6.4 ps and 1.0 ps respectively. The reduced tolerance to clock phase offset means a faster clock phase tracking mechanism is required.

If one considers the clock phase caching method demonstrated in [2], one needs an 8 times faster caching rate (incurring at least 8 times more network overhead) if the lowest performance oscillator is used as the clock source oscillator, rather than the best reference oscillator in our experiment.

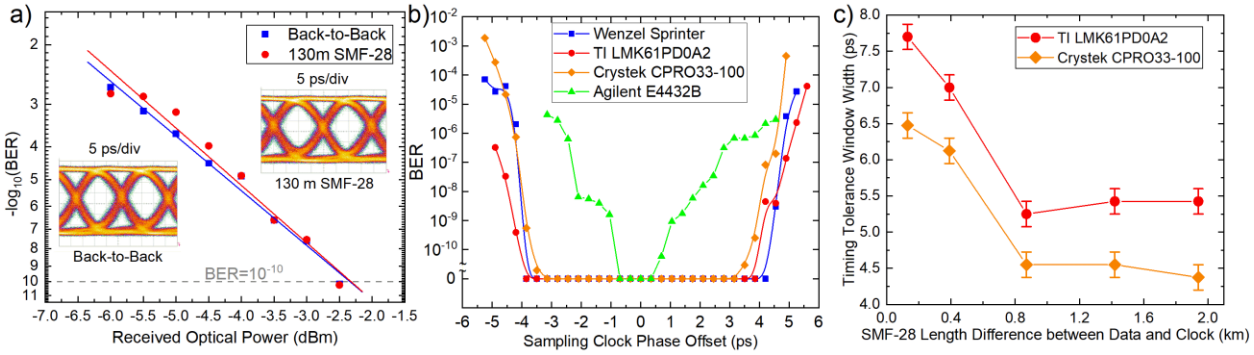

Fig. 2. a) BER characterization (Red circles: Back-to-Back, Blue squares: after 130 m SMF-28); b) Tolerable sampling clock phase offsets with different reference clocks (Red circles: LMK61PD0A2, Green triangles: Agilent E4432B, Blue squares: Wenzel Sprinter, Orange diamonds: Crystek CPRO33-100); c) Effect of fiber transmission distance mismatch on timing tolerance window

As the path length difference increases, the phase noise of distributed 400 MHz clock becomes decorrelated and results in increased timing jitter, and consequently, leads to a reduction in the tolerance window for the clock phase misalignment. This was characterized and shown in Fig.2c. As with previous experiment, the received optical power is set to 4.5 dBm. Using TI LMK61PD0A2 as the source oscillator, the timing tolerance window was reduced from 7.7 ps to about 5.4 ps when the length difference increased from 130 m to about 1 km, and remained the same as the fiber further increase to up to 1.94 km. Similarly, using Crystek CPRO as source oscillator, the timing tolerance window decreased from 6.4 ps to 4.4 ps as the fiber length difference increased from 130 m to 1 km, and stopped further degradation as the clock fiber length increases. Our results indicate that low phase noise source oscillator offers better distance scalability for clock synchronized DC interconnects. Designer should consider minimize the total jitter due to the decorrelated phase noise, rather than the absolute jitter of the clock or the data, for clock synchronized data communications.

4. Conclusion

We demonstrated clock synchronized transmission of real-time optical packets at 51.2 GBd and benchmarked system performance by using different inexpensive commercial oscillators as the system source clock. We demonstrated that a low phase noise source oscillator is crucial in clock synchronized systems as it offers a larger tolerance window for clock phase misalignment (8.0 ps in our experiment) and it allows these systems to scale to longer distances, crucial to future large-scale optically-switched cloud data centers. As the baud rate of DC interconnects increases to >50 GBd, we emphasize that low phase noise clock synchronization must be achieved across the whole DC network to enable high-throughput, sub-nanosecond clock recovery time all-optical switching.

The authors acknowledge Royal Society (PIF/R1/180001), EPSRC (EP/R035342/1, EP/R041792/1) for the funding support. The authors thank Rohde & Schwarz for the loan of the phase noise analyzer.

5. References

[1] H. Ballani, P. Costa, R. Behrendt, D. Cletheroe, I. Haller, K. Jozwik, F. Karinou, S. Lange, K. Shi, B. Thomsen, and H. Williams, "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 (ACM, 2020), pp. 782–797.
[2] K. A. Clark, D. Cletheroe, T. Gerard, I. Haller, K. Jozwik, K. Shi, B. Thomsen, H. Williams, G. Zervas, H. Ballani, P. Bayvel, P. Costa and Z. Liu , " Synchronous subnanosecond clock and data recovery for optically switched data centres using clock phase caching, " Nat. Electron., **3,**426–433(2020).
[3] R. Slavik, G. Marra, E. Fokoua, N. Baddela, N. V. Wheeler, M. Petrovich, F. Poletti and D. J. Richardson., "Ultralow thermal sensitivity of phase and propagation delay in hollow core optical fibres," Sci. Rep. **5,** 1-7 (2015).
[4] R. S. Sohanpal, K. A. Clark, B. J. Puttnam, Y. Awaji, N. Wada, P. Bayvel, and Z. Liu, "Clock and Data Recovery-Free Data Communications Enabled by Multi-Core Fiber With Low Thermal Sensitivity of Skew," J. Lightwave Technol. **38,** 1636-1643 (2020)
[5] K. A. Clark, Y. Chen, E. Fokua, T. Bradley, F. Poletti, D. J. Richardson, P. Bayvel, R. Slavík, and Z. Liu, "Low Thermal Sensitivity Hollow Core Fiber for Optically-Switched Data Centers," J. Lightwave Technol. **38,** 2703-2709 (2020)
[6] X. Pang, O. Ozolins, R. Lin, L. Zhang, A. Udalcovs, L. Xue, R. Schatz, U. Westergren, S. Xiao, W. Hu, G. Jacobsen, S. Popov, and J. Chen, "200 Gbps/Lane IM/DD Technologies for Short Reach Optical Interconnects," J. Lightwave Technol. **38,** 492-503 (2020)
[7] Y. Matsui, R. Schatz, D. Che, F. Khan, M. Kwakernaak and T. Sudo, " Low-chirp isolator-free 65-GHz-bandwidth directly modulated lasers," Nat. Photonics, **15,** 59-63(2021).