Traffic Monitoring System for 100-Gbps Virtualized Optical Networks

Yusuke Sekihara, Namiko Ikeda, Hiroyuki Uzawa, Shoko Ohteru, Saki Hatta, Shuhei Yoshida, Kimikazu Sano

NTT Device Innovation Center
NTT Corporation, 3-1 Morinosato-wakamiya, Atsugi, Kanagawa, 243-0198, Japan,
yuusuke.sekihara.kv@hco.ntt.co.jp

Abstract: We propose a system to get flow information of virtualized logical networks in support of 100-Gbps optical networks. By using packet sampling, we can create monitoring rules even when the monitoring target is unknown.

1. Introduction

With the development of cloud services, the number of application services that use networks is increasing. Data centers that provide cloud services are adopting network virtualization technology to shorten the lead time for service delivery by reducing the time required for network construction and modification in response to such diversification of services. The network virtualization technology enables us to share network resources among multiple network services and thus brings benefits such as reduced CAPEX and OPEX. Moreover, each service provider can become free to design and build its own individual network. On the other hand, its technology causes making it difficult to obtain routing information in each network service. This means that service providers have operational problem that make it difficult to distinguish whether a service problem is caused by the logical network or the physical network. To enable checking which logical network is flowing through which physical network, there is a strong need for a system that monitors the traffic flowing through the nodes on the physical network. For this reason, some traffic monitoring systems have been developed in the past [1]. However, the conventional system has been implemented with software and is unsuitable for the traffic monitoring in the optical network for datacenters. Because 100-Gbps class traffic flows over the optical network and the software-based system cannot handle high-load packet processing such as analyzing encapsulated packets and classifying network traffic using many header fields at high speed.

In this paper, we propose a traffic monitoring system for virtualized traffic on 100-Gbps optical networks. In our proposed system, we use rule-based matching for traffic analysis that is used for high speed processing. Rule-based matching is an analysis method that detects flows by comparing input traffic with pre-defined rules that specify monitoring targets. In this type of analysis method, high speed processing is achieved but it is necessary to know in advance the flows to monitor, even if the flow pass through the physical node cannot be identified, such as virtualized traffic. Therefore, this paper simultaneously proposes rule generation by packet sampling. This make possible our system to monitor virtualized traffic where the flow could not be identified. In addition, we proposed system consists of an FPGA-based smart NIC and a general-purpose server, whose cost is lower than ASICs or other dedicated products.

2. Proposed system

The proposed system aims to monitor flow over a single node of a physical network in a virtualized network environment. The system consists of a general-purpose server and an FPGA-based smart NIC, and thus provides both cost efficiency and high speed processing. The FPGA analyzes received packets, generates statistical information necessary for traffic monitoring. Statistics generated are, for example, packet count, byte count, and so on. Statistics generated by the FPGA are sent to the traffic-monitoring software by middleware. Middleware provides communication function between the FPGA and the server OS. The traffic-monitoring software consists of open source software, which collects and analyzes received statistical information and build a database and also provides GUI to visualize database queries. Thus, a high-speed and flexible system can be constructed by using FPGAs to perform heavy-load packet analysis processing and software to perform statistical information processing that requires flexibility. Configuration of this system and target network is shown in Fig.1. As shown in Fig.1, this system is connected to an optical TAP device installed on physical optical network, and monitors logical network flowing on physical path by analyzing branched optical signals. One of features of this system is that it can monitor main line without affecting it, since it targets optical signals branched by TAP. Input optical signal packets are subjected to a rule matching process using a hash-based search algorithm in the FPGA.

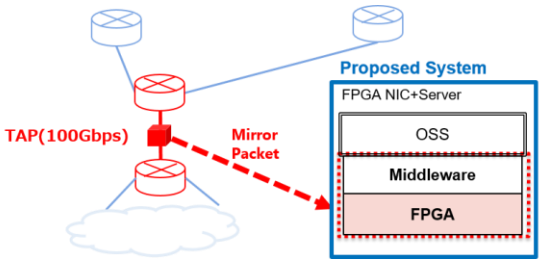

Fig. 1. System configuration and target network

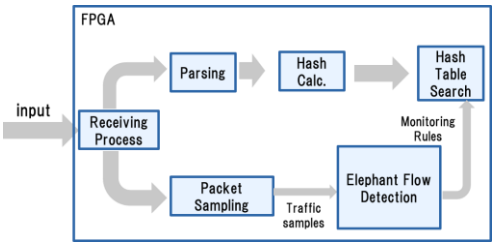

Fig. 2. Functional block structure inside the FPGA

3. Proposed rule generation method

The hash-based search algorithm builds a hash table based on pre-input rules, and performs a high-speed comparative matching between input values and the hash table. However, since the search target to detect must be known in advance, it is difficult to use this algorithm for monitoring networks such as virtualized ones, where it is not possible to specify on which path the traffic is flowing as it changes from time to time. Therefore, we propose rule generation by packet sampling. Packet sampling is a technology proposed for sFlow monitoring [2], which analyzes packets passing through a device at a rate of one to N and uses the results to infer the traffic before sampling. The traffic inferred from the sampling gives a rough idea of the volume of the flow in the interval time. However, since the exact behavior of the traffic before sampling is not known, sampling is insufficient to cover the operation of networks that require a high level of availability and reliability, such as optical networks. Here, we propose to identify elephant flows that dominate the network behavior by analyzing sampling packets, and to use the identified elephant flows as rules. The rule-matching-based monitoring allows us to monitor all packets of an elephant flow and to understand its exact behavior.

