Introducing Data Processing Units (DPU) at the Edge [Invited]

Luca Barsellotti*, Faris Alhamed†, Juan Jose Vegas Olmos‡,
Francesco Paolucci*, Piero Castoldi†, Filippo Cugini*
*CNIT, Pisa, Italy.
†Scuola Superiore Sant'Anna, Pisa, Italy.
‡NVIDIA, Israel.
*Corresponding author: filippo.cugini@cnit.it

***Abstract*—The recent availability of smart network interface cards (smart NICs) and Data Processing units (DPUs) providing hardware-accelerated networking and computing functionalities is opening the way towards new applications and use cases beyond the traditional data center scenarios.**

In this paper, three different use cases for edge scenarios that leverage on the innovative programmability enabled by DPUs are presented and discussed. The first use case focuses on a pervasive monitoring infrastructure to support accurate and decentralized network awareness for low-latency 5G services. The second one focuses on the implementation of power-efficient edge-to-cloud continuum. The third use case refers to effective network security functions at the DPU.

***Index Terms*—P4, smart NIC, in-network functions, DPU, telemetry, security, programmability, programmable data plane.**

I. Introduction

Software Defined Networking (SDN) and data plane programmability, originally designed for data center applications, are becoming ubiquitous, also targeting edge computing scenarios [1]. So far, data plane programmability (e.g., using the P4 technology [2]) has been mainly implemented in bare metal and software switches [3]. However, the recent evolution of network interface cards (NIC) towards smart NICs and Data Processing Units (DPU, e.g. [4]) is driving the introduction of novel programmable network functions accelerated in hardware also within edge computing nodes. This opens the way to innovative solutions aiming at efficiently support innovative low-latency 5G services as well as providing accurate network awareness and security capabilities to computer and orchestration systems at the edge [5].

In this paper, we first provide an overview of the innovative capabilities enabled by DPUs. Then, we present three different use cases for edge scenarios that leverage on the innovative programmability enabled by DPUs.

The first application consists in an innovative pervasive monitoring infrastructure enabling accurate network performance monitoring across the entire end-to-end network. The monitoring infrastructure relies on (i) telemetry extended up to the user equipment, (ii) telemetry enhanced as in-network communication channel, (iii) decentralized telemetry among network and edge devices and modules, and (iv) novel hardware-accelerated telemetry collector.

The second application consists in a power-efficient scenario of smart edge nodes able to aggregate computing, networking and optical transmission technology in a single energy-efficient element at the edge.

The third application consists of fast HW-accelerated cyber-security operations relying on machine learning processing directly performed within the DPU.

This work has been supported by the BRAINE Project, funded by ECSEL JU under grant agreement No. 876967.

II. Previous works on smart NIC technology

The introduction of smart NIC has been received with a large interest in the networking community, thanks to the flexibility and the novel degree of programmability that such devices may bring to the network-computing node boundary. The main research topics are related to application-aware forwarding, security, converged nodes architectures, telemetry and 5G functions.

The work in [6] analyzes the use of smart NIC for offloading software packet processing in the kernel user space for distributed denial-of-service (DDoS) attack mitigation purposes. The work in [7] proposes the adoption of smart NIC for in-band Network telemetry to offload the chained latency values processing at the INT sink node interface. The work in [8] adopts a P4-enabled smart NIC attached to an optical Bandwidth Variable Transponder to perform L2-L3 operation (i.e., routing, forwarding) directly at the optical node. The work in [9] introduces the smart NIC with local FPGA processing for disaggregated hardware acceleration of Distributed Units and Centralized Units in the 5G Radio Access Network segment. The work in [10] introduces the serverless dynamic latency-critical application migration at different edge nodes coordinated by a joint packet-optical monitoring infrastructure combining packet-level INT and quality of transmission telemetry of optical channels. Full migration and recovery of specific processing-intensive VNF belonging to a chains of Function-as-a-Service modules is enforced in less than 10ms.

