DEMONSTRATION OF 100 GBIT/S ACTIVE MEASUREMENTS IN DYNAMICALLY PROVISIONED OPTICAL PATHS

Jorge E. López de Vergara[1,2]* , Mario Ruiz[2], Lluís Gifre[2], Marc Ruiz[3], Luis Vaquero[1,2],
José F. Zazo[1], Sergio López-Buedo[1,2], Óscar González de Dios[4], Luis Velasco[3]

[1]Naudit High Performance Computing and Networking, S.L., Madrid, Spain
[2]Universidad Autónoma de Madrid, Madrid, Spain
[3]Universitat Politècnica de Catalunya, Barcelona, Spain
[4]Telefónica GCTO, Madrid, Spain
*E-mail: jorge.lopezdevergara@naudit.es

Keywords: Monitoring and data analytics, network probe, optical path capacity, network delay, packet loss.

Abstract

New techniques, based on Software-Defined Networks, are used to deploy optical paths dynamically. This demonstration shows how active measurements at 100 Gbit/s are performed to check before operation that the performance requirements are met in terms of capacity, delay or packet loss.

1 Introduction

The H2020 Metro-Haul project [1] aims at defining the metro optical network architecture to interconnect incoming 5G access networks with the core network. For this, it is developing new techniques, based on Software-Defined Networks (SDN), to provision optical paths dynamically. This metro network has to provide Key Performance Indicators (KPI) such as capacity, delay or packet loss, with very stringent requirements.

In order to check that the deployed optical paths meet these requirements, this demonstration shows how active measurements at 100 Gbit/s are performed to check that the requested KPIs are met before the optical path is put in operation. Given the foreseen heterogeneity in the optical network, with different equipment providers, these preliminary measurements are essential for the network operator to assure the QoS (Quality of Service) provided to the users, and complement passive measurements performed during the optical path lifetime [2].

In the demonstration, integration of SDN controllers with a Monitoring and Data Analytics (MDA) controller, and active monitoring devices will be shown. This demonstration exhibits novel innovations, by adding active measurements to the MDA capabilities of optical SDN-based network scenarios. This combination can help network providers to operate 5G optical networks. Hence, it will be of interest to a broad ECOC audience, attracting the community's attention.

The rest of this paper is structured as follows: First, we provide a general overview of the architecture proposed in the Metro-Haul project, focusing on the deployment of new optical paths and testing them. Next, we explain how the network measurements are implemented. Then, we propose the demonstration workflow to be presented. Finally, we conclude with the main ideas here provided.

2 Architecture

This demonstration focuses on an infrastructure where a multi-layer network interconnects Central Offices (CO). In the control, orchestration, and management (COM) plane, a Network Function Virtualisation (NFV) orchestrator (NFVO) deals with the deployment of NFV Network Services (NS). The multi-layer interconnection network is controlled by a hierarchical SDN architecture, where a parent SDN controller is on top of per-layer SDN controllers, one for the packet layer and another one for the optical layer. Finally, an MDA controller [3] is in charge of collecting monitoring data from the underlying infrastructure and applying data analytics techniques on these data, while generating useful notifications of detected events and issuing recommendations to the SDN controller for network automation [4] (see Fig. 1).

In particular, one of the main aspects of network automation is to guarantee that provisioned connections actually meet the requested performance, e.g., in terms of end-to-end latency and throughput in the case of the packet layer. In that regard, this demonstration shows the use of active probes at the packet layer to measure the performance of packet layer connections. The active probes are installed in every CO and connected through a 100GE interface of a layer 2 (L2) switch in the internal CO network. Given that the active probes are connected to interfaces configured in trunk mode, the probes can tag the generated Ethernet frames with the desired VLAN ID, selecting the VLAN to be measured. The process of measuring a packet connection, consists in injecting a train of numbered and timestamped Ethernet frames, which arrive the other end of the connection, where the remote active probe loops back them. Following this procedure, packet circuit performance,
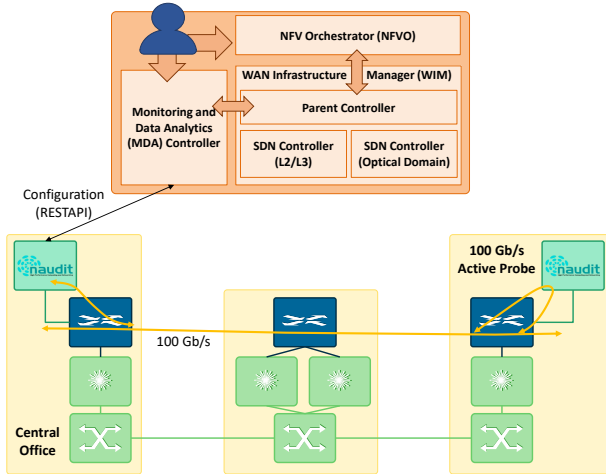
Fig. 1. Architecture of the demonstration scenario.