The system configuration inside the FPGA, including the above proposal, is shown in Fig.2. The system receives optical signals through a QSFP28 IF. The received signals are interpreted as an Ethernet packet by the FPGA. Input packets are sent to both parsing and packet sampling blocks inside the FPGA. Packets sent to parsing block are parsed by header field, and hash values are calculated for each field. The hash value is added to each valid field set for each monitoring rule, and the result of the addition is used as the unique Key value for each monitoring rule. The rule matching process is performed by searching the hash table for the unique Key value for each monitoring rule. In this way, by hashing and adding each field value, the search for each monitoring rule can be accomplished in a single hash table search, making the rule matching process faster. On the other hand, packets sent to packet sampling block are discarded with a certain probability. Packets that are not discarded are sent to elephant flow detection block. The elephant flow detection estimates the input traffic based on the received sampled packets. Then, among the estimated input traffic, those that are determined to be elephant flows are input to the hash table as monitoring rules. We use the Top-k selection algorithm for the elephant flow detection. Top-k selection is an algorithm for extracting the top k values from a population. In this work, we used the bit count of each flow as the value, and the traffic inferred from packet sampling as the population.

Table. 1. Hardware configuration

FPGA	Intel Stratix 10 SX
CPU	Intel Xeon Silver 4110
Memory	64Gbyte
PCIe	Gen3 × 8,x16 × 1
IF	QSFP28 × 2

Table. 2. Synthesis result

ALMs	Block Memory	Utilization
287469.3	33565956	38.5%

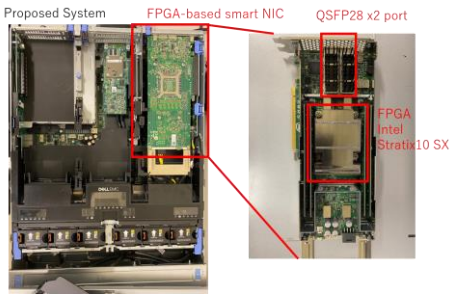

Fig. 3. Photograph of proposed system

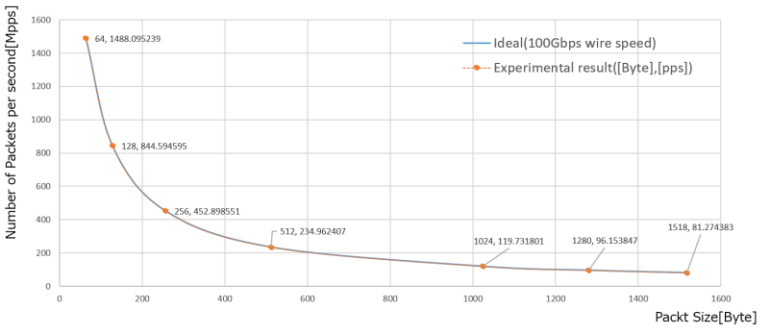

Fig.4. Experimental results

4. Experimental results

The hardware configuration of our proposed system is summarized in Table1. Table2 shows the results of synthesis on the FPGA. Fig.3 shows a photograph of proposed system. We evaluated the traffic monitoring throughput of the proposed configuration using VLAN and MPLS simulated traffic as test traffic. Fig.4 shows the experimental results. Fig.4 shows that the proposed configuration can monitor an entire 100-Gbps Ethernet without packet loss, and can monitor traffic at the wire-speed of an optical network. In addition, the results were identical for MPLS and VLAN, indicating that wire-speed monitoring is possible regardless of the protocol or header configuration. The elephant flow was extracted by sampling the CAIDA [3] Internet backbone data at a rate of 1/100 and sequentially extracting the top-100 flows in terms of the count of bits in the interval. As a result, it was found that about 60% of traffic could be monitored by using top-100 flows as the monitoring rule. This result is roughly the same as the elephant flow ratio reported in various traffic [4]. Therefore, it is considered that the necessary traffic can be sufficiently monitored. In addition, the latency for sampling is about 1ms in this implementation, which is a value that can sufficiently follow traffic fluctuations in general networks.

5. Summary

In this paper, for a 100-Gbps optical network, we proposed a system that monitors traffic from optical TAPs on the physical network and obtains flow information of virtualized logical network. The system consists of an FPGA-based smart NIC and a general-purpose server, which achieves economic efficiency, and we proposed to utilize packet sampling for monitoring rule generation to identify flows on the logical network. The experimental results show that the proposed system can monitor all the virtualized traffic at 100-Gbps wire speed without any packet loss. We also confirmed that detailed monitoring of elephant flows, which are 60% of traffic and dominant in the network behavior, can be achieved by using packet sampling to generate monitoring rules for unknown traffic.

6. References

[1] One Technologies Company, https://www.toyo.co.jp/onetech/incubation_english/detail/neteyez/

[2] IETF, https://datatracker.ietf.org/doc/html/rfc3176

[3] CAIDA, https://www.caida.org/

[4] Péter Megyesi, Sándor Molnár. Analysis of Elephant Users in Broadband Network Traffic. 19th Open European Summer School (EUNICE), Aug 2013, Chemnitz, Germany. pp.37-45