Streaming Nuclear Physics at Scale

Jonathan Zarling, Vardan Gyurjyan, David Lawrence
Data and Computational Science
Jefferson Lab
Newport News, Virginia, USA
{jzarling, gurjyan, davidl}@jlab.org

David Abbott, Benjamin Raydo
Fast Electronics
Jefferson Lab
Newport News, Virginia, USA
{abbottd, braydo}@jlab.org

Derek Howard, Yatish Kumar, Stacey Sheldon
Energy Sciences Network
Lawrence Berkeley National Laboratory
Berkeley, California, USA
{dhoward, yak, stac}@es.net

***Abstract*—Nuclear physics experiments have long challenged the limits of high-rate and low-latency logic. Presently, streaming readout can move decision making from fixed detector electronics to software, allowing raw detector data to be processed continuously as they are produced. We highlight the state of Jefferson Lab and ESnet collaborative work with example workflows. First, existing data from the CLAS12 detector were replayed across ESnet through the EJFAT network load balancer to the Perlmutter supercomputer at NERSC, reaching 100 Gb/s as jobs dynamically scaled the workflow. Second, live data directly from the CLAS12 electromagnetic calorimeter were streamed, showing end-to-end capabilities and real-time cosmic ray identification. These demonstrations go beyond earlier transport benchmarks to physics-based streaming workflows, thus combining high-rate data transport with online scientific processing. We also discuss applying this approach to Jefferson Lab's development of streaming-readout modules for Positron Emission Tomography (PET) processing, enabling future whole-area, whole-body scanning otherwise considered impractical.**

***Index Terms*—streaming readout, high-performance computing, real-time data processing, wide-area networking, load balancing, nuclear physics**

I. Introduction

Traditional data acquisition systems use dedicated FPGA-based hardware to decide when continuous measurements coming from an experiment should be saved as physics events. While this has proven successful at significantly reducing the raw data rate, it is constrained by the limitations of fast electronics. This limits both the complexity and flexibility of the logic when making a decision to save or discard data coming from nuclear experiments. Streaming readout instead sends detector data continuously to software, where event identification, filtering, monitoring, and reconstruction can be performed using scalable computing resources [1].

This shift adds complexity to networking and compute steps along the experimental data path, placing high reliability requirements on both. Earlier work demonstrated that high-rate detector data could be transported over wide-area networks and distributed efficiently among remote compute resources [2], [3]. Here we take the next step toward complete scientific workflows: processing CLAS12 data at 100 Gb/s on a remote supercomputer and live identification of cosmic ray signals directly from a continuous detector stream.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics under Contract No. 89243126CSC000213

II. High-Throughput Reconstruction

Triggered CLAS12 data were replayed from Jefferson Lab across ESnet through EJFAT to NERSC's Perlmutter supercomputer. EJFAT distributed incoming data among processing jobs as they became available. The workflow sustained 100 Gb/s while using up to 40 Perlmutter nodes comprising more than 10,000 CPU cores [3].

This not only demonstrates network throughput. It also connected one scientific data source directly to a large and distant computing system and allowed processing capacity to change while data continued to flow. EJFAT can distribute data among workers that join or leave the processing pool, allowing networking and computing resources to scale independently [2], [3]. This demonstrated 100 Gb/s rate is therefore an end-to-end result rather than probing a limit of the architecture.

III. Real-Time Event Identification

A second demonstration addressed hardware integration into the workflow, while also finding scientifically useful signals directly within a continuous detector stream. The software identified time windows containing patterns consistent with cosmic rays, converted the readout into detector coordinates, and filled live occupancy and monitoring histograms as the stream was processed. In this workflow, candidate events were discovered in software from the continuously sampled detector data rather than supplied by a hardware trigger.

Although recorded data were replayed for this demonstration, the processing was performed online as the stream arrived. Integrating the same workflow with live detector data is the next step toward a complete streaming-readout path.

IV. Streaming Whole-Body PET

Jefferson Lab is also developing modular streaming readout for PET using scintillators, silicon photomultipliers, front-end electronics, and FPGA-based readout [4]. Software-based PET coincidence processing has been demonstrated previously [5]; the opportunity here is to combine modular detector readout with modern high-bandwidth networking and scalable computing.

This approach offers the capability to move coincidence finding, calibration, monitoring, and later analysis into a common software processing layer whose capacity can grow with detector size and data rate. Such scalability becomes increasingly important as PET systems expand in channel count and detector coverage. We therefore propose this solution to PET readout as an emerging application of the streaming architecture rather than a completed full-system demonstration.

V. Outlook

These demonstrations move streaming readout from transport-oriented proof of concept toward physics-facing workflows. The high-rate CLAS12 test shows that detector data can be delivered to dynamically scalable remote HPC resources, while the calorimeter demonstration shows that particle signals can be identified and monitored directly from a continuous data stream.

Together, they illustrate a model in which high-speed networking and remote computing are not simply used after data collection, but participate directly in the experimental workflow. We expect to scale these tests further to 300-400 Gb/s, which approaches the limits of detector readout hardware. This approach offers a path toward increasingly data-intensive nuclear-physics experiments and scalable medical applications such as a streaming-readout PET.

References

[1] F. Ameli *et al.*, "Streaming readout for next generation electron scattering experiments," *Eur. Phys. J. Plus*, vol. 137, no. 8, Art. no. 958, 2022, doi: 10.1140/epjp/s13360-022-03146-z.

[2] M. Goodrich, C. Timmer, V. Gyurjyan, D. Lawrence, G. Heyes, Y. Kumar, and S. Sheldon, "ESnet/JLab FPGA accelerated transport," *IEEE Trans. Nucl. Sci.*, vol. 70, no. 6, pp. 1096–1101, Jun. 2023, doi: 10.1109/TNS.2023.3243871.

[3] I. Baldin *et al.*, "Low latency, high bandwidth streaming of experimental data with EJFAT," arXiv:2510.12597, Oct. 2025. [Online]. Available: https://arxiv.org/abs/2510.12597

[4] W. Xi, J. McKisson, B. Raydo, and A. G. Weisenberger, "Streaming readout PET architecture and development," in *2023 IEEE Nuclear Science Symposium, Medical Imaging Conference and International Symposium on Room-Temperature Semiconductor Detectors (NSS MIC RTSD)*, Vancouver, BC, Canada, Nov. 2023, pp. 1–1.

[5] B. Goldschmidt, C. W. Lerche, T. Solf, A. Salomon, F. Kiessling, and V. Schulz, "Towards software-based real-time singles and coincidence processing of digital PET detector raw data," *IEEE Trans. Nucl. Sci.*, vol. 60, no. 3, pp. 1550–1559, Jun. 2013, doi: 10.1109/TNS.2013.2252193.