SC26 Network Research Exhibition: Demonstration Publishable Abstract

# NRE112: Reimagining advanced capabilities in high-bandwidth research networks

Rafael Coelho Lopes de Sa, University of Massachusetts Amherst

**Abstract**

This paper presents a field-programmable gate array (FPGA)-based SmartNIC architecture designed to address critical network challenges in data-intensive scientific computing. Our approach focuses on implementing network visibility, optimization, and orchestration capabilities through FPGA acceleration. The goals are to develop a flexible firmware-centric implementation of high-bandwidth network functions on AMD/Xilinx SmartNICs, leveraging P4-based packet-processing methods with support for packet parsing, marking, telemetry, local steering, and control hooks for machine learning inference.

The next generation of data-intensive science will be constrained by networks as much as by compute and storage. Major computing sites are preparing for sustained wide-area traffic in the 400 Gb/s region and beyond. Visibility is needed to identify the owner and purpose of traffic; optimization is needed to make more efficient use of shared bandwidth; and orchestration is needed so workflows can request elevated service when required.

As a concrete example, we have implemented and evaluated a preliminary network telemetry architecture based on Scientific Network Tags (SciTags). SciTags associate network traffic with scientific context through packet marking. The FPGA-based bridge counts IPv6 packets according to the experiment information in their SciTag values, carried in the IPv6 Flow Label, and performs programmable virtual local area network (VLAN) translation. Flowd marks packets on the external nodes. Host software configures the tables and periodically reads the counters into a local database. Figure 1 shows the setup, implemented and evaluated on AMD Alveo U280 FPGAs in the Open Cloud Testbed (OCT) at the Massachusetts High Performance Compute Center (MGHPCC).

Using iperf3, we compared the FPGA-based bridge with a reference path between regular network-interface cards (NICs). Receiver throughput was 37.0 Gb/s for both paths on 40 Gb/s NICs, and 95.1 Gb/s for the reference versus 95.3 Gb/s for the bridge on 100 Gb/s NICs. These results show no observable throughput degradation. The FPGA counters agreed exactly with offline packet-capture analysis: ATLAS, 854,311 packets; CMS, 744,050; LHCb, 765,889; and ALICE, 841,103.

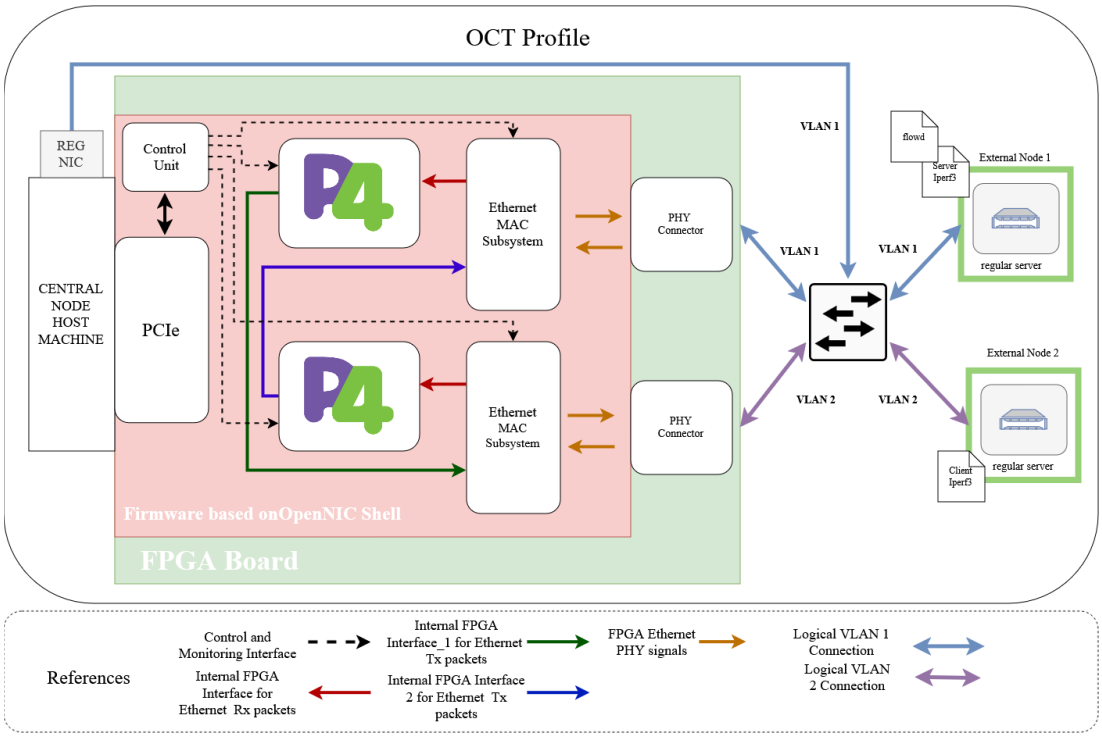



Figure 1. Network telemetry based on SciTags, tested on the OCT platform.

