Fast reconfigurable optical network with edge data center nodes for low-latency 5G applications

H. Santana,[1] B. Pan,[1] R. Kraemer,[1] A. Mefleh,[1,2] and N. Calabretta[1]

[1] Eindhoven University of Technology, Eindhoven, The Netherlands
[2] KPN, The Hague, The Netherlands

To fulfil stringent latency requirements of time-sensitive applications in 5G networks, data has to be processed in a decentralized way. The Edge Cloud Network is composed of computational resources placed at most at tens of kilometers far from the sources of the data flows they need to process. Besides the physical proximity, nanosecond-scale reconfiguration time of optical switches and a fast control of the optical networks are also required to guarantee dynamicity with latencies on the order of tens of microseconds. In this work, we propose an Edge Cloud network composed of a metro-access ring and optically switched edge data center. SOA-based ROADMs are used in the ring for fast add/drop of wavelengths. The edge data center is composed of top-of-rack switches interconnected via an SOA-based optical switch. A supervisory channel is used by the network nodes to exchange control packets in a time-slotted synchronous fashion, and FPGA-based controllers guarantee nanosecond-scale reconfiguration decisions.

Introduction

To meet the latency and bandwidth requirements of 5G applications, data centers have been moving from centralized operation to edge operation [1]. This means that smaller data centers with computational and storage resources are placed close to where the data flows need to be processed, therefore reducing the propagation delay by design of the network.

Edge data centers are typically used for specific applications, for example, content caching. However, as the segment grows, we can envisage a scenario where edge data centers operate in a flexible way as coordinated computational units able to meet strict latency requirements for applications that require parallel processing (Fig. 1).

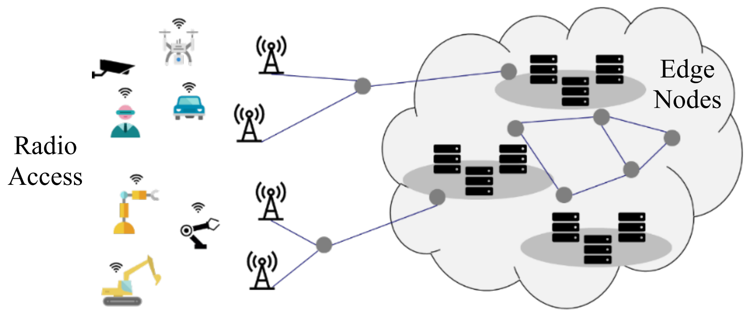

Fig 1. – Vision of federated edge data center networks

As key parameters for the time domain performance of edge data center networks, we can name the latency, defined as the delay between source and destination, the jitter, defined as the variation in the latency, and the delay for network reconfiguration.

In the edge data center operation, one node can receive thousands of flows generated by connected users and machines. There will be competition for the available resources and some flows will wait in a queue until the computational resources become free. Therefore, the longest waiting times, i.e., the tail of the latency distribution, is the relevant performance parameter to determine the service quality.

Cloud gaming is one application that illustrates the importance of consistent latencies among the users. In this application, the actions of the game are processed in a data center and then rendered to the players in their particular displays. Their experience can be critically impacted if the time they experience to obtain a response from the data center is too long.

Also, in the Industry 4.0, the machines have a digital representation. To perform as accurately as possible, the digital representation of the system should be trustworthy and reflect variations in the state of the machine immediately, therefore showing the importance of low-latency in the delivery of information.

The new applications leveraged by 5G also bring new requirements in terms of time synchronization. Several applications have strong requirements in terms of distributed processing. For example, real-time applications based on artificial intelligence, such as object recognition and tracking. In these applications, the processing has to be broken in smaller pieces and distributed among computational resources. These have to be orchestrated to output an action within a time delay [2].

Transparent connections in the network are also a desirable feature in providing latency guarantees. Establishing an optical circuit path between the source and the destination has as consequence a precise latency in the link, namely the optical propagation time, bringing determinism to the connection [3]. Therefore, optical switches which can reconfigure fast enough not to affect the performance of the network are required. However, optical switches with high port count are typically slow. On the other hand, optical switches based in electro-optical effects can reconfigure in nanosecond-scale. An architecture for optical switches based on a broadcast-and-select using semiconductor optical amplifiers (SOA) approach has been presented in [4]. Because of the amplification provided by the SOAs, lossless and fast optical switches can be designed with this architecture.

