A Path-Aware Routing for Data Intensive Science: Proposal, Deployment and Evaluation in High-Performance Testbed

Daniel Ventorim [1,2], Everson S. Borges[1,2], Rafael S. Guimarães[2], Magnos Martinello[1], Moises R. N. Ribeiro [1], Cristina K. Dominicini[2], Marcos Schwarz[5], Bruno Xavier[2], Jeronimo Bezerra[6], Mariam Kiran[3] and Harvey Newman[4]

[1] Universidade Federal do Espírito Santo (UFES)

{daniel.ventorim,everson,magnos.martinello, moises.ribeiro}@ufes.br

[2]Instituto Federal do Espírito Santo (IFES)

{rafael.guimaraes, cristina.dominicini, bruno.xavier}@ifes.edu.br

[3]Oak Ridge National Laboratory, USA (kiranm@ornl.gov)

[4]Caltech, California Institute of Technology, USA (newman@hep.caltech.edu)

[5]Rede Nacional de Ensino e Pesquisa (RNP) (marcos.schwarz@rnp.br)

[6]Florida International University (FIU), USA (jeronimo@fiu.edu)

Abstract. *This paper introduces a path-aware architecture designed to enhance traffic engineering through path controllability and visibility for Data-Intensive Science (DIS). The approach relies on a stateless core network architecture that leverages line-rate packet forwarding reusing the cyclic redundancy check (CRC) implemented with P4 language. A path-aware approach was developed to enable path selection and hardware network metric collection. As a proof of concept, our approach was deployed in a 100+ Gbps multipath intra-datacenter topology. The evaluation includes experiments on i) a 100+ Gbps multipath intra-datacenter topology and ii) inter-datacenter networking across a Pan-American hybrid infrastructure connecting Vitória, ES, Brazil, and Atlanta, GA, USA. The traffic steering is demonstrated for intra-datacenter at the Caltech testbed, where concurrent flows achieve approximately 40 Gbps each at line rate. Inter-datacenter communication is evaluated over a 10,000 km link as part of a record-breaking 10 Tbps traffic stress at Supercomputing 2024.*

Resumo. *Este artigo apresenta uma arquitetura de rede ciente do caminho projetada para aprimorar a engenharia de tráfego por meio de controlabilidade e visibilidade do caminho em ciência intensiva de dados (DIS). A abordagem se baseia em uma arquitetura de rede com núcleo sem estado que baseia-se em encaminhamento de pacotes em taxa de linha reutilizando o cyclic redundancy check (CRC) implementado em linguagem P4. Uma abordagem path-aware foi desenvolvida para habilitar a seleção de caminho e a coleta de métricas de rede em hardware. Como prova de conceito, nossa abordagem foi implantada em uma topologia intra-datacenter multicaminho de 100+ Gbps. A avaliação inclui experimentos em i) uma topologia intra-datacenter multicaminho de 100+ Gbps e ii) inter-datacenter em uma infraestrutura híbrida pan-americana conectando Vitória, ES, Brasil, e Atlanta, GA, EUA. O direcionamento de tráfego*

é demonstrado para intra-datacenter no testbed da Caltech, onde fluxos simultâneos atingem aproximadamente 40 Gbps cada em line rate. A comunicação entre datacenters é avaliada em um link de 10.000 km como parte de um estresse de tráfego recorde de 10 Tbps na Supercomputing 2024.

1. Introduction

Data-intensive sciences (DIS) are gaining increasing importance as they drive the exploration of fundamental scientific questions by collecting and processing vast amounts of scientific data [Dunefsky et al. 2022]. In the context of large-scale data-intensive sciences, scientists collaborate on various projects to design infrastructure for data generation, establish data centers for storage and processing, and, develop a network infrastructure for efficient data transfer. An exemplary instance is the global system implemented by *Conseil Européen pour la Recherche Nucléaire* (CERN), featuring 170 data centers distributed across 40 countries [Newman et al. 2003].

As demand for high-performance networks continues to grow, the need for innovative solutions in the field of computer networking becomes increasingly critical [Barisits et al. 2019]. However, the network layer provides a best-effort service to the endpoints, with no control of the path properties between the endpoints. The path is assumed to be invisible, homogeneous, singular, with dynamics determined by the connectivity of the endpoints. Endpoints have very little information about the paths over which their traffic is carried and no control at all beyond the destination address.

Over the last 20 years, a variety of path-aware Internet architectures has been proposed to enhance transparency and provide end hosts with more control over the network paths [Barrera et al. 2017, Anderson et al. 2014]. Unlike the current Internet, where end hosts have no influence on routing decisions for their packets, path-aware networks expose routing paths to end hosts, enabling them to specify the desired path within the packet header (packet-carried forwarding state). This approach ensures that the data plane enforces compliance with these source-defined paths as packets traverse the network.

Despite efforts by the IETF, which established the path-aware Networking Research Group (PANRG) [1], and the Internet Research Task Force (IRTF) to raise awareness of path-aware networking within the protocol design community, no standardized protocol is currently ready for deployment to fully support path-aware routing. As noted, "Even if endpoints cannot control the path, at least they can determine the properties of the path by choosing among paths available to them." However, existing solutions still fall short of meeting the demanding requirements of data-intensive science, particularly in maintaining *high throughput over long-distance networks* .

The most common solutions rely on static routing mechanisms, where switches' tables are preconfigured to define paths (e.g., MPLS [Viswanathan et al. 2001]) or depend on a centralized architecture, less scalable and susceptible to failure in the central controller (e.g., Software-Defined Network for End-to-end Networked Science at the Exascale (SENSE) [Monga et al. 2020]). While designed for compatibility with various Software Defined Network (SDN) systems and routing protocols, SENSE demonstrates advanced network orchestration and path-aware routing capabilities. However, its re-

[1] https://datatracker.ietf.org/rg/panrg/about/

liance on a centralized intelligent orchestration model introduces potential single points of failure, posing scalability challenges in large-scale networks.