To the best of our knowledge, only a single work analyzes the performance of the recently released DPU platforms, highlighting the benefits mainly in the field of network encryption/decryption offloading, data compression and decompression, and inter-process communication [11].

III. Programmable smart NIC and Data Processing Unit (DPU)

The different smart NIC and DPU offloading flavours are shown in the examples of Fig. 1. Typical state-of-the-art offloading of network functions are employed at the network switch levels (i.e., at the programmable P4 switch/routers nodes) and partial offloading is performed at the edge node. While traditional NICs provide the low transmission protocol stack acceleration (e.g., Ethernet MAC) and smart NICs provide some programmability at L3-L4 (e.g., exploiting P4). Furthermore, advanced in-network functions are run inside the edge node resorting to the computational capabilities of the node, e.g., running a chain of VNF or containers, optionally resorting to GPU resources to enforce AI-based problem solutions. In this partial offloading, application packets are L1-L2 processed at the NIC (including also L3 and L4 in case of smart NICs) (1), passed to containers processing involving CPU cores (2), optionally run inference and classification resorting to GPU (3) and outcome post-processed again by containers (4), and finally sent out to the next hop performing L1-L2 encoding at the NIC (5). With the advent of DPU, full offloading flavour may be envisioned. In this case, the packet is received by the DPU providing all the programmable ASIC-based acceleration features for packet dissection and L1-L7 protocol dissection and processing. Moreover, in-network functions are run inside the DPU itself resorting to local processing capabilities, such as embedded CPU cores, optional GPU availability and additional acceleration stages such as Deep Packet Inspection filters (2). Finally, the processed packet is re-transmitted exploiting the programmable pipeline ASIC acceleration stages (3). In the case of full offload, the edge node computational resources are not employed for in-network packet processing, thus allowing improved processing availability for tenants and application services.

IV. DPU-based pervasive monitoring infrastructure

The use of programmable smart NIC has the potential of supporting novel telemetry solutions specifically designed for the edge continuum ecosystem. A pervasive monitoring infrastructure can be envisaged to provide distributed knowledge and innovative decentralized decision-making solutions.

The infrastructure will leverage on the following innovations, shown in Fig. 2: (1) telemetry extended up to the user equipment, (2) telemetry enhanced as in-network communication channel, (3) decentralized telemetry among network and edge devices and modules, (4) novel hardware-accelerated telemetry collectors.

A. Telemetry at the user equipment

In-band Network Telemetry (INT, [12], [13]) whose processing is performed in hardware by the DPU has the potential to be extensively exploited not only to retrieve accurate statistics/metadata from network nodes, but also enforced at the user equipment or Internet-of-Things (IoT) terminal. This

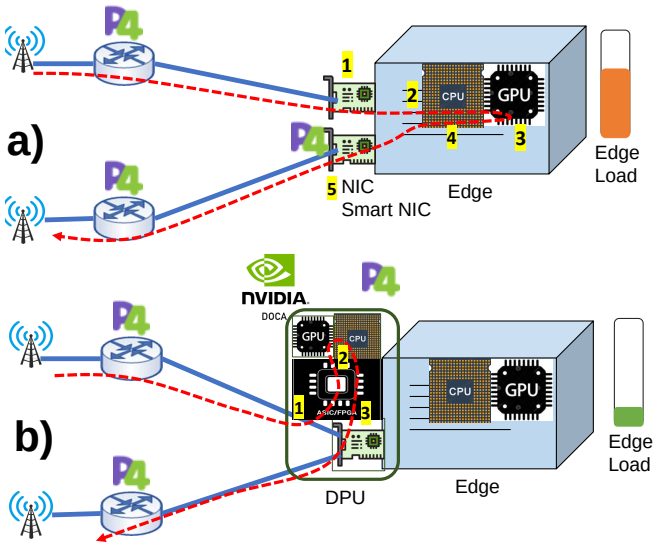

Fig. 1. Partial (a) and full (b) in-network function offload at the edge.

