

Dec 2 – 4, 2024
University of Tokyo
Asia/Tokyo timezone

XII edition of Streaming Readout Workshop (Summary)



TAKU GUNJI

QUARK-NUCLEAR SCIENCE INSTITUTE

CENTER FOR NUCLEAR STUDY, THE UNIVERSITY OF TOKYO



Series of SRO workshop

- ▶ This workshop was initiated and led by **EIC R&D Streaming Readout Consortium eRD23** and has been focusing on the development of streaming readout technologies for the Electron-Ion Collider (EIC).
- ▶ XII edition of the SRO workshop brings together **DAQ specialist and experimentalist** from all over the world, to discuss the learning experience from existing streaming DAQ system and collaborate on future Streaming DAQ system at many facilities and experiments and **in particularly the EIC**.
- ▶ This XII is the first time to be held in Asia.

SRO I online, Jan. 2017

SRO II at MIT, Jan. 2018

SRO III at JLab ([link](#)), Dec. 2018

SRO IV at Camogli ([link](#)), May, 2019

SRO V at BNL ([link](#)), Nov. 2019

SRO VI online ([link](#)), May, 2020

SRO VII online ([link](#)), Nov, 2020

SRO VIII online ([link](#)), Apr, 2021

SRO IX online ([link](#)), Dec, 2021

SRO X at JLab ([link](#)), May, 2022

SRO XI at Hawaii/Satellite of APS/JPS ([link](#)), Nov-Dec, 2023

SRO XII at Tokyo ([link](#)), Dec, 2024

Streaming readout

3

- ▶ **Streaming readout is the new paradigm of the DAQ in HEP and NP community.**
 - ▶ Need to cope with high-rate interactions and large data volume
 - ▶ Needs to record complex events, which are hard to be implemented in the hardware triggers
 - ▶ Thanks to new developments in computing and network technologies
- ▶ The system will be very complex to handle continuous readout from detectors, continuous data processing in real-time (reconstruction, calibration, physics analysis), and orchestration of large systems and large number of processes.
- ▶ **This needs strong collaborations between DAQ experts, software experts, computing experts, and physics analyzers!!**



Streaming readout at ePIC

- ▶ Streaming readout is the base for ePIC.
- ▶ Compute-Detector Integration
 - ▶ Acceptance of 100% events and final state particles
 - ▶ Rapid turnaround of 2-3 weeks from raw data stream to physics analyses with full calibrations

ePIC streaming computing model

Integration is challenging due to complex composition of SRO systems: RDO, DAM, switch/load balancer, Computing, HW accelerators, software framework for timeframe building, data transport and workflow-based distributed processing, Orchestration (PanDA + Rucio)

Streaming readout for continuous data flow of the full detector information.

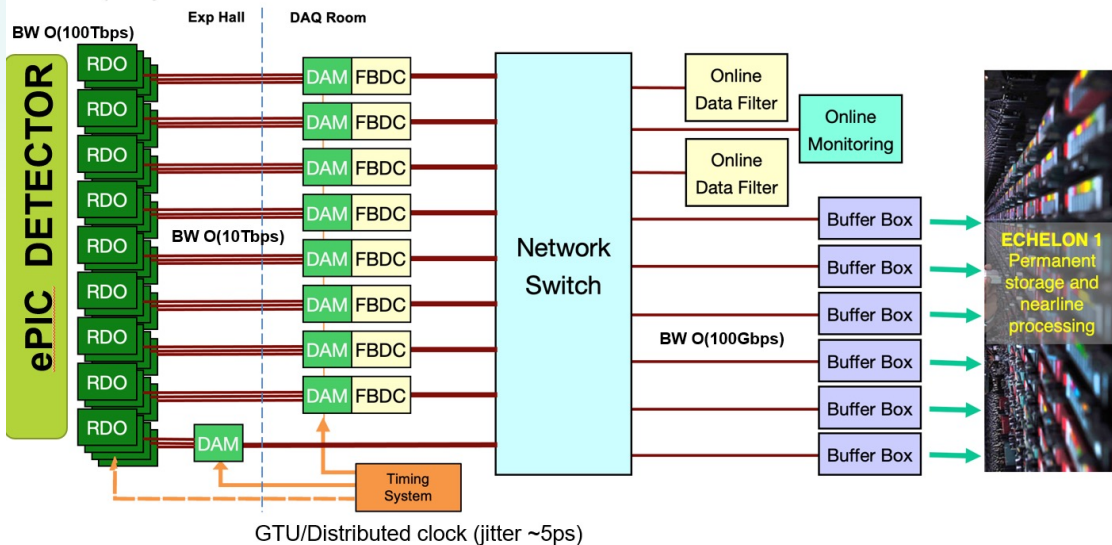
AI for autonomous alignment, calibration, and validation for rapid processing.