The contribution of this paper is an innovative approach to enabling path-awareness for DIS. This is achieved through the following architectural components:

1. Edge - that supports the control and stores the network state with its policy;
2. Stateless core - that achieves line-rate packet forwarding by a technique that reuses the CRC mechanisms implemented with the P4 language [Dominicini et al. 2020a];
3. Management layer - that outlines the interfaces, messages and a dashboard that provides path visibility and path controllability;
4. Path-Awareness in Multipath Environments - supports hosts with multiple interfaces by enabling explicit path selection and steering across available routes.

This enables the high-performance DIS networking with enhanced path controllability and path visibility. As proof of concept, our approach was deployed in a 100+ Gbps multipath intra-datacenter topology. The evaluation encompasses experiments on: (i) a 100+ Gbps multipath intra-datacenter topology, and (ii) inter-datacenter networking across a Pan-American hybrid infrastructure linking Vitória, ES, Brazil, with Atlanta, GA, USA [GÉANT GP4L 2024]. The approach demonstrates seamless traffic steering for intra-datacenter traffic engineering and evaluates inter-datacenter communication over a 10,000 km link, as part of a record-breaking 10 Tbps traffic stress experiment showcased at Supercomputing 2024.

The remainder of this paper is structured as follows. Section 2 provides background information and reviews the state of the art in path-aware networking. The deployment of the path-aware solution in an experimental testbed is detailed in Section 3. Section 4 introduces the proposed architecture and its implementation. Section 5 presents the experimental results, followed by a discussion on the future implications for high-performance networks in Section 6.

2. Background and state of the art

Path-aware networking is an emerging architectural paradigm that seeks to enhance traditional network communication by explicitly exposing available network paths and their properties to endpoints [Scherrer et al. 2021], enabling informed path selection and visibility. Unlike conventional networks, where routing decisions are abstracted from the user and often rely on single best-path paradigms, path-aware networks empower applications and operators with greater control over traffic flows. This control helps meet specific performance, reliability, and trust requirements, offering exciting properties to applications, such as path transparency, and fine-grained path control [Sunshine 1977].

At the core of path-aware networking lies the ability to expose and utilize multiple paths between the endpoints. This approach builds upon foundational work in network programmability and traffic engineering but extends it by explicitly incorporating path properties, such as latency, bandwidth, or trust, into the decision-making process at endpoints. The RFC 9217 [2] highlights the potential of segment routing and source routing

[2]https://www.rfc-editor.org/rfc/rfc9217.html

architectures in achieving path-awareness by embedding routing decisions within packet headers, enabling efficient and scalable control of traffic flows [Trammell 2022].

This lack of path-aware and control is a significant challenge for DIS networks, where predictable and reliable network performance is crucial. Virtual Circuits (VCs) represent a fundamental technology for overcoming these challenges and achieving the network performance required in DIS [Newman et al. 2015]. VCs establish dedicated network paths with guaranteed bandwidth between specific endpoints, emulating point-to-point connections within a packet-switched network [Johnston et al.]. These virtual circuits are typically established using *a static routing* mechanism, where switching tables are pre-configured to define the path (e.g., MPLS and OpenFlow). However, DIS applications require more dynamic solutions, considering the nature of DIS workloads, which often involve large data transfers between multiple locations [Zurawski et al. 2021].

In the context of SENSE services, the "network" encompasses both the switching and routing elements as well as the network stacks of the end systems, such as Data Transfer Nodes within Science Demilitarized Zone (DMZ) facilities. The data plane capabilities tied to these services include Layer 2 point-to-point with Quality of Service (QoS), Layer 2 multi-point with QoS, and Layer 3 Virtual Private Network (VPN) and Flow QoS [Monga et al. 2020]. Nevertheless, the system's dependence on a centralized intelligent orchestration model could create vulnerabilities, as it introduces potential single points of failure and may hinder scalability in expansive network environments.

In this direction, Source Routing (SR) schemes deserve attention in the context of DIS because they can manage virtual circuits dynamically more efficiently than traditional protocols [Jyothi et al. 2015]. This is achieved by updating rules only at the network edge and inserting a route label in the packet header to select a path [Jin et al. 2016]. Furthermore, source routing reduces control overhead and latency related to path establishment convergence, since path migration is simply a matter of changing the state at the source [Dominicini et al. 2020b]. In summary, our proposal introduces a 2.5 layer functionality by incorporating source routing tunnels at the IP layer, enabling enhanced routing controllability. Furthermore, we advance towards path-awareness by providing link and path visibility so that informed routing decisions can be taken by server applications in the context of DIS.

2.1. Technology for Source Routing in DIS

The traditional way of performing source routing is the list-based approach. This approach represents the route label or identifier (*routeID*) as a list of ports or addresses. The forwarding operation is represented as a pop to rewrite the route label, updating the list [Dominicini et al. 2020b]. Although this approach drastically reduces the burden of managing network states by eliminating tables in core nodes, it still needs to maintain a state in the packet by using a route label rewrite operation in each node to update the list. This operation can be costly for packet networks and difficult to implement in optical networks [Wessing et al. 2002]. In addition, a classical problem in SR is how fast it reacts to node or link failures [Gomes et al. 2016].

Currently, the most widespread list-based source routing protocol is Segment Routing [Ventre et al. 2020], but it has some limitations. Although a promising solution, Segment Routing is not widely supported by all network equipment and vendors.