will enable the collection at the edge of accurate end-to-edge information, such as the overall latency experienced crossing both the wireless and the wired links. Furthermore, additional information, traditionally not easily available at the edge, will be retrieved. This includes (i) geo-localization data of terminals and devices with low age of information (AoI), (ii) user equipment performance such as system parameters, queue occupation, CPU load, etc. with sub-second granularity, (iii) quality parameters enabling AI-based investigation of the overall level of customer satisfaction. These additional parameters will enable not only detailed mobility monitoring and improved cloud-edge or edge-edge steering policies, but also forecasting and completely new MAS-based distributed knowledge and decision-making approach at the edge. [14].

B. Telemetry as fast in-network communication channel

Hardware-accelerated telemetry has the potential to carry not only monitoring parameters, but will also to support fast in-network communication channels among network elements or from the user equipment. For example, as soon as performance degradation is detected (or forecasted) at the user equipment, rather than waiting for monitored data to be collected and processed by a central collector or at the edge, the user equipment can directly request service adaptation though specific fields included in an extended version of in-band telemetry. Indeed, by leveraging on P4-based packet manipulation, specifically added header flags/fields can be directly filled in by the UE/terminal and consumed by intermediate switching elements or edge nodes [15]. This has the potential to significantly reduce the reaction time by the infrastructural elements as well as remarkably improved the accuracy, since sporadic cases leading to false warning/alarm conditions might be avoided/limited through such explicit in-network communication channel.

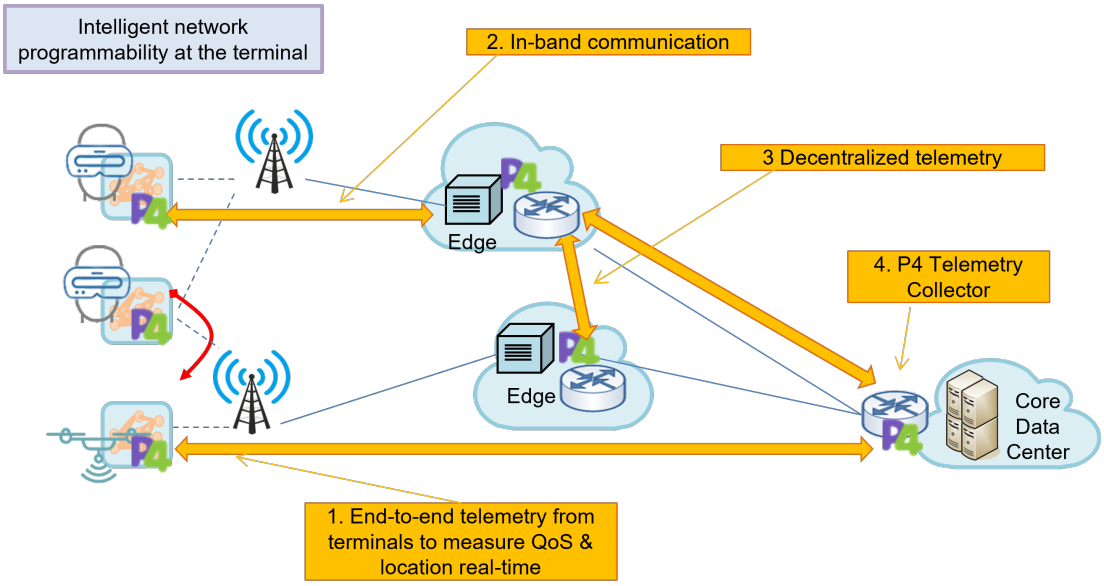

Fig. 2. Next-generation in-network functions exploiting programmable data plane networking at the terminals.

