Demonstration of Low-Latency ETH-switched DataCenter and 5G Fronthaul Networks Using the 1024-port Hipoλaos Optical Packet Switch

George Giamougiannis[(1)], Apostolos Tsakyridis[(1)], Nikos Terzenidis[(1)], Dimosthenis Spasopoulos[(1)], Christos Vagionas[(1)], Nikos Pleros[(1)]

(1) Department of Informatics, Center for Interdisciplinary Research & Innovation, Aristotle University of Thessaloniki, Thessaloniki, Greece

giamouge@csd.auth.gr

Abstract *We experimentally evaluate the performance of a 1024-port Hipoλaos OPS with Ethernet traffic in real-world DC and 5G fronthaul testbeds, successfully demonstrating video transmission and iperf communication through low-latency optical switching at 10Gb/s per port.*

Introduction

The major driving forces for the evolution of the Ethernet switching landscape are currently identified as the Data Center (DC) networks, along with the emerging 5G fronthaul architectures. Traditional server-centric DC architectures but even more the emerging DC resource disaggregation paradigm[1] are constantly pushing the performance envelope on the switching infrastructure, with ultra-low latency, multi-Tb/s capacity and high-radix connectivity considered as critical requirements. At the same time, high-performance switch infrastructures are also penetrating the access segment, with edge computing and 5G Ethernet-switched fronthaul networks being expected to skyrocket the number of devices connected to IP networks, while necessitating end-to-end latency performance in the msec-range[2].

In this realm, switching continues to be the stronghold of electronic ASIC-based solutions that recently managed to allow for up-to 25.6Tb/s[3]-[5] aggregate switch capacity. However, with latency and high-radix connectivity values extending to new boundaries, it remains still an open and highly challenging problem on whether and how electronics can meet next-generation switching targets. To this end, significant efforts have been already devoted by the research community towards optical switch technologies[6]-[10] that would allow to address next-generation latency, radix, capacity and energy requirements in both the DC and 5G fronthaul network segment. Optical Circuit Switch (OCS) designs can definitely offer high-radix connectivity, but their millisecond-scale switching times[11],[12] renders them inappropriate for packet-switched environments. On the other hand, Optical Packet Switch (OPS) architectures have witnessed a rapid progress during the last years, allowing for full-scale deployments that offer both low-latency switching and high-port counts[13],[14]. Along this line, we have demonstrated the Hipoλaos OPS architecture[15]-[18] that was shown to operate successfully with sub-μsec latency performance in both 256- and 1024-radix layouts at 10Gb/s[17],[18] and, more recently, at 25Gb/s port line-rates[15]. However, only few OPS systems managed to be validated in real DC and/or 5G fronthaul application environments performing with real Ethernet traffic, like the one described in[6].