Its implementation is costly, as it requires significant reconfiguration of existing network infrastructure, including support for new protocols and mechanisms such as MPLS or IPv6. Its MPLS version still depends on tables in the core nodes and variable headers, which limits the number of maximum hops that can be implemented in hardware switches [Al-Najjar et al. 2024].

To overcome these challenges, some SR approaches [Liberato et al. 2018, Dominicini et al. 2020b] utilize the Residue Number System (RNS), where the output port is determined by calculating the remainder of the division (i.e., a modulo operation) between the route label and the node identifier. An essential characteristic of this strategy is performing fully stateless forwarding, meaning that no changes are made to packet headers during the process [Dominicini et al. 2020b]. Related works have integrated this scheme with SDN [Martinello et al. 2014], developed fast failure recovery mechanisms [Gomes et al. 2016], investigated techniques to improve *routeID* scalability [Ren et al. 2017], and applied the scheme to enable traffic redirection [Dominicini et al. 2020c]. However, all these works rely on integer RNS arithmetic, and the integer modulo operation cannot be implemented in current network hardware [Dominicini et al. 2020b]. Therefore, they utilize software switch implementations [Martinello et al. 2014, Dominicini et al. 2020c], or rely on the synthesis of integer division for ASICs or NetFPGAs [Liberato et al. 2018].

The PolKA protocol [3] has been tested on programmable commercial equipment distributed in Europe, connected by 10Gbps links [Dominicini et al. 2021], with the same data plane performance as traditional forwarding methods. Subsequently, the creation of an overlay network with PolKA tunnels was demonstrated to validate intensive data transfer at 10Gbps and 100Gbps in a testbed composed of programmable switches in Europe, the United States, and Brazil [Borges et al. 2022].

3. Deployment of a Path-Aware Solution in Experimental Testbeds

Since the earliest editions of the Supercomputer Conference, Caltech has played a prominent role, particularly by conducting real-time transmission tests and fostering collaboration among research groups from around the globe. In 2024, Caltech showcased a robust infrastructure comprising 10 DTN servers of various models primarily used for data transmission, 2 switches with 400 Gbps capacity, 5 programmable switches, and an additional switch dedicated to network management, as shown in Figure 1 component (b) High capacity conventional network.

The topology illustrated in Figure 1 component (a) Caltech/UFES/IFES at P4 testbed at Supercomputing. It comprises two Dell R7030XD servers and five programmable switches based on Intel chipsets. This infrastructure enabled the execution of path-awareness experiments directly at the conference booth. With support from RNP, an efficient communication channel was also established between UFES and Caltech's booth Figure 1 component (c), utilizing the production network for data transmission tests.

In the development and testing of advanced networking protocols, testbeds serve as indispensable tools, providing controlled environments to simulate real-world network conditions [Borges et al. 2022]. These platforms enable researchers to evaluate the performance and scalability of innovative designs, such as path-aware networking, where

[3] https://nerds-ufes.github.io/polka/

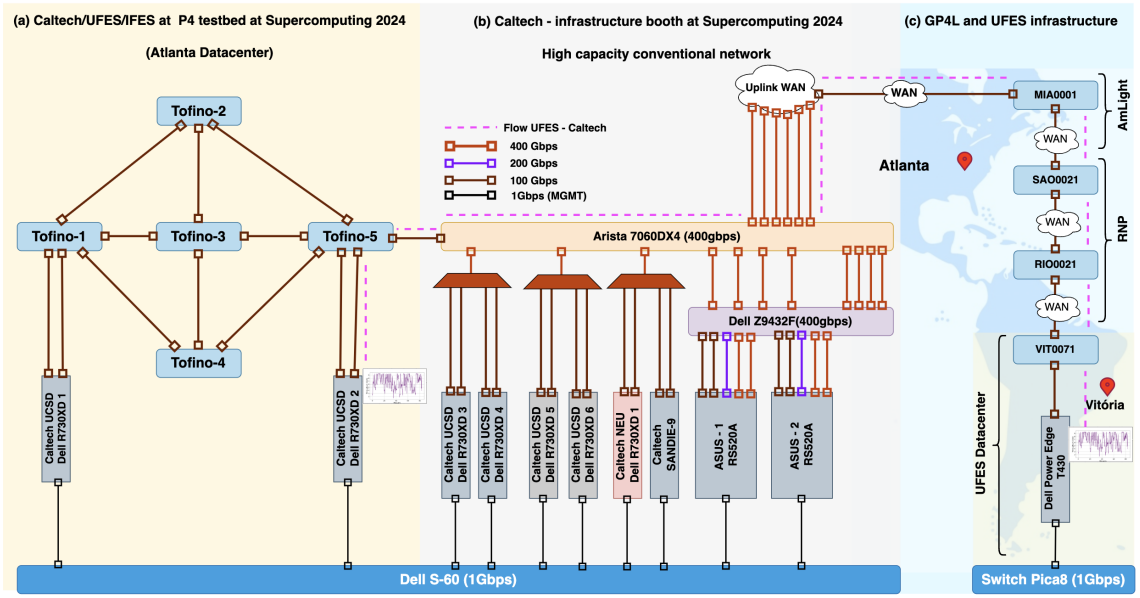

Figure 1. Infrastructure of the Caltech booth with a GP4L conection

routers and endpoints dynamically adjust traffic based on the characteristics of available paths. By leveraging testbeds, researchers can use diverse network path scenarios, such as varying latency, packet loss, and bandwidth constraints, numbers of hops which are critical for validating the robustness of path-aware solutions. This approach not only accelerates the innovation cycle but also ensures the practicality of these protocols in complex, real-world deployments.