C. Decentralized telemetry

Having the network react in real time to network problems and changes is crucial to keeping the services provided by the network uninterrupted, and to maintain an acceptable low latency that falls within the terms of the Service-Level Agreement offered by the provider. For this reason many systems have been previously proposed for the monitoring of network devices and for the collection and analysis of telemetry data to diagnose abnormalities. One example is Pingmesh [16] that works at the endpoints and aims to measure the latency across all endpoints in a Data Center Network. However, Pingmesh requires an Agent on every element which, according to [16], has to be carefully implemented so that the CPU, memory, and bandwidth overhead is small and affordable. Another example of a system that collects and manages decentralized telemetry data is NetView [17]. In this case, a *Telemetry Antenna* injects a *probe* into the network. The probe is routed using source routing to collect telemetry data along specific network routes, and delivered to the *Telemetry Analyzer*. The above solutions are relevant for metro-edge scenarios supporting low latency applications, particularly if accelerated in hardware. Indeed, data plane programmability and SmartNICs/DPUs open the doors to hybrid solutions in which telemetry data is still exported to telemetry collectors, while the DPU allow to monitor the edge computing resources in a decentralized way. For example, network performance indicators such as latency and congestion can be measured and processed in real-time to allow traffic to be re-routed at load-balancers. This operation can be accelerated by the DPU itself without affecting the CPU usage on the end server.

D. P4 telemetry collector

Traditional centralized collection of telemetry data offers the benefits of global visibility and effective correlations, potentially leading to optimal decisions. However, the implementation of such collectors is extremely critical, due to the amount of received data to be processed at wire speed (see Fig. 3(a)). The preliminary work in [18] proposes a two-stage telemetry collector where the first stage of data collection, pre-processing, and aggregation is performed by a programmable P4 switch/DPU which pre-processes the data at wire-speed, also leveraging on its multiple interfaces and on embedded stateful capabilities for preliminary correlations, as shown in Fig. 3(b). This way, the amount of data received by a telemetry server for subsequent data storage and elaboration is significantly reduced. For example, for selected 5G(+) services, min/avg/max latency statistics can be effectively extracted from N post-card or in-band telemetry messages, reducing by a factor of N the bandwidth at the telemetry server without requiring significant P4 memory resources. Furthermore, significant benefits are expected at the server CPU. For example, a preliminary test showed that a telemetry stream of 30k pps overwhelms a CPU to 100% load, while reducing it of a factor of N=15 reduces the CPU to a value of 50%. This solution will also provide an automated solution to assess the level of aggregation per critical service accounting for service requirements, rate, introduced latency in the delivery of data, and usage of P4 resources.

V. Converged Packet, Optical, and Edge Computing Infrastructure

In the context of edge/metro networking, the introduction of packet-optical white box, i.e. packet forwarding nodes (e.g., IP routers) equipped with coherent pluggable modules, is driving the design and implementation of interoperable low-cost converged packet-optical transport solutions effectively capable of removing boundaries between different network domains. For example, edge to cloud interconnection can

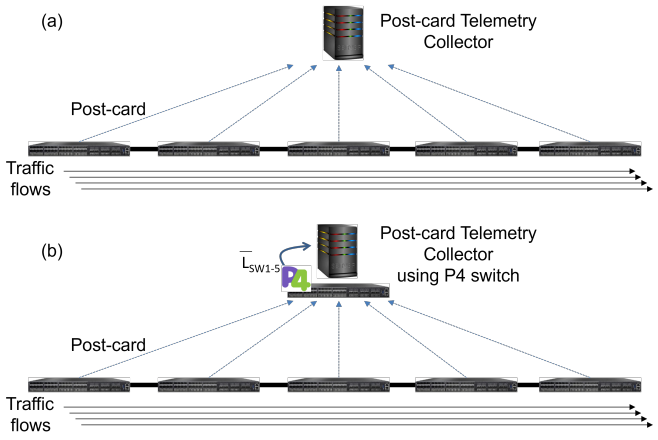

Fig. 3. (a) Traditional centralized telemetry collector on a server. (b) P4-based HW-accelerated telemetry collector.