To maintain the dynamicity of the network, bottlenecks should be avoided in all levels. That means that we should have a hardware that can reconfigure fast and a control plane that can react fast and implement decision in μs-scale. Control plane solutions used in software-defined networking (SDN) are slow in this regard and they rely on general purpose computer systems, which hinders deterministic operation. Field-programmable Gate Arrays (FPGAs) offer the possibility of precisely driving several parallel interfaces according to its internal clock. Therefore, we can leverage fast hardware-level control and solve the bottleneck in the control plane [5].

Proposed Edge Data Centers System

The proposed system consists of edge nodes arranged in a ring network, as shown in Fig. 2(a). The edge nodes use a 2-degree ROADM to add and drop wavelengths to/from the ring, depicted in Fig. 2(b). The ROADM is based on SOAs, therefore providing fast reconfigurability to the network. The incoming wavelengths are separated by an AWG (arrayed waveguide grating) and each wavelength has a dedicated SOA which can block it or let it through. The SOA-based ROADM has been validated in a previous work for metro-access operation using SDN [6].

FPGA-based controllers provide fast reaction time to the control commands. FPGAs excel at driving several interfaces in parallel, as many actions can be taken within a clock cycle. In the system, the FPGAs operate with a clock of 156.25 MHz, which means that control actions can be taken every period of 6.4 ns.

A dedicated supervisory channel is used to inform the state of the network and to share the slots for operations. This channel is dropped at every node and uses control packets in the XGMII format, containing 64 bits of data and 8 bits of control flags [7]. Between two control packets, the control flag 0x07 is inserted to maintain the link.

The supervisory channel is used to reserve slots for the transmission of nodes and to distribute the time in the system. With the transmission slots, the system operates in a time-domain multiplexed fashion. When the controller in the node receives a control packet, it decodifies the schedule contained in it and reconfigures the SOAs in the node accordingly to allow the transmission of specific wavelengths. One control packet can be used to reserve one or more slots after it. The number of reserved slots per control packet as well as the slot duration affects the dynamicity of the network, but can increase the network efficiency. The latency due to the control plane can be divided in the latency processing the control packets and the optical switch reconfiguration latency. Since the longest delay is spent processing the control packets, by scheduling several slots in a row, we increase the ratio of time spent transmitting data.

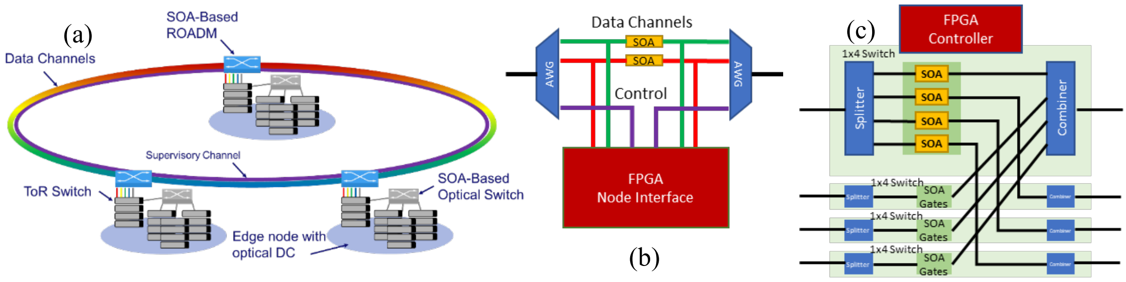

Fig 2. – (a) Proposed optical ring network interconnecting edge data center nodes, (b) SOA-based ROADM and (c) SOA-based intra-datacenter optical switch (for simplicity only one 1x4 instance is shown in detail).

The computational resources are organized in racks. The interface to the ring network is a top-of-rack (ToR) switch and serves as a gateway for the metro-traffic. In every edge node, several racks are organized in clusters to form the edge data center. For the intra-datacenter communication, an SOA-based optical switch is used, as shown in Fig. 2(c). The switch works in a broadcast and select approach. The optical signals are copied in different optical paths via the splitters. SOAs placed in every path then gate the light according to the packet destination.

A scheduler protocol of requests and grants between the ToR switches and the switch controller is used to prevent packet contention. A previous work has shown 5 µs of server-to-server latency and has validated the usage of optical switches for communication between cluster of ToRs [8].

Results

Fig. 3(a) reports the simulated dynamics of the supervisory channel to control the network with 3 nodes. One node is the leader (green) and generates the control packets (highlighted in the red rectangles). The followers (yellow and orange traces) receive the packets, process them, set the SOA-based ROADM for dropping/adding traffic to/from the node and further transmits the modified control to the ring.