A cutting-edge topology was assembled using intel Tofino switches to facilitate advanced experimentation. This setup demonstrated the potential of Source Routing with a stateless core, enabling efficient and scalable data forwarding. By leveraging the capabilities of P4-programmable switches, the deployment achieved forwarding at line rate by reusing CRC [Dominicini et al. 2020a], ensuring high-performance operation without compromising accuracy.

The topology, implemented on Caltech/UFES/IFES at P4 testbed, utilized Tofino switches to validate real-world applicability. The network design featured multiple paths available in the topology Figure 2 and corresponding to Figure 1 component (a), allowing for dynamic traffic distribution and enhanced fault tolerance. With its easy-to-configure tunnels and seamless integration into the freeRtr OS platform, the solution showcased both operational simplicity and flexibility, paving the way for innovative approaches in high-performance networking.

4. Proposed Architecture for DIS Path-Aware Networking

In Figure 3, we present an overview of the path-aware Architecture, highlighting three key physical components: two Intel Barefoot Tofino switches and a DTN (Data Transfer Node) server. In terms of management plane, there are inband and out-of-band messages. PolKA routing messsages (red color) implemented in the freeRtr OS are related to the data plane configuration that consolidates the network state distribution based on ordinary IGP protocol (e.g. OSPF). The out-of-band messages (green color) are implemented for the

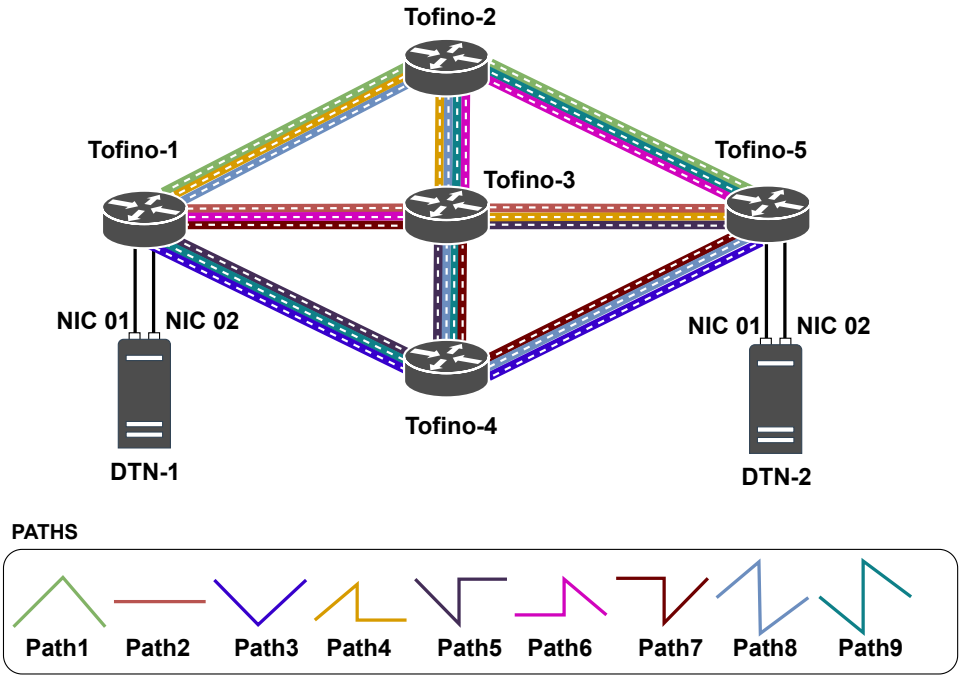

Figure 2. intra-domain Caltech testbed at the booth of supercomputing 2024

support of path-aware approach, collecting diverse metrics (e.g. P4 hardware counters) from data plane and also exposed by freeRtr OS agents.

The endhost interacts with the path-aware Controllability through a REST API provided by freeRtr. This interaction enables users to perform path selection, link occupancy queries, and real-time network performance monitoring. When a user initiates a path selection, requests link utilization data, or retrieves network metrics, the *dashboard* communicates with the REST API, which queries freeRtr OS and returns the relevant information. This design ensures that the system remains accessible and responsive in both local and remote environments.

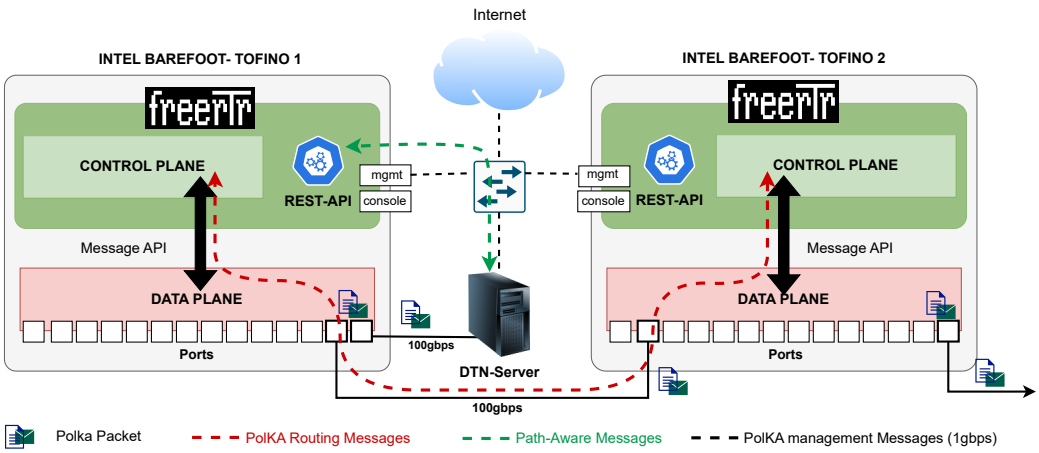

Figure 3. Path-Aware Architecture

A path-aware controllability presented in Figure 4 was developed to provide a control over network routing, allowing users to select and modify paths as needed. This intuitive interface leverages the freeRtr REST API, enabling seamless interaction with the network topology. Through this dashboard, users can manage paths directly from any of the DTNs available in the topology, ensuring flexibility and ease of operation. This functionality enhances the adaptability of the network, enabling real-time adjustments to optimize performance and accommodate changing traffic demands.