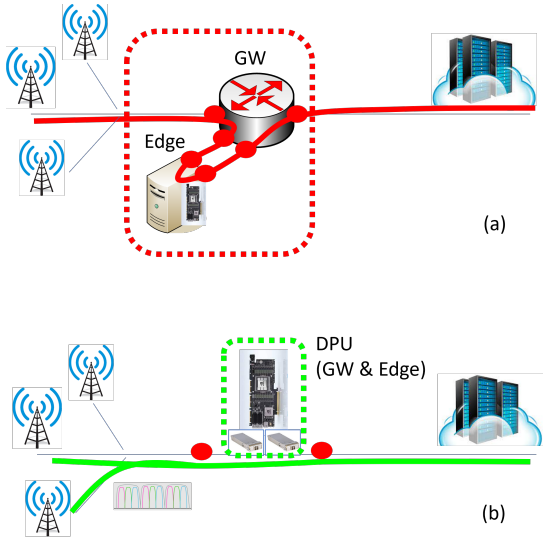

Fig. 4. (a) Traditional approach for routing and computing at the edge, leveraging on different network and computing equipment; (b) converged networking and computing scenario at the edge using DPU. Highlighted in red the Optical-electro-optical conversions.

be implemented using a single router/gateway at the edge performing packet forwarding as well as long-reach optical transmission without requiring dedicated optical equipment (e.g., transponders), as shown in Fig. 4(a).

DPUs have the potential to further improve the edge-to-cloud continuum removing the barriers between computing and networking resources, as shown in Fig. 4(b). The delivery of 5G and beyond services is indeed driving Telco Central Offices (CO) to host not only networking equipment such as routers, but also edge computing resources. However, such separation of networking and computing resources is expensive (both in terms of CAPEX and OPEX), power hungry, and not latency efficient. For example, multiple opto/electro/optical conversions have to be experienced in the computing continuum between (B)5G services and edge and cloud resources, as highlighted by the red dots in Fig. 4(a).

DPUs provide a unified network-computing element providing both selected IP functionalities (e.g., Layer-2/3 and specific MPLS features with native time synchronization protocols such as IEEE-1588 v2 and Synchronous Ethernet) as well as edge computing resources. Indeed, DPUs, as evolution of smart network interface cards designed for intra-data centre networking, provide advanced networking capabilities (e.g., up to 4 interfaces at up to 400Gb/s, advanced timing and synchronization, HW encryption and embedded security features, P4 programmability, etc.) as well as computing and acceleration capabilities for selected ultra-low latency services at the edge (e.g., up to 16 ARM CPUs, programmable acceleration for AI processing). The benefit of relying on a single equipment for edge computing, networking, and long-reach optical transmission is presented in Fig. 4(b). To this goal, DPUs have to support optical coherent technologies for cloud-edge interconnection. In particular, both point-to-point (e.g., ZR/ZR+) and point-to-multipoint (P2MP) pluggables (e.g., openXR) would be needed. The latter is of particular interest to provide high-capacity connections towards the access and the cloud while using a few physical interfaces, thus overcoming the constraint on the limited number of physical interfaces available in DPUs.

Pervasive telemetry may benefit from the collapse of both packet and optical layer in a single converged DPU, enabling augmented correlation between optical monitoring and data analytics [19] and novel packet-based telemetry acceleration.

VI. DDoS Detection and Mitigation Use Case

Cyber security is another relevant topic that is efficiently addressed using in-network data plane programmability at the DPU. So far, security feature extraction have been proposed to be accelerated through P4-based switch programmability [20]. Focusing on smart NICs, in [21] a DDoS traffic classification and filtering schema that identifies malicious packet signatures based on Machine Learning algorithms and that generates filtering rules is proposed. This schema is composed by four steps: Signature Extraction (SE), Signature Classification (SC), Signature Reduction (SR) and Anomaly Mitigation (AM). SE and AM components are provided by a smart NIC in the data plane using the eXpress Data Path (XDP) framework for high-performance, whereas SC and SR components are implemented in the control plane. SC normalizes and classifies signatures through supervised Machine Learning models, whereas SR reduces the number of malicious signatures to accelerate the mitigation performance of the AM component. This is done by resolving a multi-objective Pareto problem in which the objective is to minimize:

- The number of malicious signatures (filtering rules)
- The percentage of benign traffic drops

A fast evolutionary approach based on Non-dominated Sorting Genetic Algorithm II is adopted to iteratively try in each step to further reduce the objectives, using as stopping condition a time limit.