Fig. 2 Active network probe implemented in a small factor server. FPGA is placed in the bottom corner, with both QSFP28 interfaces connected. Management interfaces are on the left.

i.e., round-trip packet delay, jitter, packet loss and throughput, can be measured (see [5, 6] for details).

The active probes have been developed to be integrated in the above described COM system. To that end, they expose a REST-API-based northbound interface (NBI) through which the MDA controller can initiate a measurement session on a specific packet circuit. Due to its complexity, the demonstration is done remotely, connecting to the equipment deployed in a testbed in Spain that includes all the elements presented above. Fig. 2 shows how the probe looks like.

3 Active measurements

In order to achieve precise measurements at 100 Gbit/s, we have implemented the active probe in hardware, using a Field Programmable Gate Array (FPGA). To do so, a Virtex Ultrascale+ FPGA device is being used. In particular, the ADM-PCIE-9V3 High-Performance Network Accelerator card, which includes two 100GE interfaces, 8 GB of DDR4 memory and a XCVU3P-2-FFVC1517 FPGA. The QSFP28 physical cages are mapped directly to the FPGA. An integrated 100G Ethernet interface is in charge of communicating the FPGA side with the physical network. The FPGA internally works with a bus of 512-bit width of data and is clocked at 322 MHz (3.1 ns) to reach the needed throughput, even with smallest frames.

The packet train technique [6] has been implemented in the FPGA at 100 Gbit/s. Packets are timestamped at the transmitter and at the receiver. Receiver timestamps are useful to calculate the achieved throughput, whereas transmitter timestamps allow a precise calculation of round-trip delay with the clock implemented in the FPGA, without needing to synchronise the transmitter and receiver clocks.

The development is split into two independent designs, which reside in the same FPGA. In one hand, the transmitting side, a synthetic packet generator has been developed. On the other hand, the receiving side is in charge of filtering the packets, analysing them and generating a summary. In what follows, both elements are detailed.

3.1 Synthetic Packet Generator

This piece of hardware is written in Verilog and implemented as a Finite State Machine (FSM). In this way, the behaviour is completely deterministic and it provides the most accurate measurements. This module is in charge of generating UDP packets that will carry useful information for the measurement. The packets can be generated at the maximum throughput, which is extremely close to the theoretical value in 100G Ethernet links.

When a measurement is requested, some of the fields in the packets can be set, such as VLAN, source and destination IP addresses, source and destination ports, packet size, and Bit Error Rate Test (BERT) type used for the payload. Moreover, each UDP packet carries an `iperf` [7] compatible payload, with the following fields:

- Extended identifier (4 bytes).
- Local timestamp (8 bytes).
- Total amount of packets configured by the user (4 bytes).
- Burst identifier (1 bytes).
- BERT type (1 bytes).
- BERT payload, filling the rest of the packet size.

Additionally, other configurable parameters are the amount of packets in the burst and the inter-packet gap. All those configurations are done through an interface from a program running in the computer hosting the FPGA card.

The following BERT types have been implemented in the payload of the packets, so the active measurements can also be used to check if the optical modulation and the Forward Error Correction (FEC) implemented in the optical path are working properly:

- **PRBS** (Pseudo Random Binary Sequence): binary sequence that is difficult to predict and exhibits statistical behaviour similar to a truly random sequence.
- **All zeros**: A sequence of all zeros.
- **All ones**: A sequence of all ones.

- **1:1**: A sequence composed of alternating ones and zeros.
- **1:7**: Also referred to as "1 in 8". It has only a single one in an eight-bit repeating sequence.
- **2:8**: A pattern that contains a maximum of four consecutive zeros and only two ones.
- **3:24**: A pattern that contains the longest string of 15 consecutive zeros with only three ones.

3.2 Packet Filtering

In the receiver side, packets are timestamped in the moment they reach the FPGA side. The packet handler module is in charge of filtering packets by type. For instance, ARP and UDP packets are taken into account in this implementation. After that, UDP packets are parsed and those fields that are configurable are verified in order to check if the packets match with the required configuration. If so, the useful information is extracted to compute the packet train parameters. This information is fed to the statistics generator, which collects it. The results of the burst are then returned to the MDA once the burst has been completely received or after a timeout.

4 Demonstration workflow

The workflow of the demonstration is as follows (see Fig. 3):

1. The operator requests the deployment of a new NS through the NFVO's Graphical User Interface (GUI), which includes a number of interconnected Virtual Network Functions (VNFs), placed in different COs. This step triggers the set-up of a number of packet layer circuits, which the NFVO requests through the parent controller's NBI; the NFVO specifically selects to enable monitoring on those circuits. When the circuits are set-up, the parent controller sends a confirmation to the NFVO indicating the *circuit Id* that can be used for further actions.
2. The operator requests to measure the performance of the just created packet circuits through the MDA controller's GUI; to this end, the circuit is selected by its *circuit Id*. The MDA controller uses the *circuit Id* to ask the parent controller about the VLAN IDs that have been used for the circuits, which are next sent to the active probe to be used for tagging the generated Ethernet frames.
3. To measure the performance of a circuit, the active probe in a CO injects trains of Ethernet frames that are received by the active probe in the remote CO, which loops back them to the sender. Once the performance of a circuit is measured, the obtained values are replied. Such values are shown in the MDA controller' GUI to the operator that can in such way verify the performance of the set-up circuits.