Heterogeneous computing for acceleration (CPU, GPU).

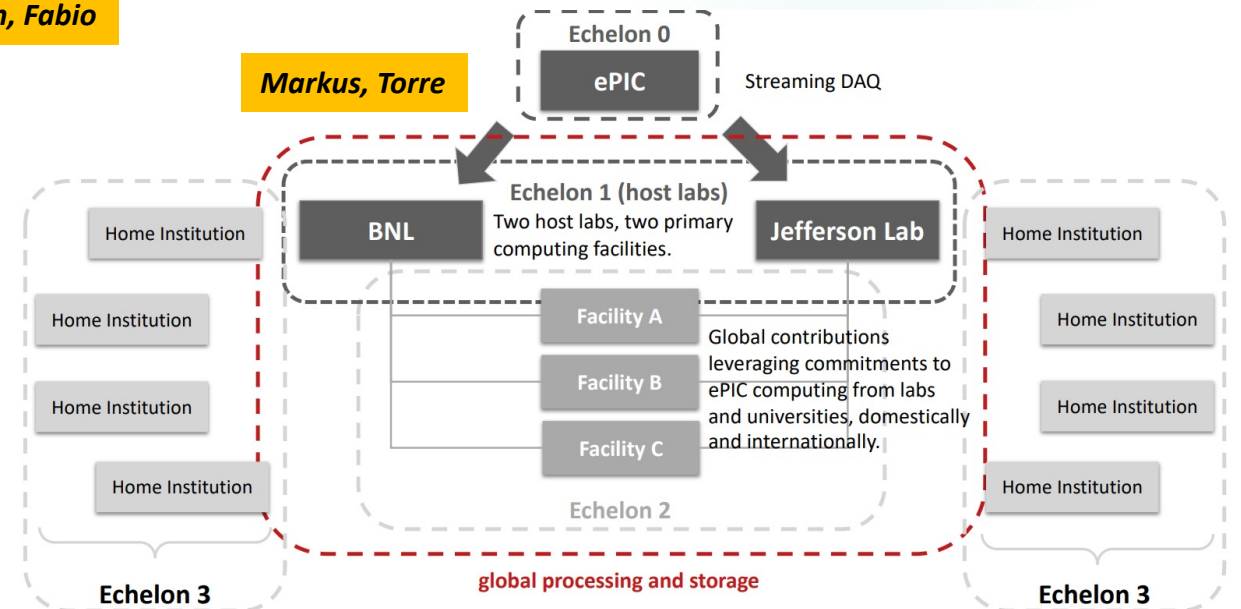
Bunch Crossing ~ 10.2 ns/98.5 MHz
Interaction Rate ~ 2 us/500 kHz
Low occupancy