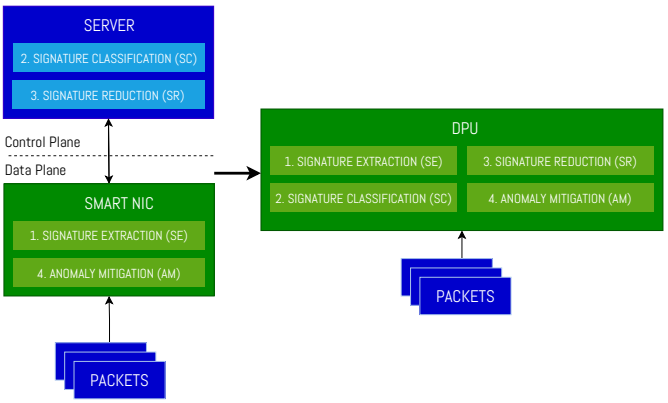

Fig. 5. Converging the Signature Extraction, Signature Classification, Signature Reduction and Anomaly Mitigation steps in the DPU.

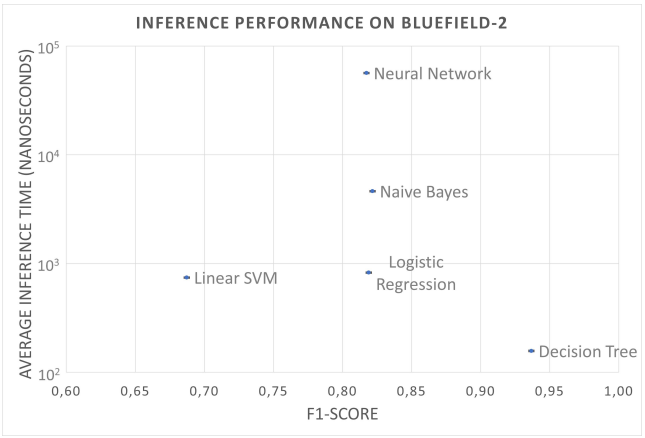

Fig. 6. Time and F1-Score performance of Machine Learning models applied on UNSW-NB15 Network Intrusion Detection dataset using a NVIDIA BlueField-2.

In our proposal, all the four steps are implemented inside a DPU, leading to several advantages:

- Reduction of the overhead in terms of time and bandwidth provided by the exchange of information between the smart NIC and the control plane.
- The control plane is completely unaware about the DDoS detection and mitigation and does not consume its resources for this task.
- The DDoS detection and mitigation process is independent from the connection between the DPU and the control plane, and can be implemented in every point of the net.

The Signature Extraction and Anomaly Mitigation steps can be handled, respectively, by the DOCA Flow and DOCA Deep Packet Inspection (through Suricata rules) features available on the NVIDIA BlueField to take advantage of the hardware accelerations. For the Signature Reduction step, the time limit used as stopping criterion can be set as a trade-off between the time constraints of the network and the performance obtained in terms of number of filtered malicious packets and percentage of benign traffic dropped with the solutions provided by the NSGA-II algorithm.

The problem that needs to be addressed consists in executing the Signature Classification on top of the DPU, which is currently not equipped with a GPU. Five models, Decision Tree, Linear SVM ($\lambda = 0.0001$, $\delta = 1.0$), Logistic Regression, Naive Bayes and the Artificial Neural Network proposed in [22] are evaluated on UNSW-NB15 Network Intrusion Detection dataset using a NVIDIA BlueField-2, equipped with 8 ARMv8 A72 Cortex cores. The ANN model is trained in TensorFlow and converted in TensorFlow Lite using a post-training quantization to exploit operations based on 8-bit integers. Then, it is parsed and executed through the optimized ArmNN library. The other models are tested using their MLPack implementation. The dataset is composed by 700000 samples, and a random 20% is assumed as fixed test set that contains 135564 benign samples and 4437 malignant samples.

