SDN-Enabled Reconfigurable Optical Data Center Network with Automatic Network Slicing to Provision Dynamic QoS

Xuwei Xue[(1)], Bitao Pan[(1)], Fernando Agraz[(2)], Albert Pagès[(2)], Xiaotao Guo[(1)], Fulong Yan[(1)], Salvatore Spadaro[(2)], Nicola Calabretta[(1)]

[(1)] IPI-ECO Research Institute, Eindhoven University of Technology, Netherlands. x.xue.1@tue.nl
[(2)] Universitat Politècnica de Catalunya, Barcelona, Spain

Abstract *An SDN-controlled optical DCN enabling automatic QoS-driven network-slice provisioning, reconfiguration and flow-priority updating is demonstrated for the first-time. Network-slice reconfigurations with guaranteed QoS based on monitored statistics are achieved within 125 ms with 3.485 μs server-to-server latency and zero packet-loss at 0.7 load for high-priority.*

Introduction

The traffic-boosting applications emerging with the 5G related services, cloud computing and internet of things have imposed stringent requirements to the data centers (DCs) in terms of high capacity, low latency and flexible resource utilization[1]. To satisfy these requirements, the DC networks (DCNs) are expected to evolve towards new architectures and switching technologies to upturn their network performance. Recently, we proposed and numerically investigated OPSquare, an optical DCN architecture based on fast Semiconductor Optical Amplifier (SOA)-based switches and Optical Flow Control (OFC) protocol aiming to solve the packet contentions[2, 3]. Benefitting from the parallel intra-/inter-cluster switching capabilities, the proposed optical DCN features high capacity and low latency performance.

The plethora of applications with heterogeneous traffic needs, which are hosted in the DC, impose their own set of requirements, particularly for bandwidth, packet loss and latency, to the DCN infrastructure. Hence, such infrastructure has to offer tailored, guaranteed and dynamic quality of service (QoS). Network slicing (NS) of the optical DCN is a promising solution to deal with these challenges, forming a complete instantiated logical network to meet certain network characteristics required for a specific application[4]. The virtual sliced network units need to be flexibly reconfigured to provide dynamic QoS to different tenants in a DC. Benefiting from the flexible use of the data plane infrastructure, network slicing can significantly improve resource utilization.

In this work, we experimentally demonstrate for the first time the full operation of optical OPSquare DCN, empowered by a software-defined networking (SDN) control and orchestration plane for reconfigurable NS deployment with dynamic QoS guarantees. The data plane switches the traffic flow at nanoseconds scale, benefitting from the fast configuration speed of SOA switches. The SDN control and orchestration plane collects the resource topology of the connected data plane and then accordingly computes the optimized virtual network functions (VNFs) allocation and path interconnections to provide the NS provisioning service. Meanwhile, the forwarding priority is flexibly assigned to the traffic flows associated with each NS, according to the latency requirements of deployed applications. Moreover, the statistics (counts of retransmitted and lost packet) of data plane are monitored and reported to the Orchestrator to trigger the reconfiguration of the NS connection to reduce packet loss, thus achieving the required QoS.

SDN-enabled OPSquare DCN

The SDN-enabled OPSquare architecture

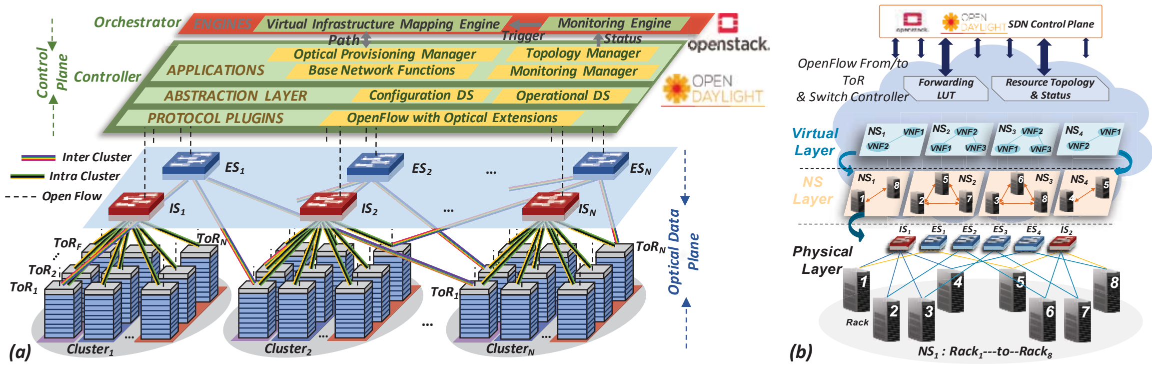

Fig. 1: (a) Reconfigurable SDN-enabled optical DCN. (b) NS allocation over infrastructure.