Jeff, William, Fernando

Luca, Cristian, Fabio



Markus, Torre

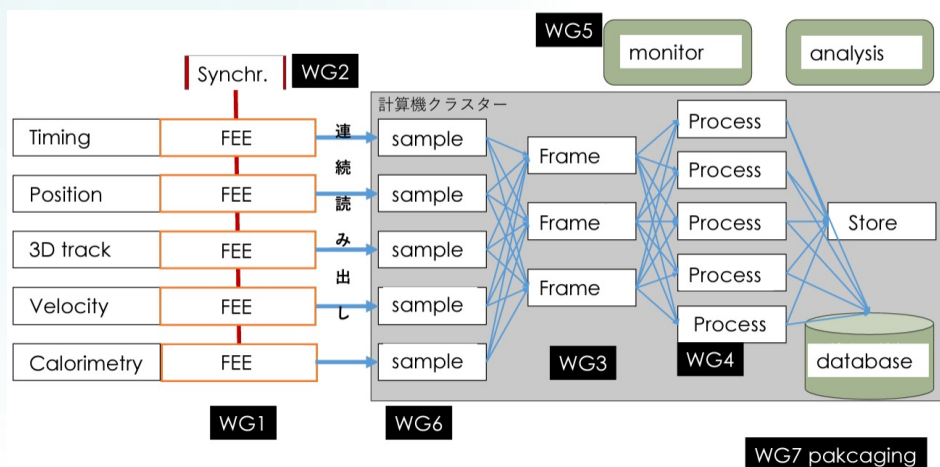


SPADI-Alliance in Japan

- ▶ Streaming readout are common needs in the entire nuclear physics of Japan
 - ▶ Low energy physics at RIBF in RIKEN and at RCNP in Osaka University
 - ▶ Hadron physics at the hadron-hall facility in J-PARC and RARiS at Tohoku University
 - ▶ High-energy QCD physics at the facilities outside Japan (LHC-ALICE, EIC-ePIC, FAIR-CBM)
- ▶ SPADI-Alliance was established in 2022.

Shinsuke, Kotaro, Tomonori and posters

 - ▶ >140 researchers and 24 institutes from different experiments and different facilities
 - ▶ 7 working groups (FEE, Timing, framework, real-time processing, UI, Computing ...)



<p>WG1 Frontend Electronics</p> <p>Streaming type Charge ASD board Voltage ASD board WF Digitizer board Control Firmware dev.</p>	<p>WG2 Clock synch. / Data Transfer</p> <p>General Clock Synch. High throughput Intra-board transfer</p>	<p>WG3 Acquisition software framework (NestDAQ + ...)</p> <p>Streaming type FairMQ-based Scalable DAQ Sampling, Time frame build, Event build, Monitoring... Format</p>	<p>WG4 Event processing</p> <p>Acceleration using GPU/FPGA Zero suppression Calibration, Clustering, Tracking, PID,</p>	T. Gunji as convener
<p>WG5 User Interface</p> <p>Control, Monitor, Configure, Trial with SlowDash</p>	<p>WG6 Computing infrastr.</p> <p>High throughput Large volume Flow and Archive Power consumption Interconnect Networking</p>	<p>WG7 Packaging</p> <p>Standalone system Popularization Standardization Market research User feedback</p>	<p>Analysis</p> <p>Trial with Artemis</p>	

XII edition of SRO workshop

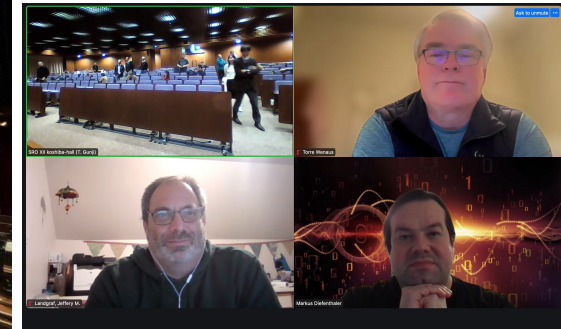
- ▶ XII edition of SRO workshop addresses:
 - ▶ Streaming DAQ and experiences at many facilities
 - ▶ Real-time calibration and data processing in SRO and heterogeneous computing
 - ▶ Application of AI/ML technologies
 - ▶ ASICs, FECs, Data Aggregation, new challenges for SRO
 - ▶ Establishment of work plans for the future SRO system
- ▶ At this time, we had a lot of contributions from **AI4EIC** to discuss the development and implementation of AI/ML based technologies in the SRO. **This was successful and will be very likely considered in future editions.**
- ▶ T. Gunji (chair), M. Battaglieri, J. Bernauer, A. Camsonne, M. Diefenthaler, C. Fanelli, T. Horn, T. Hachiya, Y. Goto, Y. Sekiguchi, S. Ota (SPADI-A), H. Baba (SPADI-A)



XII edition of SRO workshop

7

- ▶ <https://indico.bnl.gov/event/24286>
- ▶ The number of registrants = 107
 - ▶ In-person attendance = 45
(31 from Japan, 14 outside)
 - ▶ Zoom attendance = 62
(16 from Japan, 46 outside)
- ▶ 38 presentations + 7 posters
 - ▶ 23 talks online, 15 talks in-person



Streaming readout experiences

▶ JLab: CODA, JANA2, EJFAT, ERSAP

EJFAT: ESnet-JLab FPGA Accelerated Transport

ERSAP: Environment for Real-time Streaming, Acquisition, and Processing Framework

