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

Data-Centric Transmission with Adaptive FEC for Ultra-Low Latency Resource Sharing in Wide Area

Toshiya Matsuda, Kota Nishiyama, Takeshi Seki, Takashi Miyamura

NTT Network Service Systems Laboratories, NTT Corporation, toshiya.matsuda.fa@hco.ntt.co.jp

Abstract *We propose individual error correction techniques for headers and data to adapt to various transmission requirements of data. We also experimentally demonstrate lossless transmission via 100-GT/s optical interfaces up to 90 km with less than 1.5% increase in latency due to network equipment.*

Introduction

The expansion of cloud services strongly demands large amounts of resources, leading to a trend of sharing of resources on wide area networks.

Some of these services require low latency, e.g., trading or object tracking video surveillance [1], while others require reliability, e.g. scientific calculations and remote surgery [2]. The concept of data-centric [3] has been proposed to enable various services to efficiently share the wide variety of data present in the nodes in a network. Dense wavelength-division multiplexing (DWDM) systems are commonly used to guarantee transmission distance and capacity in data centre interconnects (DCIs). Although reliable, DWDM systems are ineffective in terms of latency for sub-100 km transmission distances because of frame conversions between client devices and DWDM systems and signal processing for high-performance forward error corrections (FECs) [4].

NTT proposes the Innovative Optical and Wireless Network (IOWN) [5] to create an innovative information processing infrastructure. The All-Photonics Network (APN), which is one of the technologies comprising IOWN, converts all information transmission and relay processing to photonics-based processes. We studied an optical bus platform architecture [6] based on PCI Express (PCIe) packets that provide an ultra-low latency optical path over the APN. The platform enables the FEC signal processing to be optimized to the distance by selecting the FEC type independent of transmission frames.

In this paper, we propose individual error correction techniques for frame headers and payload data in the platform to improve the frame loss rate (FLR). We conducted transmission experiments with concept demonstration machines focused on the transmission function of the platform and showed experimentally that the network equipment latency is negligible compared to that for a 90-km fibre only with no frame loss.

Optical bus platform architecture

Figure 1 shows a schematic of the optical bus platform architecture.

The platform provides an external optical bus to connect endpoint devices over an internal bus within servers located throughout the metropolitan area. The network topology should be based on a ring topology for use in metro areas.

The optical bus platform has three technical features for low latency and high reliability. First, direct optical conversion of bus signals simplifies the layer structure. Data is encoded with FEC as necessary, accommodated in PCIe frames and transferred from the endpoint device to the optical bus platform via a multi-lane PCIe bus. The platform replaces the physical layer frame and directly converts the PCIe frames to optical signals by transmitters (Txs). Then, Txs use WDM or space-division multiplexing (SDM) techniques to generate optical paths equal to the lane width of the bus. Buffers for frame conversion are not needed since signals do not go through Ethernet or an Optical Transport Network (OTN) frame. Secondly, optimized FEC

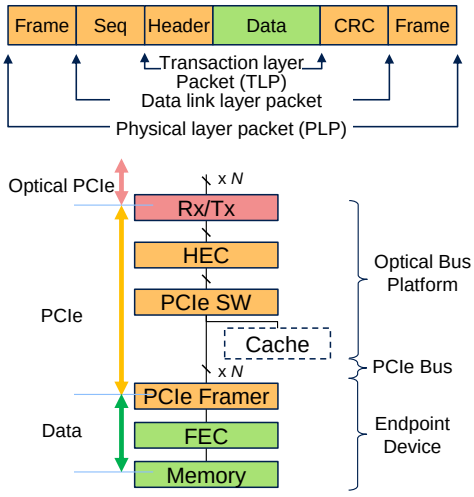

Fig. 1: Schematic diagram of optical bus platform architecture

separated from transmission frames can minimize signal processing. Separation of FEC from transmission frames enables the FEC signal processing to be optimized depending on data requirements. Header error correction (HEC) can simply correct bit errors in the headers of received frames by copying the header of a normal frame to the others based on the CRC check in each lane regardless of FEC. Finally, buffering for timing adjustment is only at packet add in the vicinity of the data. Non-blocking through packet minimizes packet processing and buffering in intermediate PCIe switches (SWs).

Experiments and results

We conducted transmission experiments to evaluate the effectiveness of our architecture.

Figure 2 illustrates a block diagram of the concept demonstration machine and experimental setup. Each FPGA emulated the PCIe SW and a cache memory with a PCIe framer and FEC. BCH (64, 51) code was used for FEC as optimized short packets for instructions or control signals. The PCIe framer generated 4-lane PCIe Gen3 compliant packets with the same header. The type of TLP implemented is only a memory request, and the format bits in the header identify “read” and “write”. The framer can generate a packet to measure the latency. A counter calculated the transmission latency without the FEC block from the clock numbers when the packet for latency measurement was transmitted and received. The SW can adjust the effective transfer rate by insertion of dummy PCIe packets. The machine has two QSFP28 ports as 4 x 25-GT/s optical interfaces.