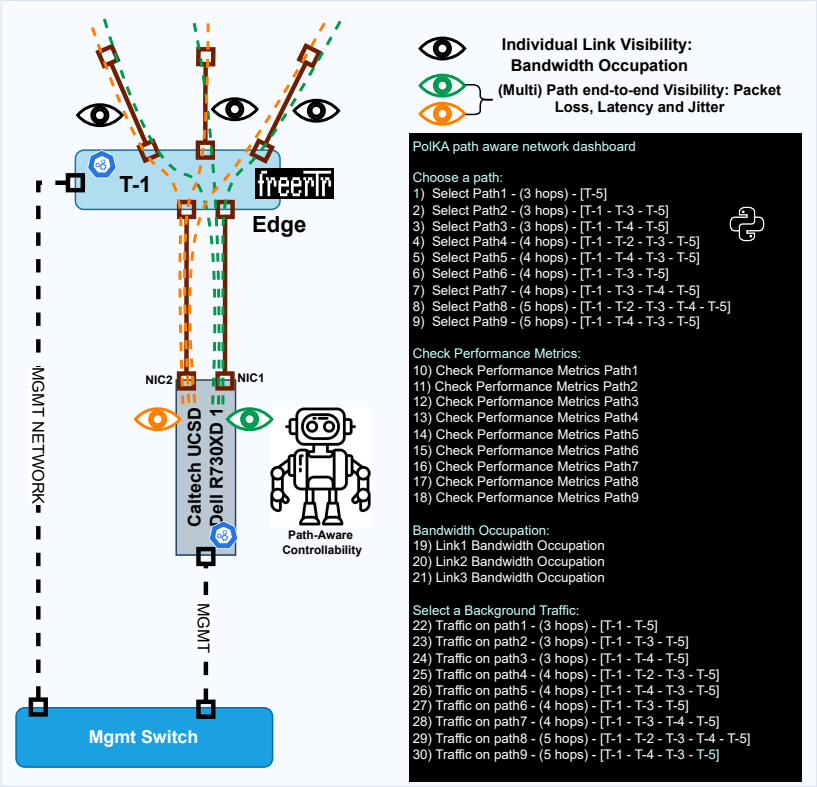

Figure 4. Network Visibility and Path Controllability

To facilitate path selection decisions, the dashboard provides path metrics via the "Check Performance Metrics Path" option, which can be used in conjunction with the "Link Bandwidth Occupation" option to assess path performance and inform routing choices. This allows users to not only dynamically switch between paths but also to understand the current performance characteristics of each available path, enabling more strategic traffic engineering and optimization. The dashboard implementation and the data from the experimentations are available in a public repository[4]

When a different path is selected for network flows using the Path-Aware dashboard, the flows can be distributed across the available paths, as highlighted in the topology shown in Figure 4. This capability allows better load balance and optimized resource utilization within the network. By leveraging the dashboard's intuitive interface, operators

[4]The path-aware dashboard and experimental data can be accessed at: https://github.com/eversonscherrer/sc2024/

can efficiently manage traffic, distributing flows across two distinct paths, thus improving overall performance and reducing potential bottlenecks.

Furthermore, the Path-Aware Controllability offers a comprehensive suite of features for monitoring and managing network performance. It enables endhost users to check performance metrics for each path, such as throughput, latency, and packet loss, as well as to analyze the link occupation, providing critical insights into network utilization. All these functionalities are seamlessly integrated by leveraging the capabilities of P4 counters, which collect and aggregate real-time data directly from programmable switches. This integration ensures precise and up-to-date visibility into the network's behavior, facilitating data-driven decision-making. By combining advanced monitoring with dynamic path selection, the dashboard not only enhances operational efficiency but also supports the optimization of traffic flows under varying network conditions, thereby addressing the demands of modern high-performance networking environments.

5. Performance Evaluation of Path-Aware Routing

This section presents a detailed evaluation of the proposed path-aware networking solution, powered by the PolKA protocol, covering its design, prototyping, validation, deployment, and integration within a production-grade testbed showcased at Supercomputing Conference 2024. To enable real-time traffic monitoring and performance analysis, the InMon Traffic Monitor Tool [InMon Corporation 2023] was utilized, providing detailed visualization of flow characteristics, congestion trends, and bandwidth utilization across multiple network segments. Complementing this, Fast Data Transfer (FDT) [Harutyunyan et al. 2009] facilitated high-speed, long-distance data movement, leveraging parallel TCP streams and direct memory access (DMA) to optimize throughput and efficiency. This combination of monitoring and transfer technologies was particularly suited for large-scale scientific and research environments, ensuring efficient bulk data transfers while maintaining network stability and performance.

5.1. Datacenter intra-domain demonstration at Caltech Booth

This subsection evaluates the effectiveness of a path-awareness solution in addressing the limitations of traditional Interior Gateway Protocols (IGPs), which generally use a single path for forwarding multiple traffic flows. To overcome these constraints, we integrate the path-aware dashboard with PolKA (Polynomial Key-Based Architecture). PolKA employs modular arithmetic operations at each hop to determine packet forwarding decisions, enabling dynamic, fine-grained traffic distribution across multiple available paths and improving overall network efficiency.