▶ BNL: RCDAQ

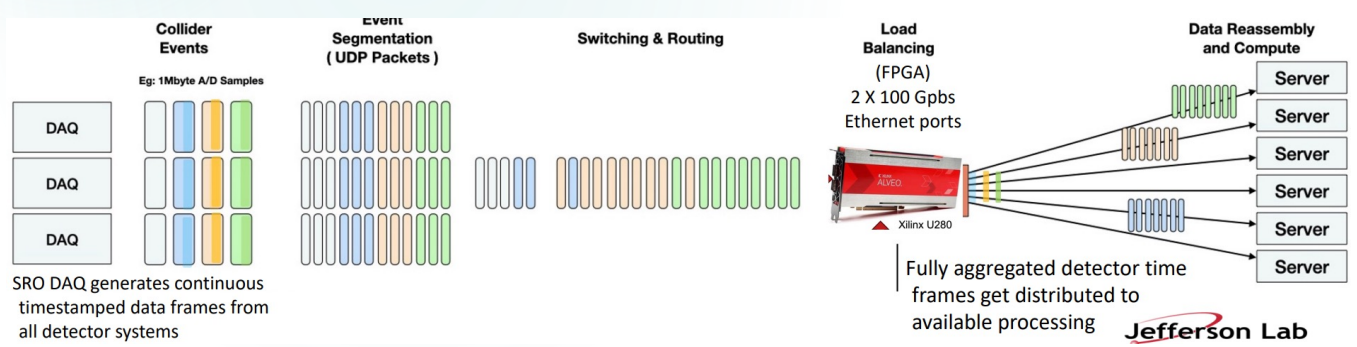
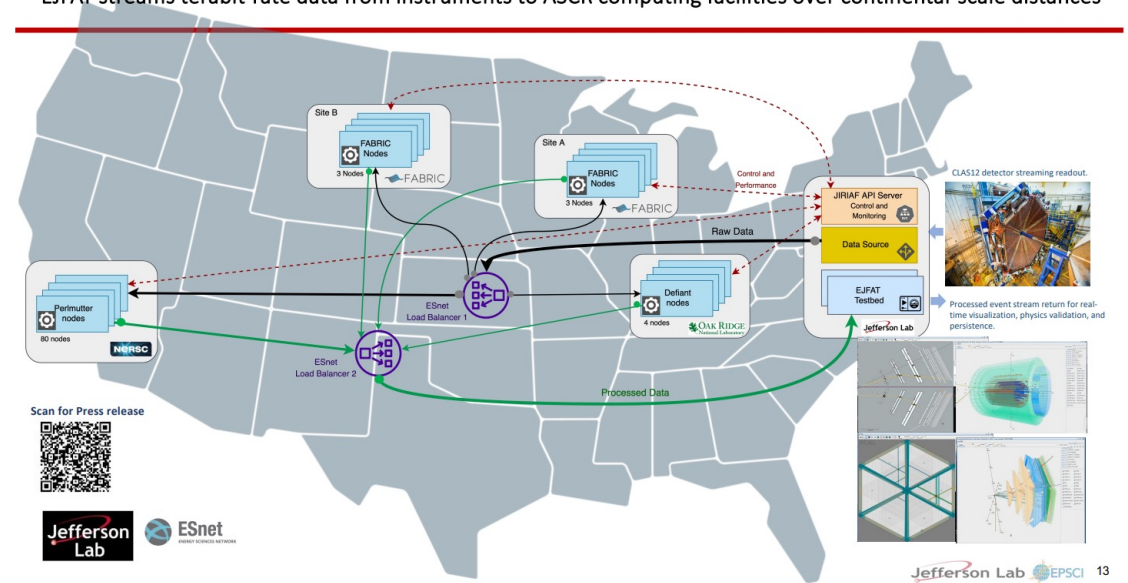
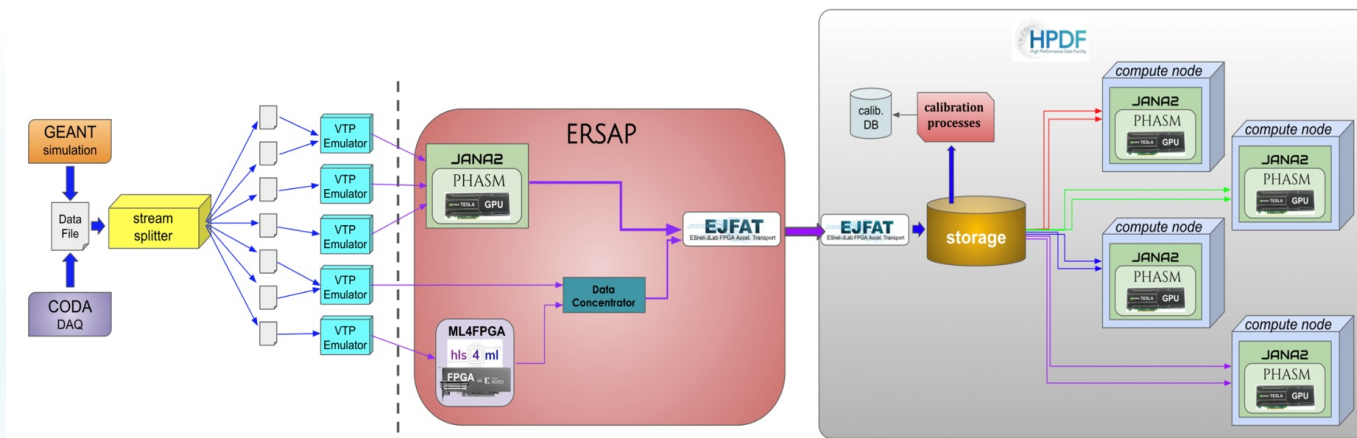
▶ Japan: nestDAQ, artemis (offline-online SW)

