Enhancing cross layer monitoring on open optical transport networks

Nathan Ellsworth[1], Tianliang Zhang[1], Sebastian Troia[2], Guido Maier[2], and Andrea Fumagalli[1]

[1] *Open Networking Advanced Research (OpNeAR) Lab, UT Dallas, TX, USA* [2] *Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milan, Italy*

Email: nathan.ellsworth@utdallas.edu

Abstract: Continuous monitoring of key network elements is instrumental in intelligent control and predictive analysis. This demonstration illustrates implementation challenges that are encountered in cross-layer monitoring of optical transport networks in an open-source network operations platform.

OCIS codes: (060.4250) Networks, (060.2310) Fiber optics.

1. Overview

Today, we are witnessing the emergence of new services and applications that require low latency, high reliability, increased connection density and improved energy efficiency in an unprecedented way. Examples such as augmented reality and the Metaverse are paving the way to applications increasingly in need of high data rate and ultra-low latency networks. Cloud computing provides on-demand, real-time access to a variety of applications, each with different network requirements. The optical transport network, which enables these services, must be more agile and capable of achieving dynamic and efficient allocation of resources, such as the zero-touch provisioning architecture for distributed data centers envisioned in [1]. As a result, more precise and intelligent network monitoring mechanisms are increasingly required to help control and predict the application quality of service (QoS) at runtime.

In the context of open optical transport networks, and more specifically in the OpenROADM implementation eco-system, providing consistent and real-time information for effective and efficient QoS monitoring of cloud applications remains challenging. In this demonstration, we propose an enhanced cross-layer monitoring system developed and validated inside the Network Operation Platform (NOP) [2]. The enhanced NOP (eNOP) is designed to collect cross-layer network data in an OpenROADM compliant network comprising network elements from multiple vendors in real time. Specifically, it is able to collect the following performance metrics:

1. Optical layer: ROADM link input/output power and transponder pre-FEC corrected errors
2. Network layer: Per-port L2 switch data rates
3. Transport layer: TCP parameters of active traffic flows, such as congestion window, receive window, number of segment re-transmissions, round trip time.

The goal of this demonstration is to showcase how eNOP is able to collect the aforementioned performance metrics in real time on a real-world OpenROADM based optical network testbed.

2. New Functionalities Offered by eNOP

From a network operator perspective, visibility of fault, configuration, availability, performance, and security (FCAPS) is critical to maintaining the high service levels demanded by customers and regulators alike. Many of the Operations/Business Support Systems (OSS/BSS) tools used by commercial operators are proprietary and closed source. In order to facilitate exploration and further innovation in this space from an academic research perspective, our team has been gradually adding monitoring features to our all open-source Network Operations Platform (NOP), see Fig. 1. Here we introduce two new features.

2.1. Continuous Retrieving of Optical Network Performance Metrics

First is the capture, storage, and presentation of NETCONF-based Open ROADM Performance Metrics (PMs), including OTS optical input and output power and transponder pre-FEC corrected block counts. These PMs are

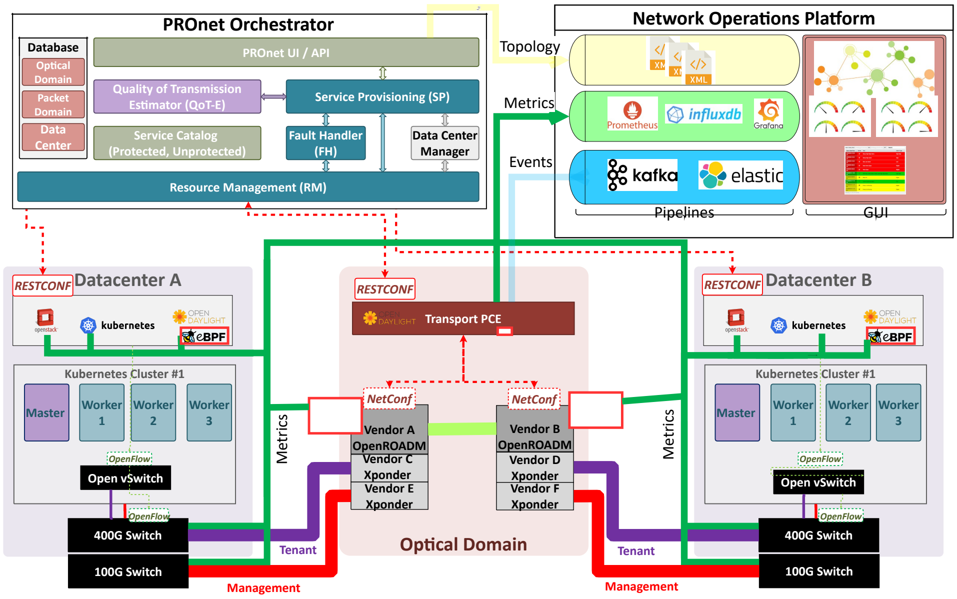

Fig. 1. eNOP Architecture Diagram

graphed live in a UI console and stored in a "data lake" for data analysis. The pre-FEC corrected block count is important as it is proportional to the bit error rate (BER) when signal rate and FEC overhead is taken into account.

In the traditional application of NETCONF-based monitoring, an SDN controller, such the open-source SDN controller for OpenROADM – Transport PCE (T-PCE), performs the translation of a northbound RESTCONF user interface to the southbound NETCONF interface with the network elements [3]. In our effort to obtain real-time PM values, we initially used this approach, which appeared to work with Honeynode-based emulated network elements. But when we began using real hardware, we noticed disturbingly frequent NETCONF session locks which caused both the PM collection and a parallel NETCONF lightpath service provisioning request to fail. In Fig. 2, which shows the pre-FEC corrected errors during a hardware-based experiment, the missing PM data is obvious.