comprising the distributed SOA-based switches as well as the SDN controller and orchestrator is illustrated in Fig. 1(a). The *N* *N*×*N* intra-cluster optical switches (ISs) and *N* *N*×*N* inter-cluster optical switches (ESs) interconnect the intra- and inter-cluster top of racks (ToRs), respectively. FPGAs are used to implement the ToRs and the switch controllers of the IS/ES. The implemented OFC protocol [3] between ToRs and IS/ES controllers solves packet contentions at the switch nodes. In case of packet contention, the packets assigned with higher priority are directly forwarded by the switches while the packets with lower priority will be retransmitted. The flow control signals (ACK/NACK) generated at the IS/ES controllers are sent back to the ToRs for releasing the packets or requesting packet retransmission.

As shown in Fig. 1(a), the FPGAs-based ToRs and switch controllers collect and report the infrastructure resources to the OpenDaylight (ODL) based SDN controller through OpenFlow (OF) Agents with extended OF protocol [5, 6]. The data plane layout and physical distribution information is stored in the Topology Manager (TM). The Optical Provisioning Manager (OPM) module has been developed to configure the underlying devices (i.e., IS, ES and ToR) required to set up the specific network connectivity for NS deployment. The OPM cooperates with the Path Computation Engine (PCE) of the OpenStack-based Orchestrator, which relies on abstracted topological information from the TM to provide the controller with a ToR-to-ToR path and flow priority computation service. The SDN controller then provides a set of actions to the ToRs and switch controllers in order to configure the NS (updating the look-up tables), as well as to assign flow priority per NS. Furthermore, the Monitoring Manager (MM) collects optical data plane statistics (counts of retransmitted and lost packets) and aggregates them into NS level metrics. Such aggregated information is collected by the Monitoring Engine (ME) of the Orchestrator to trigger the NS reconfiguration whereby maintaining the expected QoS. Fig. 1(b) shows the concept of NS allocation, where each NS comprises several VNFs based on the functional characteristics and server availability.

Experimental demonstration and results

The experimental demonstration of the SDN-controlled and orchestrated OPSquare DCN is shown in Fig 2 (a). It consists of 8 FPGA-based ToRs equipped with OF Agents grouped in 2 clusters, and 4 SOA-based ESs and 2 SOA-based ISs with corresponding FPGA-based switch controllers and OF Agents. Each FPGA-based ToR interconnects the SPIRENT at 10 Gb/s that generate Ethernet frames with variable and controllable load. An ODL-based SDN controller is placed over the data plane to control the configuration of the infrastructure. The OpenStack-based Orchestrator implements the engines responsible for guaranteeing the performance of the NSs.

1. NS provisioning and flow priority assignment

First, the NS provisioning and flow priority assignment are demonstrated to validate the programmability of OPSquare DCN. Based on the monitored infrastructure topology as shown in Fig. 2(b), the PCE assigns NS_1, NS_2, NS_3 and NS_4 over the demonstrated infrastructure, where VNF_1 and VNF_2 of NS_1 are racks 1 and 8, respectively. Afterwards, the OPM coordinates with the PCE to configure the best path among racks as illustrated in Fig. 2(c) to provide connectivity for the VNFs. In this case, the selected paths between ToR_1 and ToR_8 for NS_1 is $ToR_1 \leftrightarrow IS_1 \leftrightarrow ToR_4 \leftrightarrow ES_2 \leftrightarrow ToR_8$.

Meanwhile, the Orchestrator also computes the flow priority of NSs based on the latency requirements of the deployed applications. Upon the computed priority order, the OPM allocates the flow priority to each NS by the extended OF protocol as shown in Fig. 2(d). The proposed sliceable DCN provides 4 classes of priority to each NS. NS_1 hosting the latency-sensitive applications is assigned with the highest priority 1, while the lowest priority 4 is allocated to NS_2 where applications are latency insensitive.