From Fig. 6 it is possible to observe that the Decision Tree model is the best solution for running inference on the BlueField-2 for each packet, providing the best results in terms of F1-Score (0.9367) and average inference time (157ns). Depending on the specific time requirements, ensembling solutions based on Decision Trees can be adopted, such as a Random Forest model, to further increase the classification performance and the robustness of the model. The Neural Network, even though it is composed only by three fully connected layers and is quantized, is the slowest model, providing an average inference time equal to 56435 nanoseconds and a F1-Score equal to 0.8173. The second best model is the Logistic Regression one, with a F1-Score equal to 0.8190 and an average inference time equal to 826 nanoseconds. Naive Bayes (0.8217, 4623ns) and Linear SVM (0.6872, 744ns) models are not particularly interesting, lacking, respectively, in average inference time and F1-Score.

Conclusions

This paper presented three different use cases for edge scenarios exploiting the recently introduced innovative programmability enabled by DPUs. The first use case refers to the setup of a pervasive monitoring infrastructure to support low-latency 5G services. The second use case focuses on the potential of implementing a power-efficient edge-to-cloud continuum through a converged computing-packet-optical solution. The third use case focuses on implementation of effective network security functions which also exploit the embedded DPU computing resources. The performance of five ML models have been assessed in the context of network intrusion detection. Results showed that the Decision Tree model can provide an average inference time of a packet in only only 157ns with an extremely high F1-Score of 0.9367. The utilization of DPUs allow for native integration of artificial intelligence driven operations directly on the network fabric.

REFERENCES

[1] F. Paolucci, F. Civerchia, A. Sgambelluri, A. Giorgetti, F. Cugini, and P. Castoldi, "P4 Edge Node enabling Stateful Traffic Engineering and Cyber Security," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 11, no. 1, pp. A84–A95, Jan. 2019.

[2] P. Bosshart, D. Daly, G. Gibb, M. Izzard, N. McKeown, J. Rexford, C. Schlesinger, D. Talayco, A. Vahdat, G. Varghese *et al.*, "P4: Programming protocol-independent packet processors," *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 3, pp. 87–95, 2014.

[3] F. Paolucci, F. Cugini, P. Castoldi, and T. Osiński, "Enhancing 5G SDN/NFV edge with P4 data plane programmability," *IEEE Network*, vol. 35, no. 3, pp. 154–160, 2021.

[4] *https://www.nvidia.com/content/dam/en-zz/Solutions/Data-Center/documents/datasheet-nvidia-bluefield-2-dpu.pdf*.

[5] A. Giorgetti, J. Chamanara, M. Albado, E. Marx, S. Ahearne, D. Scano, F. Paolucci, and F. Cugini, "Kubernetes orchestration in sdn-based edge network infrastructure," in *Optical Fiber Communication Conference (OFC) 2022*, 2022.

[6] S. Miano, R. Doriguzzi-Corin, F. Risso, D. Siracusa, and R. Sommese, "Introducing smartnics in server-based data plane processing: The ddos mitigation use case," *IEEE Access*, vol. 7, pp. 107 161–107 170, 2019.

[7] Y. Feng, S. Panda, S. G. Kulkarni, K. K. Ramakrishnan, and N. Duffield, "A smartnic-accelerated monitoring platform for in-band network telemetry," in *2020 IEEE International Symposium on Local and Metropolitan Area Networks (LANMAN*, 2020, pp. 1–6.

[8] Y. Yan, A. F. Beldachi, R. Nejabati, and D. Simeonidou, "P4-enabled smart nic: Enabling sliceable and service-driven optical data centres," *Journal of Lightwave Technology*, vol. 38, no. 9, pp. 2688–2694, 2020.