The optical modules used in the experiments were QSFP28 100GBASE-ER4. Each optical interface of the machine was connected with several tens of kilometres of single mode fibres (SMFs). The HEC at the post-stage of the optical interface on the receiver side checks the CRC of packets in each lane and copies the header of normal packets to other headers after deskew processing. Under the condition that each lane has an equal bit error rate (BER), the FLR with HEC is approximately expressed by the following equation.

$$\mathrm{FLR} \approx (d \cdot r)^N \qquad (1)$$

where d is the frame length, $r\,(r \ll 1)$ is the BER before FEC, and N is the lane width. The switch on the receiver side looked at the destination in the packet header and sent packets addressed to itself to the memory, while passing packets with other destinations. An optical variable attenuator (VAT) was inserted before the port #1 of machine A to adjust the optical received power. It has two serial interfaces to communicate with a controller for transmission data and control signals.

First, we investigated the performance of the FEC and the HEC. Port #0 and port #1 of machine A were connected back-to-back via the optical VAT. Measurement packets addressed to itself containing data, which is a 2^{15}-1 pseudo random binary sequence (PRBS), was accommodated in multiple packets with a payload size of 1024 bytes. Figure 3 shows measured BERs and FLRs as a function of the mean optical received power of four lanes. Closed and open circles indicate BERs without and with FEC, respectively. Closed and open squares indicate FLRs with and without HEC, respectively. The BCH (64, 51) code is a 2-bit correctable code with a code length of 64 bit, and the theoretical BER after FEC is shown in Fig.3 by the solid line. Measured BERs with FEC agree well with the theoretical values. If a bit error occurs in a particular field, e.g. Fmt, Type, Tag, and etc., in the header, the frame is discarded in the SW, resulting in frame loss. Therefore, the measured FLRs without HEC were a constant multiple of the BER without FEC, regardless of FEC. On the other hand, measured FLRs with HEC were frame loss free below 1e-12. The measured FLRs with HEC were lower than those estimated from Eq.(1) because some lanes had better characteristics due to variations in the

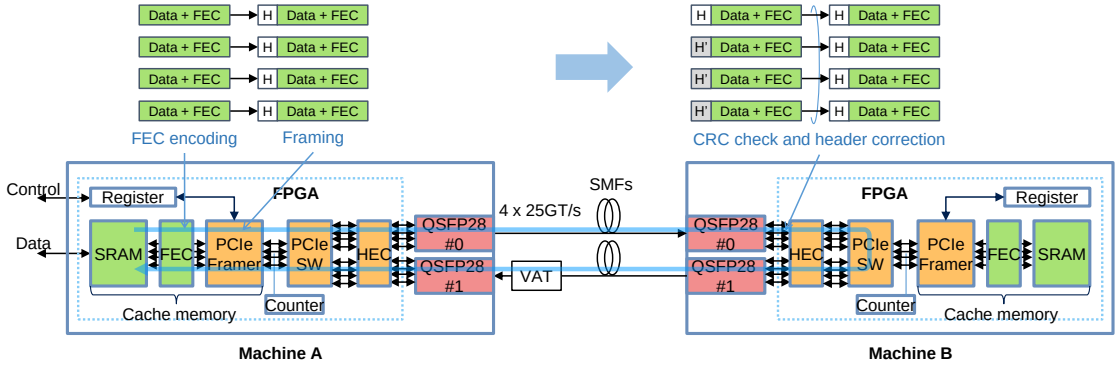

Fig. 2: Block diagram of the concept machine and experimental setup.

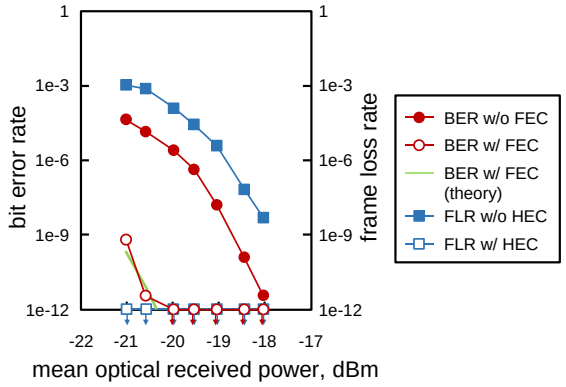

Fig. 3: Measured BERs and FLRs as a function of mean optical received power.

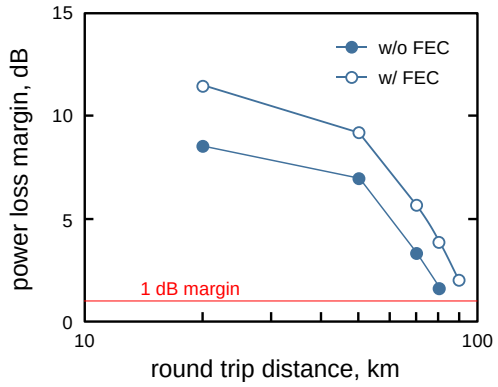

Fig. 4: Power loss margins as a function of round-trip distance.