Fig. 3(b) shows the traces for the transmission of one follower node after processing a control packet. In this case, the control packet dictates the behavior of the following 4 slots for 2 channels, which can be seen in the picture. The slots illustrate different possibilities of transmission: no channels used, one of the channels, and both channels. Between the slots, a reconfiguration state allows the reconfiguration of the SOAs.

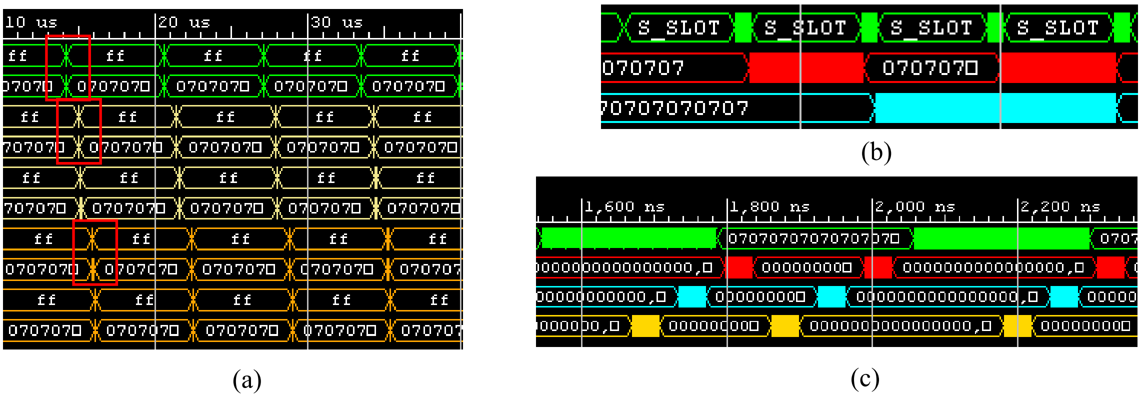

Fig 3. – Data and control simulated FPGA traces in different scenarios in the network: (a) forwarding of control packets, (b) data slots in two channels and (c) distribution by gateway ToR of incoming time-slotted data to different optical ports.

Fig. 3(c) shows the forwarding of different Ethernet frames contained in a data slot to different ToRs in the cluster (red, blue and yellow), in a scenario containing 4 ToRs. The ToR receives the traffic from the ring network and process it as Ethernet packets in its electronic switch, forwarding to others ToRs in the network according to the MAC addresses contained in the Ethernet frames.

Conclusions

We reported the system concept of an optical time-slotted network for edge data centers using FPGAs as network controllers. We presented simulations showing the distribution of Ethernet frames inside a timeslot for different ToRs when it arrives at an edge node. Further investigation is required to optimize the usage of the supervisory channel in terms of the duration of the control packets and the number of slots reserved by them.

References

[1] 5G Americas, "5G at the Edge," White paper (2019), https://www.5gamericas.org/5g-at-the-edge/
[2] N. Benzaoui, S. Sahoo and S. Bigo, "Deterministic latency networks: the enabler of edge data center synchronous operation [Invited]," in Journal of Optical Communications and Networking, vol. 13, no. 9, pp. D115-D125, September 2021
[3] N. Benzaoui et al., "Deterministic Dynamic Networks (DDN)," in Journal of Lightwave Technology, vol. 37, no. 14, pp. 3465-3474, 15 July15, 2019.
[4] N. Calabretta, W. Miao, K. Mekonnen, and K. Prifti, "SOA Based Photonic Integrated WDM Cross-Connects for Optical Metro-Access Networks," *Applied Sciences*, vol. 7, no. 9, p. 865, Aug. 2017.
[5] M. Szczerban et al., "Real-time control and management plane for edge-cloud deterministic and dynamic networks," in Journal of Optical Communications and Networking, vol. 12, no. 11, pp. 312-323, November 2020.
[6] B. Pan, F. Yan, X. Guo and N. Calabretta, "Experimental Assessment of Automatic Optical Metro Edge Computing Network for Beyond 5G Applications and Network Service Composition," in Journal of Lightwave Technology, vol. 39, no. 10, pp. 3004-3010, 15 May15, 2021.
[7] "IEEE Standard for Ethernet," in IEEE Std 802.3-2015 (Revision of IEEE Std 802.3-2012) , vol., no., pp.1-4017, 4 March 2016.
[8] X. Xue et al., "Experimental Assessments of Fast Optical Switch and Control System for Data Center Networks," 2021 Optical Fiber Communications Conference and Exhibition (OFC), 2021, pp. 1-3.