In this paper, we take the next step in the Hipoλaos switch validation chain and

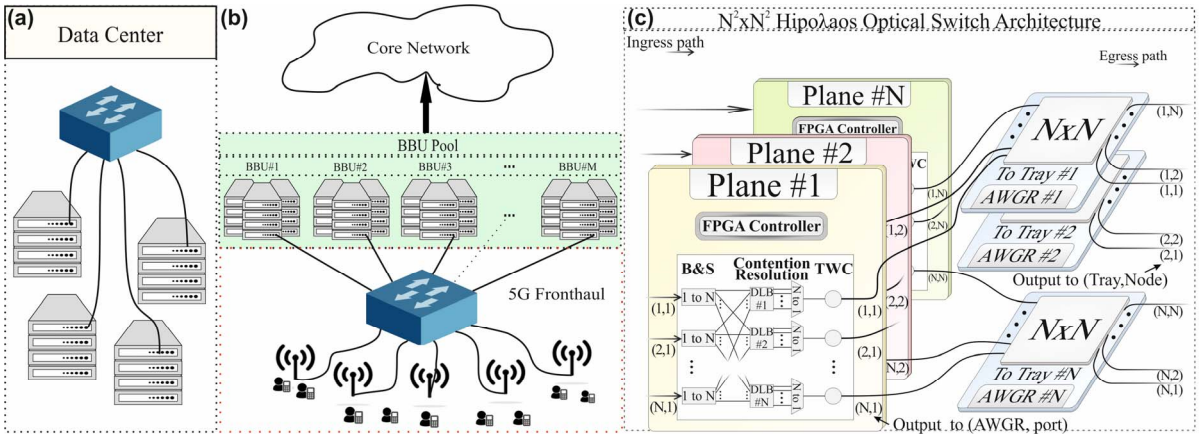

Fig. 1: (a) Intra-DC interconnection of servers by Hipoλaos OPS represented by the blue switch symbol, (b) Incorporation of Hipoλaos OPS in a 5G fronthaul network, (c) Layout of N^2xN^2 Hipoλaos OPS architecture

demonstrate for the first time its 10Gb/s Ethernet operation in both real DC and 5G fronthaul application environments. Intra-DC switch operation, as shown in Fig. 1(a), has been validated by connecting enterprise DC servers through a 10.24Tbps capacity Hipoλaos OPS. 5G Ethernet-switched fronthauling, where the switch has to route Ethernet traffic from a Baseband Unit (BBU) pool to the mobile users via Remote Radio Heads (RRHs), as shown in Fig. 1(b), was evaluated by extending the intra-DC testbed with two V-band antennas, incorporating a mmWave 5G Digital Radio Over Fiber (D-ROF) wireless link in between the fiber-based connection of the Hipoλaos outgoing port and the destination server. In both cases, a full-fledged Plane of the 1024-port Hipoλaos layout, followed by a 32x32 Arrayed Waveguide Grating Router (AWGR) device was experimentally established, as depicted in Fig. 1(c). Network performance and stability over time validation were confirmed through several measurements carried out during an HD video transmission application and when utilizing the iperf application software, achieving a sub-msec end-to-end (Layer 7) latency performance, with the Hipoλaos' contribution to the total end-to-end latency not exceeding 250ns.

Hipoλaos OPS architecture

The generic architecture of a N^2xN^2 Hipoλaos OPS layout is shown in Fig. 1(c). It comprises #N Planes with #N ports each, performing header processing and data forwarding through a Broadcast-and-Select scheme and featuring a contention resolution stage with Shared Delay Line Banks (S-DLBs) and Tunable Wavelength Converters (TWCs). Every Plane hosts an FPGA controller that performs header processing and forwarding functions. As its last stage, Hipoλaos employs #N NxN AWGR devices in order to offer collisionless forwarding via wavelength routing. A more detailed description of the Hipoλaos architecture and its constituent subsystems can be found in[19].

Experimental evaluation

A Hipoλaos OPS prototype being similar to the one reported in[18] has been deployed and incorporated in our intra-DC and 5G fronthaul testbed, with the respective experimental setups being shown in Fig. 2(a) and 2(b), respectively. In both cases 10 Gb/s Ethernet traffic is produced by Server 1, gets transmitted through an SFP+ module at λ_{serv}=1552,51nm and enters the Hipoλaos input. Then, the Ethernet traffic is routed through the Hipoλaos switch, realizing the interconnection scenarios depicted in Fig. 1(a),(b). A Xilinx FPGA was employed at the Hipoλaos plane to handle the switching functionalities. The Ethernet packets are injected to an EDFA for amplification and filtered by a 5nm Optical Bandpass Filter (OBPF), before entering the first stage of the Hipoλaos OPS, where a 1/32 splitter was used to emulate the actual losses of a 32x32 Plane within the 1024x1024 switch. After a second stage of amplification and filtering, the Ethernet traffic is divided in two identical signals via a 50/50 splitter, feeding ports D and E of the differentially-biased[20]-[22] SOA-MZI#1 and serving as control signals. A WDM signal is injected into port G, formed by multiplexing 3 CW signals via an Arrayed Waveguide Grating (AWG), whereas a CW beam at λ_4 was injected into port H to complete the differential biasing of the SOA-MZI. The wavelength-converted Ethernet packets exiting SOA-MZI#1 from port C and travel through the S-DLB stage, an EDFA and a filter until they split again to feed ports A and H of SOA-MZI #2, as control signals. An auxiliary CW beam at λ_4 is injected to port D, while a Tunable Laser Source (TLS) controlled by the FPGA generates a CW beam at λ_N, with λ_N

