

Assessing the performance of XDP and AF-XDP based NFs in edge data center scenarios

Original

Assessing the performance of XDP and AF-XDP based NFs in edge data center scenarios / Parola, F.; Procopio, R.; Risso, F.. - ELETTRONICO. - (2021), pp. 481-482. (Intervento presentato al convegno 17th International Conference on emerging Networking EXperiments and Technologies, CoNEXT 2021 tenutosi a Virtual event nel December 7 - 10, 2021) [10.1145/3485983.3493352].

Availability:

This version is available at: 11583/2970690 since: 2022-08-20T08:53:30Z

Publisher:

Association for Computing Machinery

Published

DOI:10.1145/3485983.3493352

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

ACM postprint/Author's Accepted Manuscript

(Article begins on next page)

Poster: Assessing the Performance of XDP and AF_XDP Based NFs in Edge Data Center Scenarios

Federico Parola
Politecnico di Torino
federico.parola@polito.it

Roberto Procopio
TIM S.p.A.
roberto.procopio@telecomitalia.it

Fulvio Rizzo
Politecnico di Torino
fulvio.rizzo@polito.it

ABSTRACT

While servers in traditional data centers can be specialized to run either CPU-intensive or network-intensive workloads, edge data centers need to consolidate both on the same machine(s) due to the reduced number of servers. This paper presents some preliminary experiments to determine how to improve the overall throughput of the above servers, being XDP and AF_XDP the two main technologies into play.

CCS CONCEPTS

• **Networks** → Middle boxes / network appliances; Network servers; Programmable networks; • **Software and its engineering** → Communications management.

KEYWORDS

Network Function Virtualization, XDP, AF_XDP

1 INTRODUCTION

Telco operators are increasingly building micro data centers at the edge of their network, in which the few available servers are expected to execute both telco-oriented workloads (e.g., 5G user-plane functions) and general-purpose applications (e.g., content caches, object recognition software, 5G control plane, etc.). While in traditional data centers each server is dedicated to either one of the workloads, the reduced number of servers (usually on the order of 5-10 units) suggest that they would host a variable mix of network-oriented and general-purpose workloads.

Technologies such as eXpress Data Path (XDP)[2] and AF_XDP, could provide an ideal packet processing infrastructure for the above “integrated” scenario, while other kernel-bypass technologies (e.g. DPDK) are more suitable for servers dedicated to dataplane-oriented tasks due to the necessity of dedicated CPU cores and huge memory pages, and the difficulties to support applications that require the standard TCP/IP stack. On the other side, XDP allows to process packets in the NIC driver, retaining the possibility to yield a packet to the Linux network stack, while AF_XDP sockets can be used to bypass limitations of eBPF programs.

This paper presents our preliminary analysis of the interactions between the above-mentioned packet processing mechanisms, the underlying hardware primitives (e.g., DDIO, caches), and the running workloads, deriving insights on how to optimize the throughput in the above target scenario.

2 EXPERIMENTS

This section presents our preliminary findings when running mixed workloads on the same server.

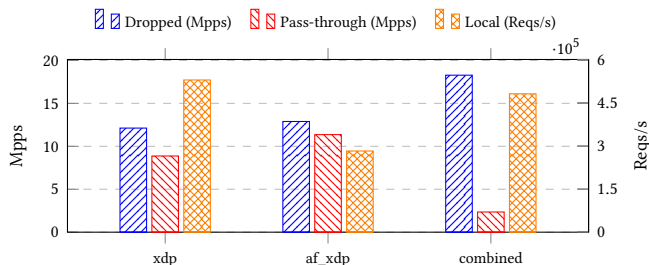


Figure 1: Throughput comparison for different technologies and different traffic destinations.

The choice between XDP and AF_XDP depends on the destination of the involved traffic. We identified three different classes of traffic that can be handled by our server, each one characterized by a different type of IO processing. **Dropped traffic** is the traffic that is discarded by the NF, typically a Firewall, a DDoS Mitigator or a traffic shaper. **Pass-through traffic** is the traffic that is forwarded by the machine after being processed by one or more network functions. This can be the case of a load balancer redirecting packets towards backends running on a different server. **Local traffic** is the traffic that has to be processed by an application running on same server as the NF, e.g., traffic that is processed by a firewall and terminated on a local pod. Even though this three scenarios are usually combined in a common deployment, we decided to study them in isolation in order to profile the behavior of the technologies under test and then define the best way to combine them in a real use case. Our experiments leverage the `xdp_rxq_info` and `xdpsock` samples from the Linux kernel tree¹ to perform some simple packet IO in XDP and AF_XDP. For AF_XDP we ran the user space packet processing thread on the same core in charge of handling interrupts of the NIC and enabled the `SO_PREFER_BUSY_POLLING` flag, that resulted in better performance. We also evaluated a combined mode, in which AF_XDP sockets are enabled but the verdict on the final destination of the packet is taken at the XDP level. For local traffic we executed an instance of the `memcached` server and measured the maximum number of requests per second that it was able to handle.

Results in fig. 1 show that the technology achieving the best throughput depends on the destination of the traffic: while the combined mode achieves the best results for dropped traffic, pure AF_XDP yields better performance when the traffic is redirected outside the machine. The reasons for the above results can be found in the following considerations: 1) all packets processed at the AF_XDP level pass through an XDP program, hence dropping the packet at this early stage can therefore save CPU cycles; 2) packets follow a different path in the kernel if redirected at the AF_XDP or

¹<https://github.com/torvalds/linux/tree/v5.14/samples/bpf>

Assessing the performance of XDP and AF-XDP based NFs in edge data center scenarios

Original

Assessing the performance of XDP and AF-XDP based NFs in edge data center scenarios / Parola, F.; Procopio, R.; Risso, F.. - ELETTRONICO. - (2021), pp. 481-482. (Intervento presentato al convegno 17th International Conference on emerging Networking EXperiments and Technologies, CoNEXT 2021 tenutosi a Virtual event nel December 7 - 10, 2021) [10.1145/3485983.3493352].

Availability:

This version is available at: 11583/2970690 since: 2022-08-20T08:53:30Z

Publisher:

Association for Computing Machinery

Published

DOI:10.1145/3485983.3493352

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

ACM postprint/Author's Accepted Manuscript

(Article begins on next page)