We conducted an experiment by generating traffic between DTN_1 and DTN_2 using FDT[5], as illustrated in Figures 5 and 6. Each DTN node is equipped with two 100 Gbps network interfaces. As depicted in Figure 7, at time $t = 0$ second, we initiated $Flow_1$ from DTN_1 through interface 1 (NIC_1), achieving a throughput of approximately 70 Gbps. Although the NIC supports 100Gbps, the actual throughput is limited by the constraints of the traffic generator. At $t = 5$ seconds, we launched $Flow_2$ from DTN_1 via interface 2 (NIC_2). Subsequently, when $Flow_1$ and $Flow_2$ began sharing the same

[5] https://github.com/fast-data-transfer/fdt/

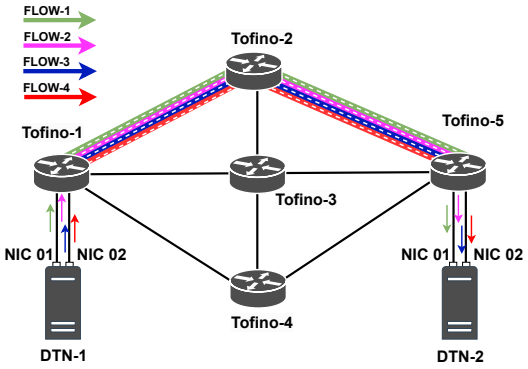

Figure 5. Without Path-Aware

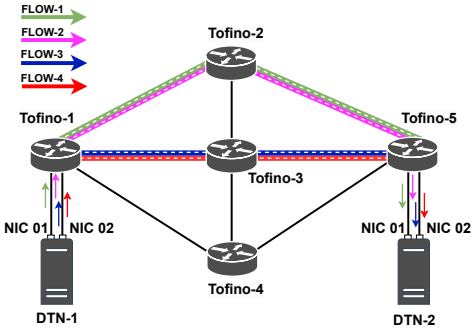

Figure 6. With Path-Aware

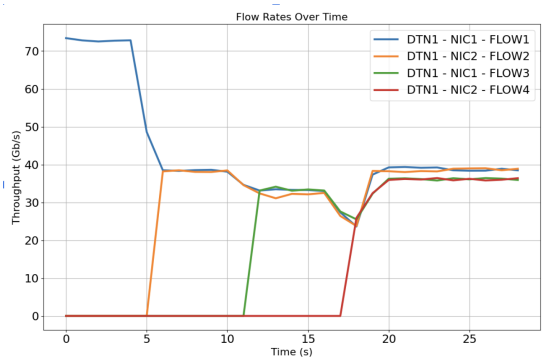

Figure 7. After change paths

path (Tofino_1→Tofino_2→Tofino_5), the throughput of each flow decreased to approximately 40 Gbps due to congestion.

At $t = 11$ seconds, we introduced $Flow_3$ from DTN_1 through NIC_1. $Flow_3$ remained on the same path, resulting in all three flows ($Flow_1$, $Flow_2$, and $Flow_3$) stabilizing at around 35 Gbps. At $t = 16$ seconds, we initiated $Flow_4$ from DTN_1 via NIC_2. At this point, using our path-aware mechanism from the endpoint, $Flow_3$ and $Flow_4$ were rerouted to an alternative path (Tofino_1→Tofino_3→Tofino_5). This led to an increase in throughput for all flows to approximately 40 Gbps, demonstrating the solution's ability to optimize resource utilization and enhance the overall network performance. The current multipath approach used in testbeds still lacks support for splitting a flow across multiple paths simultaneously and reassembling it at the edge, in contrast to the multipath routing implemented and validated in virtual switches [Guimarães et al. 2022].

Furthermore, the path-aware visibility is able to improve the network utilization by enabling the selection of alternative, uncongested paths, thereby reducing congestion and maximizing resource efficiency. To further illustrate this dynamic behavior, a real-time demonstration video[6] is provided, showcasing the system's ability to adaptively reroute traffic and its impact on throughput and path optimization.

[6]The video demonstrating the setup running can be accessed at: https://www.youtube.com/watch?v=hcZVJlXCwHE&t=220s

5.2. Performance Evaluation in a Long-Distance WAN Datacenter Inter-Domain: UFES to Atlanta

During the Global Data Transmission Speed Challenge, held at the SuperComputing 2024 high performance computing event, UFES, in collaboration with institutions such as Caltech, Amlight, ESNET, Internet2, RNP, and CERN, achieved the impressive milestone of 10 terabits per second (Tbps) in long-distance data transfer highlighted in Figure 8.