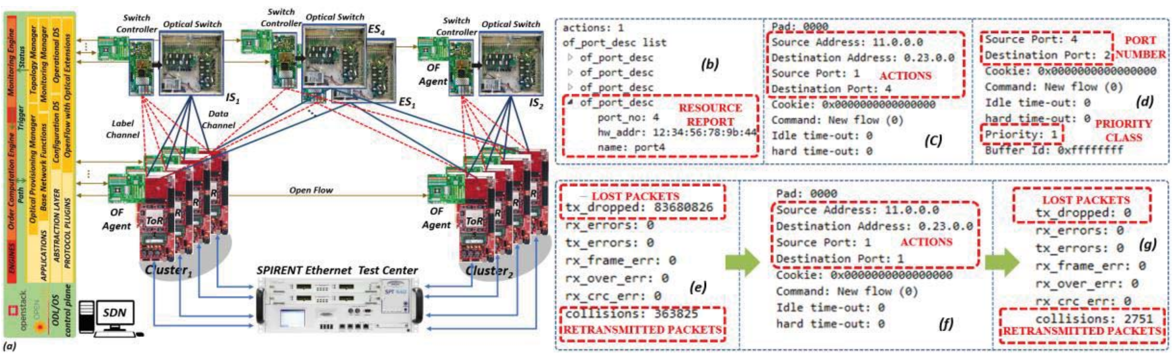

Fig. 2: (a) Experimental demonstration setup. (b), (c), (d), (e), (f), (g) Open flow commands to configure the data plane.

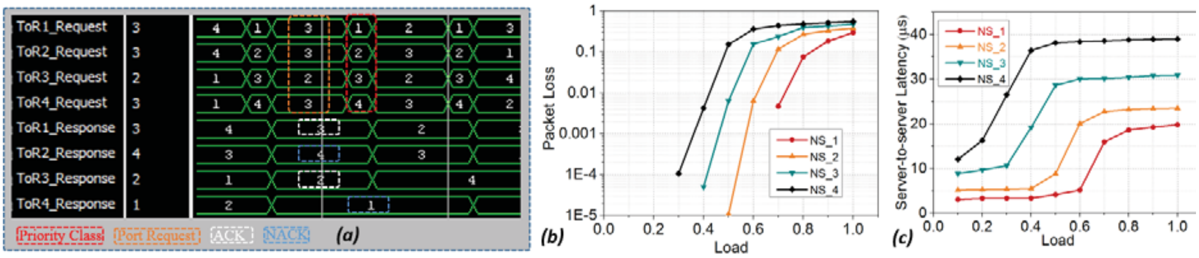

Fig. 3: (a) Flow control signals monitored switch controller. (b) Packet loss (c) latency for 4 NSs with specified priority.

Following this rule, NS_3 and NS_4 are assigned with priority 2 and 3, respectively. When contention happens, flows in NS_1 with highest priority will be forwarded directly.

2. NS reconfiguration triggered by statistics monitoring

After the NSs provisioning, the SDN controller collects the data plane statistics (the packets re-transmitted due to contention, and the packets lost at the ToRs due to buffer overflow) from the ToRs and switch controllers. Fig. 2(e) shows the monitored statistics. Once the ME detects that the packets losses at ToRs (the count of lost packets is 83680826 in this case as shown in Fig. 2(e)) are over a preventive packet loss threshold of 1E-5, ME automatically triggers the NS reconfiguration through an alternative path, to maintain the requested QoS. For instance, the OPM coordinates with the PCE to calculate the new network connectivity for NS_1. The new connectivity, $ToR_1 \leftrightarrow ES_1 \leftrightarrow ToR_5 \leftrightarrow IS_2 \leftrightarrow ToR_8$, is then reconfigured. Fig.2 (f) shows the new path configuration information at the flow used to update set of actions inside the ToRs and the switch controllers. The NS reconfiguration is completed through the new end-to-end connection within 125 ms which is less than the application switchover time. Once the new connections are reconfigured, the packet loss will be below the threshold and there will be less retransmitted packets as shown in Fig. 2(g).