## Goals

1. Demonstrate real-time inspection and monitoring of SciTags-compatible IPv6 packet markings, together with programmable VLAN manipulation.
2. Preserve high-throughput packet handling while accurately monitoring scientific data flows at 40 Gb/s and 100 Gb/s. Figure 2 summarizes the implemented P4 program.
3. Extend the data plane with Differentiated Services Code Point (DSCP) marking and processing, currently under development, to support traffic prioritization and quality-of-service policies.
4. Develop the corresponding control plane to manage SciTags, VLAN, and DSCP operations according to configurable policies at the granularity of IPv6 prefixes or subnets.
5. Extend the experiments to AMD V80s and a 400 Gb/s network interface for line-rate packet processing.

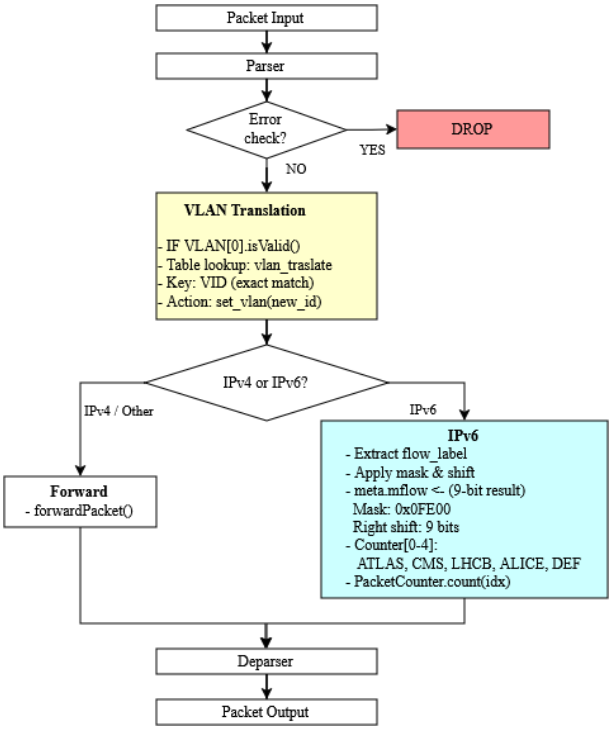



Figure 2. P4 program diagram for the network telemetry application.

## Impacts

This work establishes the foundation for an FPGA-based SmartNIC architecture with coordinated data-plane and control-plane capabilities for network visibility, optimization, and orchestration. Scientific traffic can be identified, measured, and associated with network services directly in the packet-processing path, without requiring host-CPU intervention for every packet.

These capabilities address the needs of physics experiments that couple distributed storage, high-performance computing facilities, and analysis platforms through heterogeneous wide-area networks. Scientific AI workflows create additional demand for low-latency, high-bandwidth movement of training data, model checkpoints, inference inputs, and derived products. Science-aware telemetry allows measurements to be correlated with a particular experiment, production campaign, or analysis activity.

The next stage is the development of a control plane that maintains the mapping between scientific activities and their required data-plane actions. It will configure SciTags identifiers and counters, associate services provisioned by the Software-Defined Network for End-to-end Networked Science at the Exascale (SENSE) with the appropriate VLANs, and apply DSCP policies to selected traffic classes. A programmatic interface will expose these capabilities and firmware-generated telemetry to external monitoring systems and network orchestrators. This separation preserves deterministic FPGA packet processing while allowing network policies to be created and retired dynamically in software.

## Resources

The evaluation used an AMD Alveo U280 FPGA bridge and two external nodes with regular NICs on separate VLANs. The 40 Gb/s configuration used a maximum transmission unit (MTU) of 1500 bytes; the 100 Gb/s configuration used an MTU of 9000 bytes. The FPGA design combines the OpenNIC Shell framework with AMD Vitis Networking P4. The software components are Flowd, iperf3, and the host application for configuring and reading the packet counters.

MGHPCC houses OCT and the ATLAS NET2 cluster. The planned extension uses AMD V80 FPGAs and 400 Gb/s connectivity.

## Involved Parties

University of Massachusetts Amherst: Rafael Coelho Lopes de Sa, Jessica Westclark, Eduardo Bach, Michael Zink, Carlos Ruben Dell'Aquila, and Verena Martinez Outschoorn.

Northeastern University: Miriam Leeser and Suranga Handagala.

## Acknowledgment

This work makes use of OCT, funded by the National Science Foundation under grants 1925464, 1925658, 2130907, 2130891, 2319962, and 2332423, and support from the Department of Energy award DE-SC0010004. We thank AMD for their generous donations.