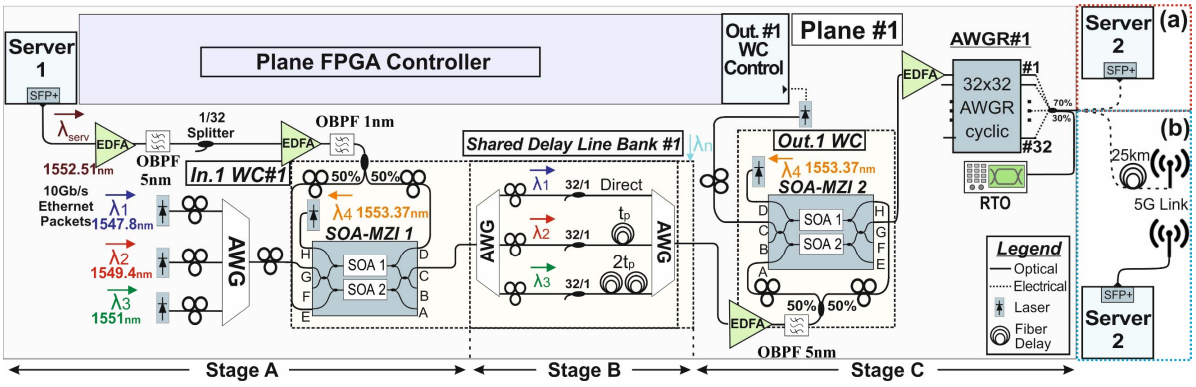

Fig. 2: Experimental setup for the evaluation of Hipoλaos layout, followed by a 32x32 AWGR. Ethernet traffic from Server 1 routed to (a) Server 2 for intra-DC interconnection emulation (in red frame), (b) to Server 2 via a 25km fiber and a V-band wireless transmission (in blue frame).

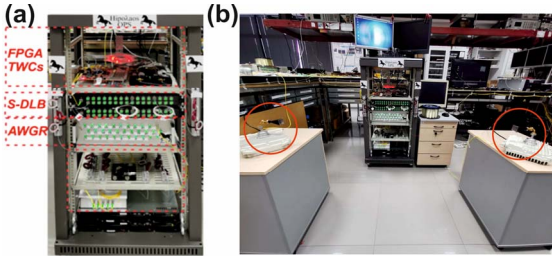

Fig. 3: (a) Hipoλaos OPS intra-DC testbed, (b) 5G fronthaul testbed with V-band antennas.

depending on the destination node and being launched as input at port C of SOA-MZI#2. The output signal, after amplification, is injected to the first port of the AWGR. Depending on the value of λ_N, the Ethernet traffic exits through the respective AWGR output and gets then split via a 70/30 coupler to allow both for monitoring the Ethernet packets on a Real Time Oscilloscope and to forward the traffic to their destination server. In the intra-DC testbed, shown in Fig. 2(a), Server 2 (destination server) connects directly to the 70%-branch of the coupler. In the 5G fronthaul testbed, shown in Fig. 2(b), the Ethernet packets coming out of the Hipoλaos OPS travel through a 25km long fiber until they reach the Antenna site, where a 5 Gb/s wireless transmission link was established between two V-band (70GHz) antennas in order to emulate a scenario where Server 2 acts as a mobile user. In both scenarios, the SFP+-Rx port of Server 1 was hard-wired to the SFP+-Tx port of Server 2 in order to allow bidirectional communication and successful protocol operation. Fig. 3(a) shows the Hipoλaos OPS prototype, while the 5G fronthaul testbed including the Hipoλaos OPS rack is shown in Fig. 3(b).