3. OPSquare DCN performance assessment

OFC protocol deployed between the ToRs and switch controllers is leveraged to solve the packets contention caused packet loss. Fig. 3(a) illustrates the monitored OFC signals at switch controller of IS_1. At every time-slot, the connected ToRs send the NS_Request signal to the controller, indicating the port destination and priority of the optical packet. Once packet contentions happen, the switch controller sends ACK/NACK signals to the corresponding ToRs. As shown in Fig. 3(a), for packets destining to ToR_3 (switch port-3) from the other ToRs, ToR_1 (allocated highest priority) receives an ACK (NS_Response=NS_Request) signal to release the stored packet, while ToR_2 and ToR_4 get NACK signals (NS_Response≠NS_Request) to trigger the retransmission of the conflicted packets.

PCE cooperating with OPM assign NS_1, NS_2, NS_3 and NS_4 with different flow priority based on the QoS requirements. Fig. 3(b) and (c) indicate the packet loss and server-to-server latency for all the considered NSs with specific flow priority. The results confirm no packet loss for NS_1 up to load of 0.7, while NS_4 with the lowest priority achieves the acceptable packet loss (0.04) at the load of 0.4. The retransmissions for the blocked packets from NS with lower priority lead to the extra retransmission delay. On the contrary, the packets with highest priority (NS_1) with no retransmissions achieves a minimum latency of around 3.485 μs, which includes the 3.181 μs of buffer queuing time at the ToRs, 140 ns delay on the 2 × 14 m fiber as well as the FPGA processing time at TX and RX (163.5 ns). NS_4 presents a max latency of 38.79 μs even at high load, which is compatible with applications with best-effort QoS requirement.

Conclusions

The full operation of SDN controlled OPSquare DCN enabling network slicing with dynamic QoS guarantees is demonstrated for the first time. For the data plane, SOA-based switches with nanoseconds configuration time are deployed to interconnect racks, and OFC protocol is implemented to solve packet contention. For the control plane, NS provisioning is automatically supplied by the SDN controller based on the collected resource topology of data plane. Meanwhile, the control plane assigns the flow priority to each NS according to the applications' latency requirement. The monitored statistics of data plane automatically triggers the dynamic NS reconfiguration within 125 ms to decrease packet loss. Results show that packets with the highest priority have high QoS, achieving zero packet loss and 3.485 μs ToR-to-ToR latency at the load of 0.7.

Acknowledgements

The authors would like to thank the Olympics (ESTAR17207) and QAMeleon (780354) project for partially supporting this work.

References

[1] Cisco, "Forecast and Methodology 2016-2021. White Paper. Cisco Systems," ed: Inc, 2018.

[2] F. Yan, B. Pan, and N. Calabretta, "Flexibility assessment of the reconfigurable OPSquare for virtualized data center networks under realistic traffics," *in 2018 European Conference on Optical Communication (ECOC)*, 2018, pp. 1-3: IEEE.

[3] W. Miao, S. Di Lucente, J. Luo, H. Dorren, and N. Calabretta, "Low latency and efficient optical flow control for intra data center networks," *Optics express*, vol. 22, no. 1, pp. 427-434, 2014.

[4] K. Kondepu, C. Jackson, Y. Ou, A. Beldachi, A. Pagès, F. Agraz, F. Moscatelli, W. Miao, V. Kamchevska, and N. Calabretta, "Fully SDN-enabled all-optical architecture for data center virtualization with time and space multiplexing," *IEEE/OSA Journal of Optical Communications Networking*, vol. 10, no. 7, pp. 90-101, 2018.

[5] X. Xue, F. Wang, F. Agraz, A. Pagès, B. Pan, F. Yan, X. Guo, S. Spadaro, and N. Calabretta, "SDN-controlled and Orchestrated OPSquare DCN Enabling Automatic Network Slicing with Differentiated QoS Provisioning," *Journal of Lightwave Technology*, vol. 38, no. 6, pp. 1103-1112, 2020.

[6] F. Wang, F. Agraz, A. Pages, B. Pan, F. Yan, S. Spadaro, and N. Calabretta, "Experimental assessment of SDN-enabled reconfigurable OPSquare data center networks with QoS guarantees," *in 2019 Optical Fiber Communications Conference and Exhibition (OFC)*, 2019, pp. 1-3: IEEE.