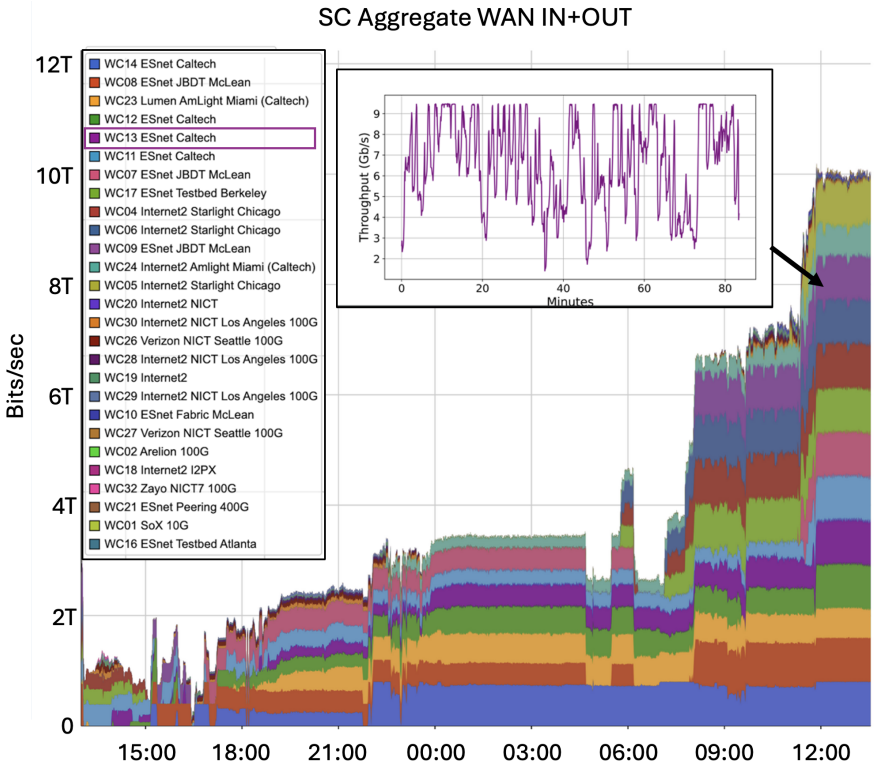

Figure 8. SC Aggregate WAN IN+OUT

GP4L (Global Platform for Lab) is a global overlay network built on top of 100G/400G programmable switches and SmartNICs from the GNA-G Auto-GOLE/SENSE Testbed and the GÉANT P4 Lab. It offers routing services through SONiC and RARE/freeRtr—fully open-source network operating systems (NOSes) that support both traditional networking protocols and extensibility for next-generation features. GP4L enables experimentation with production-oriented features and protocols in a realistic and scalable environment.

The experiment, the first pan-american high throughput transmission using PolKA, involved a continuous traffic flow transmission using FDT tool for approximately 90 minutes. Figure 8 presents an analysis of a representative one-hour obtained from InMon Traffic Monitor. The observed results indicate an average throughput of 6 Gbps, limited by the use of shared production links with transmission peaks reaching up to 9.5 Gbps, effectively achieving the maximum available capacity.

6. Conclusion

This work demonstrates significant advancements in the field of programmable networks for Data Intensive Science (DIS) by leveraging a stateless core architecture enabled

through P4 programability and CRC reuse. Our proposal offers both path visibility and path controllability, addressing key requirements for scalable and intelligent network operations. By adopting a Path-Aware Networking paradigm, the system enables flexible and agile path reconfiguration, providing a robust framework for traffic engineering and facilitating optimal flow allocation to end hosts.

Our approach demonstrated effectiveness in mitigating congestion, allowing multiple concurrent flows to achieve approximately 40 Gbps each in an intra-datacenter scenario. Resilience and efficiency over long distances were confirmed in inter-domain tests across a 10,000 km link, where the solution sustained an average throughput of 6 Gbps with peaks up to 9.5 Gbps, achieving line rate.

The implementation has explored source routing with a stateless core and the reuse of CRC in Tofino switches. It has proven its practical viability in the testbed at the SuperComputing 2024 Caltech booth. The integration of the PolKA protocol with the freeRtr OS platform showcases advanced capabilities for dynamic tunnel configuration and operational flexibility. The deployed path-aware dashboard provides an intuitive control interface, enabling real-time path selection and routing decisions with enhanced network visibility.

Future work will explore the application of this approach in different DIS scenarios, investigating protocol optimization and the development of new functionalities for the dashboard, integrating algorithms for optimizing route selection.

7. Acknowledgments

Financial support from Brazilian agencies: CNPq, CAPES, FAPESP/MCTI/CGI.br (PORVIR-5G 20/05182-3), FAPES (515/2021, 284/2021, 06/2022, 1026/2022, 941/2022, 20/2022). Prof. Magnos Martinello is supported by a CNPq Productivity Scholarship, grant number 312058/2023-3.

References

[Al-Najjar et al. 2024] Al-Najjar, A. et al. (2024). Framework for integrating machine learning methods for path-aware source routing. In *SC24-W: Workshops of the International Conference for High Performance Computing, Networking, Storage and Analysis*, pages 829–838. IEEE.

[Anderson et al. 2014] Anderson, T. et al. (2014). A brief overview of the nebula future internet architecture. *SIGCOMM Comput. Commun. Rev.*, 44(3):81–86.

[Barisits et al. 2019] Barisits, M., Beermann, T., Berghaus, F., Bockelman, B., Bogado, J., Cameron, D., Christidis, D., Ciangottini, D., Dimitrov, G., Elsing, M., et al. (2019). Rucio: Scientific data management. *Computing and Software for Big Science*, 3:1–19.

[Barrera et al. 2017] Barrera, D., Chuat, L., Perrig, A., Reischuk, R. M., and Szalachowski, P. (2017). The scion internet architecture. *Commun. ACM*, 60(6):56–65.

[Borges et al. 2022] Borges, E., Pontes, E., Mate, C., Loui, F., Martinello, M., and Ribeiro, M. (2022). Freerouter in a nutshell: A protocoland routing platform for open and portable carrier-class testbeds. In *Anais do I Workshop de Testbeds*, pages 36–46, Porto Alegre, RS, Brasil. SBC.

[Dominicini et al. 2020a] Dominicini, C. et al. (2020a). Polka: Polynomial key-based architecture for source routing in network fabrics. In *2020 6th IEEE Conference on Network Softwarization (NetSoft)*, pages 326–334. IEEE.