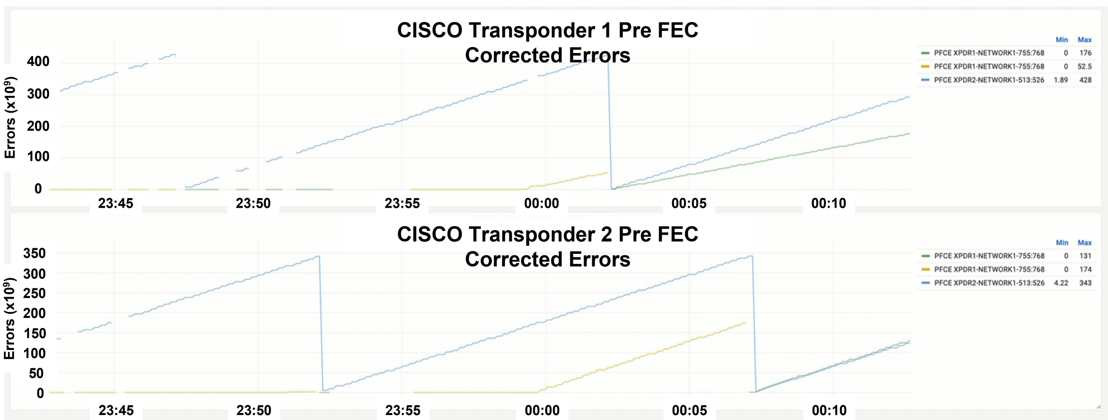

Fig. 2. Real-time graph of pre-FEC corrected errors, showing data gaps due to PM query failure

We noticed that there are certain race conditions between the SDN controller threads which perform the PM queries and the operational configuration update requests to the network devices, while one waits for the other to release a NETCONF session lock. There are several approaches to solving this problem, including modification of the SDN controller source code to implement an improved session lock management system or to avoid using the SDN controller for PM collection.

We will demonstrate the later approach, using a newly-developed software module for the Prometheus timeseries database which is at the core of our NOP. This module will use the industry-standard Prometheus exporter protocol tied together with the standard NETCONF query protocol and tuned to the OpenROADM MSA PM specifications. This module will be shared with the larger community using the Github.com code repository for Prometheus [4].

2.2. *eBPF-based TCP Parameters Collection*

We leverage the extended Berkeley Packet Filter (eBPF) [5], a novel technology that can run programs into the Linux kernel without changing kernel source code or loading kernel modules. Thanks to this technology, we are able to code specific programs from the user space to monitor different parameters of transport layer protocols, such as the number of TCP re-transmissions of traffic flows. We developed an eBPF-based Prometheus exporter capable of collecting the following TCP parameters of active traffic flows, such as congestion window, receive window, round trip time and number of re-transmitted segments. In this case too, the exporter will be shared with the larger community using the Github.com code repository for Prometheus [4].

The data lake built by Prometheus contains monitoring parameters that are collected from three layers (optical, data packet, and end-to-end transport) and are readily available for analysis of correlation.

3. OFC Relevance

This submission directly responds to the OFC Demo Zone call for demonstrations of programmable networks, by focusing on eNOP-based control and monitoring of open optical hardware. The main objective is to build an open tool capable of providing a complete and transparent view of the transport network starting from the physical (optical) layer all the way up to the transport/application layer. This holistic approach enables us to troubleshoot network problems in a targeted and conscious way, by providing intelligent optimization algorithms with the possibility of discovering relationships between cross-layer performance parameters that are not immediately visible to the human eye. In addition, the demonstration focuses on the disaggregated OpenROADM latest technologies and makes use of open-source software packages for ease of independent replication by other researchers.

4. Demo Zone Content and Implementation

The objective of this demo is to showcase the new functionalities that we have added to network operations platform. The configuration of the demo is much like that shown in Fig. 1 with hardware compute nodes communicating across a diverse multi-vendor optical network. Our eNOP display is web-based and can be shown on a large monitor attached to a laptop which will be connected to our demonstration testbed on the conference showcase floor or via VPN to the UTD lab if conditions do not allow shipping of the equipment. The demonstration itself consists of some introductory slides followed by live presentation of features discussed above, along with any late-breaking functional additions.

This year's OpenROADM demonstration brings an OpenROADM-compliant all-optical 400G regenerator for the first time. By using regenerators, long-distance and high-speed signal quality are guaranteed. At the same time, the open-source SDN controller for OpenROADM – Transport PCE (T-PCE) – is for the first time offering automatic path restoration at the physical layer. The restoration procedure is triggered when any of the links the signal goes through experiences a severe power degradation (e.g., fiber cut). The demonstration illustrates the challenges that are encountered when carrying out the automatic path restoration mechanism *while at the same time* performing continuous monitoring of the key network elements, e.g., cumulative per-oFEC error corrections at the receiver.

Acknowledgments This work is supported in part by NSF grant CNS-1956357 titled "Optics without Borders."

References

1. IOWN Global Forum, "IOWN Global Forum," https://iowngf.org/white-papers/ (2022). Last retrieved 11/15/2022.
2. N. Ellsworth *et al.*, "A non-proprietary network operations platform for openroadm environment," in *Optical Fiber Communication Conference,* (Optical Society of America, 2021), pp. M1B–3.
3. A. Sgambelluri, A. Giorgetti, F. Paolucci, P. Castoldi, and F. Cugini, "Open source implementation of openconfig telemetry-enabled netconf agent," (2019), pp. 1–4.
4. Prometheus Community, "Prometheus Community," https://github.com/prometheus-community (2022). Last retrieved 11/15/2022.
5. ebpf Community, https://ebpf.io/ (2022). Last retrieved 11/15/2022.