[9] J. C. Borromeo, K. Kondepu, N. Andriolli, and L. Valcarenghi, "Fpga-accelerated smartnic for supporting 5g virtualized radio access network," *Computer Networks*, p. 108931, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1389128622001189

[10] I. Pelle, F. Paolucci, B. Sonkoly, and F. Cugini, "Latency-sensitive edge/cloud serverless dynamic deployment over telemetry-based packet-optical network," *IEEE Journal on Selected Areas in Communications*, pp. 1–1, 2021.

[11] J. Liu, C. Maltzahn, C. Ulmer, and M. Leon Curry, "Performance characteristics of the BlueField-2 SmartNIC," *arXiv:2105.06619*, 2021.

[12] F. Cugini, P. Gunning, F. Paolucci, P. Castoldi, and A. Lord, "P4 in-band telemetry (INT) for latency-aware VNF in metro networks," in *Optical Fiber Communication Conference (OFC) 2019*. Optical Society of America, 2019, p. M3Z.6. [Online]. Available: http://www.osapublishing.org/abstract.cfm?URI=OFC-2019-M3Z.6

[13] L. Tan, W. Su, W. Zhang, J. Lv, Z. Zhang, J. Miao, X. Liu, and N. Li, "In-band network telemetry: A survey," *Computer Networks*, vol. 186, p. 107763, 2021.

[14] D. Scano, F. Paolucci, K. Kondepu, A. Sgambelluri, L. Valcarenghi, and F. Cugini, "Extending P4 in-band telemetry to user equipment for latency- and localization-aware autonomous networking with AI forecasting," *Journal of Optical Communications and Networking*, vol. 13, no. 9, pp. D103–D114, 2021.

[15] D. Scano, F. Paolucci, K. Kondepu, A. Sgambelluri, L. Valcarenghi, P. Castoldi, and F. Cugini, "Augmented in-band telemetry to the user equipment for beyond 5G converged packet-optical networks," in *2020 European Conference on Optical Communications (ECOC)*, 2020.

[16] C. Guo, "Pingmesh: A large-scale system for data center network latency measurement and analysis," in *SIGCOMM*. ACM, August 2015.

[17] Y. Lin, Y. Zhou, Z. Liu, K. Liu, Y. Wang, M. Xu, J. Bi, Y. Liu, and J. Wu, "Netview: Towards on-demand network-wide telemetry in the data center," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1–6.

[18] F. Alhamed, D. Scano, P. Castoldi, F. Paolucci, F. Cugini, I. Vershkov, and J. Vegas Olmos, "P4 Postcard Telemetry Collector in Packet-Optical Networks," in *Optical Network Design and Modeling (ONDM), 2022 International Conference on*. IEEE, 2022, pp. 1–3.

[19] L. Velasco, A. C. Piat, O. Gonzlez, A. Lord, A. Napoli, P. Layec, D. Rafique, A. D'Errico, D. King, M. Ruiz, F. Cugini, and R. Casellas, "Monitoring and data analytics for optical networking: Benefits, architectures, and use cases," *IEEE Network*, vol. 33, no. 6, pp. 100–108, 2019.

[20] F. Musumeci, A. Fidanci, F. Paolucci, F. Cugini, and M. Tornatore, "Machine-Learning-Enabled DDoS attacks detection in p4 programmable networks," *Journal of Network and Systems Management*, vol. 30, no. 21, 2022.

[21] M. Dimolianis, A. Pavlidis, and V. Maglaris, "Signature-based traffic classification and mitigation for ddos attacks using programmable network data planes," *IEEE Access*, vol. 9, pp. 113 061–113 076, 2021.

[22] F. Paolucci, L. De Marinis, P. Castoldi, and F. Cugini, "Demonstration of P4 neural network switch," in *2021 Optical Fiber Communications Conference and Exhibition (OFC)*, 2021.