The experimental results are illustrated in Fig. 4. Specifically, Fig. 4(a) depicts the eye diagram, trace and spectrum of the Ethernet traffic generated by Server 1 when transmitting an HD video file to Server 2. Fig. 4(b)-(e) illustrate the respective eye diagrams, traces and spectrums obtained when Server 2 is connected at four different Hipolaos output ports (AWGR ports #1, #9, #19, #32) within the intra-DC testbed, corresponding to different Hipoλaos switch configurations. Fig. 4(f) illustrates respective results when the 5G fronthaul testbed is employed and Server 2 is connected to the Hipoλaos output port after incorporating a 25km-long fiber link and a 2.5m long 70GHz wireless link. Uninterrupted high-quality video transmission was obtained in all cases for both the intra-DC and the 5G fronthaul testbed evaluations, with Fig. 4(g) and 4(h) depicting a snapshot of the video running at the Server 2 site.

Following the successful video transmission through different Hipoλaos switching states, the 5G fronthaul was evaluated via an iperf application that was executed between Servers 1 and 2. The latency measurement recorded a mean roundtrip time of 0.252ms, including the latency contributions from the operating system, the network interface, the propagation roundtrip time etc., with the Hipoλaos latency contribution not exceeding 250nsec. Fig. 4(i) reveals the stability of the whole system when exchanging iperf messages for a duration of 2 hours, achieving a constant 4.89 Gb/s throughput that is identical to the maximum data rate supported by the 70GHz antennas used.

Conclusion

We experimentally demonstrate Ethernet traffic transmission at 10Gb/s through the 1024-port Hipoλaos OPS on realistic DC and 5G fronthaul testbeds.

Acknowledgements

This work was supported by the EU projects 5G-PHOS (761989) and NEBULA (871658).

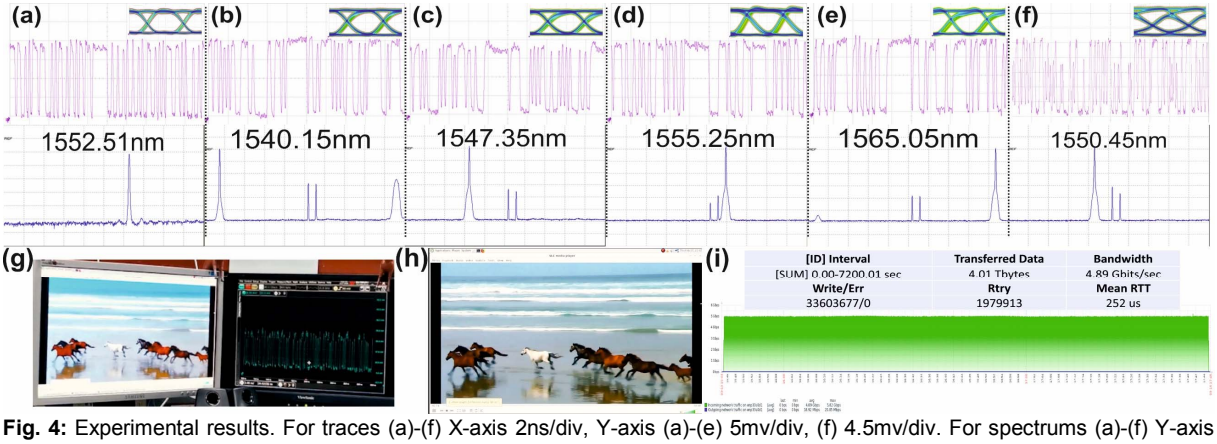

Fig. 4: Experimental results. For traces (a)-(f) X-axis 2ns/div, Y-axis (a)-(e) 5mv/div, (f) 4.5mv/div. For spectrums (a)-(f) Y-axis 10dB/div. For graph (i) X-axis 2min/div, Y-axis 0.5Gb/s/dev. (a) Data trace, eye diagram and spectrum of Hipoλaos OPS Input (Server 1 Output). Intra-DC scenario: (b)-(e) Data traces, eye diagrams and spectrums at 4 AWGR Outputs (#1, #9, #19, #32). 5G fronthaul scenario: (f) Trace, eye diagram and spectrum of data reaching the antenna site, (i) iperf measurements and 2-hour network connection graph, (g),(h) snapshots of a high quality video transmitted by Server 1 and captured by Server 2 during the experimental demonstration.