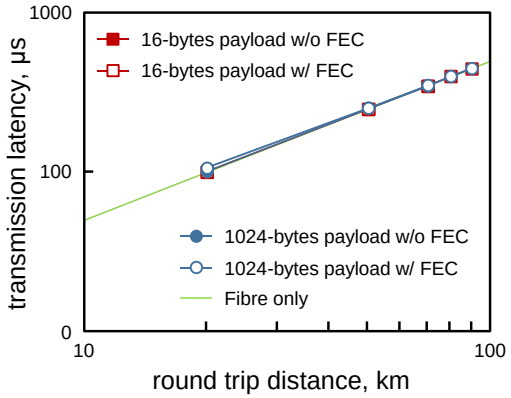

Fig. 5: Average transmission latencies as a function of round-trip distance.

received optical power of each lane.

Next, we investigated the transmission performances and the latency characteristics of the packet transport. SMFs of the same length were used for the outbound and inbound routes, and the BER was measured by adjusting the received power with the optical VAT. Figure 4 shows optical power loss margins, which are defined by the amount of attenuation with BER as 1e-12, as a function of round-trip distance. Closed and open circles indicate power loss margins without and with FEC, respectively. Without FEC, the distance over which a loss margin of more than 1 dB could be obtained was 80 km. With FEC, the distance with more than 1-dB loss margin increased to 90 km. This means that FEC is not required at transmission distances of 2 x 40 km or less.

Figure 5 shows average transmission latencies as a function of round-trip distance. Closed squares and circles indicate measured latencies without FEC for 16-bytes payload and 1024-bytes payload, respectively. The total values, including the latencies of the FEC block calculated by simulations, are shown as open squares and circles in Fig.5. The fibre only latency is also shown as a solid line. Measured latencies without FEC up to 80 km are almost consistent with fibre only latency. Estimated latencies with FEC at 90 km are almost the same as the fibre only latency. Table 1 shows increases in latency due to network equipment. For short packets with 16-bytes payload, increase in latencies was suppressed to 1.0% or less up to 90 km by adaptively using BCH (64, 51) code. In measurements where the SMFs were replaced by optical attenuators, the average latency of machine A as a transceiver was 540 ns and machine B as a repeater was 420 ns. Even for long packets with 1024-bytes payload, increase in latencies was 1.5% or less up to 90 km with FEC. The use of FECs with longer code lengths is expected to further reduce the transmission latency of long packets.

Tab. 1: Increases in latency due to network equipment.

	16-bytes w/o FEC	1024-bytes w/o FEC	16-bytes w/ FEC	1024-bytes w/ FEC
20 km	1.0%	1.3%	1.2%	7.0%
50 km	0.4%	0.5%	0.5%	2.8%
70 km	0.3%	0.4%	0.4%	2.0%
80 km	0.2%	0.3%	0.3%	1.7%
90 km	-	-	0.3%	1.5%

Conclusion

We proposed individual error correction techniques for frame headers and data in payloads that allows flexible selection of FECs according to data requirements.

PCIe packets were successfully transmitted lossless from 20 km to 90 km with or without FEC of BCH (64, 51) code. Data without FEC cloud be transmitted over short transmission distances with low latency increases to optical fibre only delay. On the other hand, data with FEC could be transmitted over longer distances without extreme delay increases.

References

[1] A. Dochhan, J. K. Fischer, B. Lent, A. Autenrieth, B. Shariati, P. W. Berenguer, and J. -P. Elbers, “Metro-haul project vertical service demo: Video surveillance real-time low-latency object tracking,” in *Proc. Optical Fiber Communication Conference (OFC)*, San Diego, CA, USA, 2020, pp. 1-3, DOI: 10.1364/OFC.2020.M2D.4.

[2] H. Laaki, Y. Miche, and K. Tammi, “Prototyping a digital twin for real time remote control over mobile networks: Application of remote surgery,” *IEEE Access*, vol. 7, pp. 20325–20336, 2019, DOI:10.1109/ACCESS.2019.2897018.

[3] M. Esler, J. Hightower, T. Anderson, and G. Borriello, “Next century challenges: Data-centric networking for invisible computing: The portolano project at the University of Washington,” *Prof. 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom '99)*, Aug. 1999, pp. 256-262.

[4] B. Teipen, M. Filer, H. Grießer, M. Eiselt, and J. P. Elbers, “Forward error correction trade-offs in reduced-latency optical fiber transmission systems,” in *Proc. of 38th European Conference and Exhibition on Optical Communications (ECOC)*, Amsterdam, Netherlands, 2012, P4.07, DOI: 10.1364/ECEOC.2012.P4.07.

[5] A. Itoh, “Initiatives concerning All-Photonics- Network-related technologies based on IOWN,” *NTT Technical Review,* vol. 18, no. 5, pp. 11-13, 2020.

[6] T. Matsuda, K. Nishiyama, K. Masumoto, M. Nakagawa, and T. Miyamura, “Ultra-low latency short packet transmission experiments with optical bus platform based on PCIe,” in Optical Fiber Communication Conference (OFC), San Diego, CA, USA, 2022, Tu3G.2.