[Dominicini et al. 2021] Dominicini, C., Guimarães, R., Mafioletti, D., Martinello, M., Ribeiro, M. R., Villaça, R., Loui, F., Ortiz, J., Slyne, F., Ruffini, M., et al. (2021). Deploying polka source routing in p4 switches. In *2021 International Conference on Optical Network Design and Modeling (ONDM)*, pages 1–3. IEEE.

[Dominicini et al. 2020b] Dominicini, C., Mafioletti, D., Locateli, A. C., Villaca, R., Martinello, M., Ribeiro, M., and Gorodnik, A. (2020b). Polka: Polynomial key-based architecture for source routing in network fabrics. In *2020 6th IEEE Conference on Network Softwarization (NetSoft)*, pages 326–334. IEEE.

[Dominicini et al. 2020c] Dominicini, C. K. et al. (2020c). KeySFC: Traffic steering using strict source routing for dynamic and efficient network orchestration. *Computer Networks*, 167:106975.

[Dunefsky et al. 2022] Dunefsky, J. et al. (2022). Transport control networking: Optimizing efficiency and control of data transport for data-intensive networks. In *Proceedings of the ACM SIGCOMM Workshop on Network-Application Integration*, NAI '22, page 60–66, New York, NY, USA. Association for Computing Machinery.

[Gomes et al. 2016] Gomes, R. R. et al. (2016). KAR: Key-for-any-route, a resilient routing system. In *2016 46th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshop*, pages 120–127.

[Guimarães et al. 2022] Guimarães, R. S. et al. (2022). M-polka: Multipath polynomial key-based source routing for reliable communications. *IEEE Transactions on Network and Service Management*, pages 1–1.

[GÉANT GP4L 2024] GÉANT GP4L (2024). Gp4l documentation. Accessed: Jan 2025.

[Harutyunyan et al. 2009] Harutyunyan, H., Swenson, C., and Newman, H. B. (2009). Fdt: A high-speed data transfer tool for scientific applications. *International Journal of High Performance Computing Applications*, 23(3):254–267.

[InMon Corporation 2023] InMon Corporation (2023). *InMon Traffic Monitor Tool*. Accessed: Jan 2025.

[Jin et al. 2016] Jin, X., Farrington, N., and Rexford, J. (2016). Your data center switch is trying too hard. In *Proceedings of the Symposium on SDN Research*, pages 1–6.

[Johnston et al.] Johnston, W., Dart, E., and Tierney, B. Addressing the problem of data mobility for data-intensive science. In *Proceedings of the Eighth International Conference on Engineering Computational Technology*.

[Jyothi et al. 2015] Jyothi, S. A., Dong, M., and Godfrey, P. B. (2015). Towards a flexible data center fabric with source routing. In *Proceedings of the 1st ACM SIGCOMM Symposium on Software Defined Networking Research*, pages 1–8.

[Liberato et al. 2018] Liberato, A. et al. (2018). RDNA: Residue-Defined Networking Architecture Enabling Ultra-Reliable Low-Latency Datacenters. *IEEE TNSM*.

[Martinello et al. 2014] Martinello, M., Ribeiro, M. R., de Oliveira, R. E. Z., and de Angelis Vitoi, R. (2014). Keyflow: a prototype for evolving sdn toward core network fabrics. *IEEE Network*, 28(2):12–19.

[Monga et al. 2020] Monga, I., Guok, C., MacAuley, J., Sim, A., Newman, H., Balcas, J., DeMar, P., Winkler, L., Lehman, T., and Yang, X. (2020). Software-defined network for end-to-end networked science at the exascale. *Future Generation Computer Systems*, 110:181–201.

[Newman et al. 2015] Newman, H., Mughal, A., Kcira, D., Legrand, I., Voicu, R., and Bunn, J. (2015). High speed scientific data transfers using software defined networking. In *Proceedings of the Second Workshop on Innovating the Network for Data-Intensive Science*, pages 1–9.

[Newman et al. 2003] Newman, H. B., Ellisman, M. H., and Orcutt, J. A. (2003). Data-intensive e-science frontier research. *Commun. ACM*, 46(11):68–77.

[Ren et al. 2017] Ren, Y. et al. (2017). Flowtable-free routing for data center networks: A software-defined approach. In *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, pages 1–6.

[Scherrer et al. 2021] Scherrer, S., Legner, M., Perrig, A., and Schmid, S. (2021). Enabling Novel Interconnection Agreements with Path-Aware Networking Architectures. In *Proceedings of the 51st Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, pages 1–12. IEEE.

[Sunshine 1977] Sunshine, C. A. (1977). Source routing in computer networks. *ACM SIGCOMM Computer Communication Review*, 7(1):29–33.

[Trammell 2022] Trammell, B. (2022). Current open questions in path-aware networking. RFC 9217, IRTF.

[Ventre et al. 2020] Ventre, P. L. et al. (2020). Segment routing: A comprehensive survey of research activities, standardization efforts, and implementation results. *IEEE Communications Surveys & Tutorials*, 23(1):182–221.

[Viswanathan et al. 2001] Viswanathan, A., Rosen, E. C., and Callon, R. (2001). Multiprotocol Label Switching Architecture. RFC 3031.

[Wessing et al. 2002] Wessing, H. et al. (2002). Novel scheme for packet forwarding without header modifications in optical networks. *Journal of Lightwave Technology*, 20(8):1277–1283.

[Zurawski et al. 2021] Zurawski, J., Brown, B., Carder, D., Colby, E., Dart, E., Miller, K., Patwa, A., Robinson, K., Rotman, L., and Wiedlea, A. (2021). 2020 high energy physics network requirements review final report.