References

[1] Y. Cheng, R. Lin, M. De Andrade, L. Wosinska, J. Chen, "Disaggregated Data Centers: Challenges and Tradeoffs," IEEE Communications Magazine, (2020).

[2] https://www.samsung.com/global/business/networks/insights/news/5g-is-now-part-1-2018-the-year-of-5g/

[3] https://www.broadcom.com/products/ethernet-connectivity/switching/strataxgs/bcm56990-series

[4] https://www.nextplatform.com/2020/05/11/innovium-stays-on-broadcoms-heels-with-teralynx-8-switch-chips/

[5] ”Barefoot Networks Unveils Tofino™ 2, the Next Generation of the World's First Fully P4-Programmable Network Switch ASICs,” (2018)

[6] P.Bakopoulos et al., “NEPHELE: An End-to-End Scalable and Dynamically Reconfigurable Optical Architecture for Application-Aware SDN Cloud Data Centers”, IEEE Communications Magazine (2018)

[7] K. Ueda et al., "Large-Scale Optical Switch Utilizing Multistage Cyclic Arrayed-Waveguide Gratings for Intra-Datacenter Interconnection," PJ. 9, 1-12 (2017)

[8] J. Gripp et al., "Photonic terabit routers: The IRIS project," in Proc. Conference on Optical Fiber Communication (OFC), San Diego, CA (2010)

[9] O. Liboiron-Ladoucer et al., “The Data Vortex Optical Switched Interconnection Network,” J. Lightw. Technol., vol. 26, no. 13, pp. 1777-1789 (2008)

[10] M. Moralis-Pegios et al., 4-channel 200 Gb/s WDM O-band silicon photonic transceiver sub-assembly," Opt. Express 28, 5706-5714 (2020)

[11] Polatis, SERIES 7000 - 384x384 port Software-Defined Optical Circuit Switch, 2016 [Online]. Available: https://www.polatis.com/series-7000-384x384-port-software-controlled-optical-circuit-switch-sdn-enabled.asp?

[12] “S Series Optical Circuit Switch” | CALIENT Technologies", Calient.net (2018). [Online]. Available: http://www.calient.net/products/s-series-photonic-switch/.

[13] M. Ding et al., “Hybrid MZI_SOA InGaAs/InP Photonic Integrated Switches*," IEEE J. Sel. Top. Quant.*, vol.24, pp. 1-8 (2018)

[14] L.Qiao et al., "32 × 32 silicon electro-optic switch with built-in monitors and balanced-status units*," Sci. Rep.*, vol. 7, pp. 42306 (2017)

[15] N. Terzenidis et al., “A 25.6 Tb/s capacity 1024-port Hipoλaos Optical Packet Switch Architecture for disaggregated datacenters”, OFC (2020)

[16] A. Tsakyridis et al., “25.6 Tbps capacity and sub-μsec latency switching DataCenters using >1000-port Optical Packet Switch Architectures” Accepted at JSTQE (Invited) (2020)

[17] N. Terzenidis et al., "High-port low-latency optical switch architecture with optical feed-forward buffering for 256-node disaggregated data centers," Opex 26, 8756-8766 (2018)

[18] M. Moralis-Pegios et al., "A 1024-Port Optical Uni- and Multicast Packet Switch Fabric," *JLT 37,* 1415-1423 (2019)

[19] N. Terzenidis et. al., “High-port and low-latency optical switches for disaggregated data centers: The Hipoλaos switch architecture,” JOCN 10, 7, B102-B116 (2018)

[20] A. Tsakyridis, et. al., “10 Gb/s optical random access memory (RAM) cell,” Opt. Lett. 44, 1821-1824 (2019).

[21] G. Mourgias-Alexandris, et. al., "An all-optical neuron with sigmoid activation function", Opt. Express 27, 9620-9630 (2019).

[22] A. Tsakyridis et. al, “Theoretical and experimental analysis of Burst-mode Wavelength Conversion via a Differentially -biased SOA-MZI”, Journal of Lightwwave Technology in early access. 10.1109/JLT.2020.2995471 (2020).