Fig. 4 shows examples of the JSON messages exchanged between the MDA controller and the parent SDN controller, as well as the MDA controller and the active probe. As shown, for the latter case, the MDA controller request a measurement with some parameters, such as the reliability —six '9' will generate measurements with one million packets—, allocated bandwidth —measurement can be adjusted to values
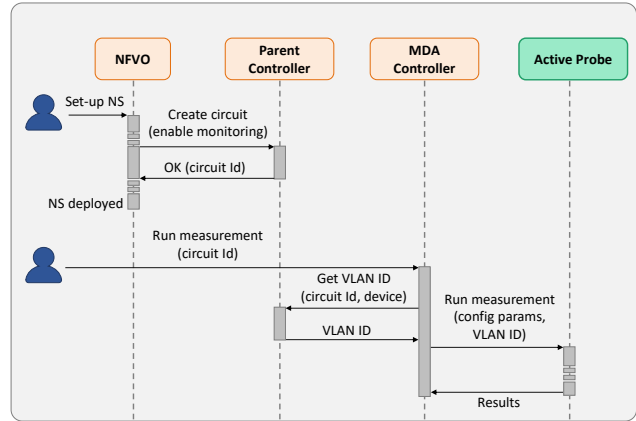

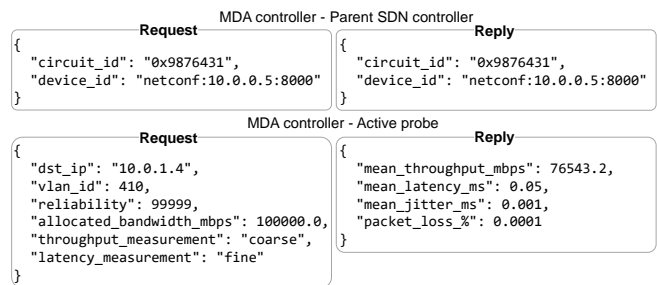Fig. 3. Workflow of the demonstration scenario.

Fig. 4 JSON messages exchanged in the measurement workflow.

below 100 Gbit/s if necessary—, and if the measurement is throughput or latency oriented —maximum transmission unit (MTU) packets will be used for fine throughput measurements, whereas minimum size packets will be used for fine latency measurements. By default, if not indicated, the used BERT type is PRBS. As a result, the probe provides the following values: achieved throughput in Mbit/s, latency and jitter in μs, and observed packet loss percentage.

5 Conclusions

In this demonstration, we show how the optical paths that are dynamically deployed in an SDN infrastructure can be checked with active measurements before they are put in operation. The use of an FPGA provides means to measure speeds of 100 Gbit/s with accuracy in the order of ns. As future work, we are working on the integration of this architecture in one of the use cases defined in Metro-Haul project, devoted to Real-Time Low-Latency Object Tracking and Security, where the dynamically provisioned optical circuits have to provide high throughput and low latency.

6 Acknowledgements

7 References

[1] Lord, A. 'Optical Metro Networks in a 5G world', 44th European Conference on Optical Communications (ECOC 2018), Rome, Italy, 22-23 September 2018.

[2] Velasco, L., Gifre, Ll., Izquierdo-Zaragoza, J.L., et al., 'CASTOR: An Architecture to Bring Cognition to Transport Networks', Proc. Optical Fiber Communications Conference and Exhibition 2018 (OFC 2018), San Diego, CA, USA, 11-15 March 2018, 10.1364/OFC.2018.Tu3D.5.

[3] Gifre, Ll., Izquierdo-Zaragoza, J.L., Ruiz, M. Velasco, L. 'Autonomic Disaggregated Multilayer Networking', IEEE/OSA Journal of Optical Communications and Networking, 2018, 10 (5), pp. 482-492, DOI:10.1364/JOCN.10.000482

[4] Rafique, D., Velasco, L. 'Machine Learning for Network Automation: Overview, Architecture, and Applications', IEEE/OSA Journal of Optical Communications and Networking, 2018, 10, (10), pp. D126-D143, DOI:10.1364/JOCN.10.00D126

[5] Leira, R., Aracil, J., López de Vergara, J.E., et al., 'High-speed optical networks latency measurements in the microsecond timescale with software-based traffic injection', Optical Switching and Networking, 2018, 29, pp. 39-45, DOI:10.1016/j.osn.2018.03.004

[6] Ruiz, M., Ramos, J., Sutter, G., et al. 'Accurate and affordable packet-train testing systems for multi-Gb/s networks', IEEE Communicatons Magazine, 2016, 54, (3), pp. 80-87, DOI:10.1109/MCOM.2016.7432152

[7] ESnet, 'iperf3: A TCP, UDP, and SCTP network bandwidth measurement tool', https://github.com/esnet/iperf, 2019.