David, Jeng-Yuan, Carl, Hanjie, Nathan

Martin, Charles, Genki

Shinsuke, Tomonori, Kotaro

EJFAT streams terabit-rate data from instruments to ASCR computing facilities over continental-scale distances



Scan for Press release

Jefferson Lab ESnet

<https://portal.fabric-testbed.net/about/about-fabric>

FABRIC

International infrastructure for cutting-edge research in the areas of networking, storage, distributed computing, machine learning, science applications, etc.

The world map highlights the global reach of the FABRIC infrastructure, with nodes located in Seattle, Salt Lake City, StarLight, New York, Los Angeles, Dallas, Atlanta, FIU, University of Tokyo, University of Amsterdam, University of Bristol, and CERN.

Streaming readout experiences

▶ JLab: CODA, JANA2, EJFAT, ERSAP

EJFAT: ESnet-JLab FPGA Accelerated Transport
 ERSAP: Environment for Real-time Streaming, Acquisition, and Processing Framework

▶ BNL: RCDAQ

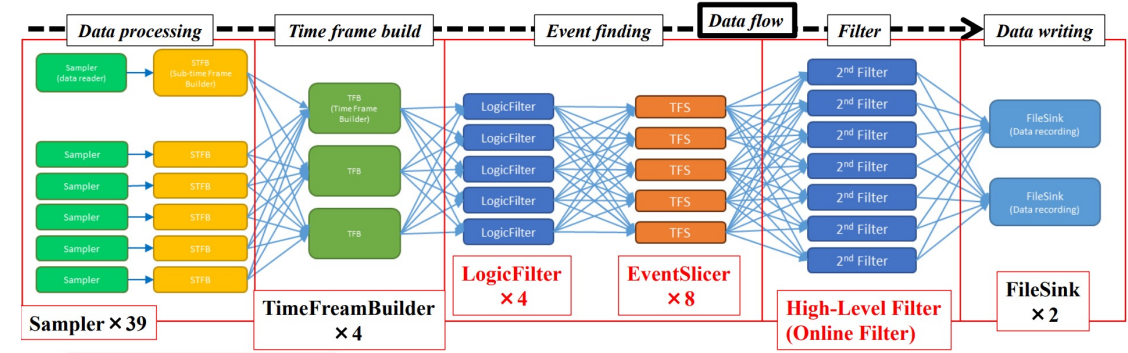
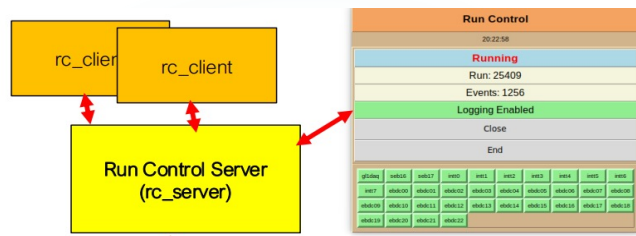
David, Jeng-Yuan, Carl, Hanjie, Nathan

Martin, Charles, Genki

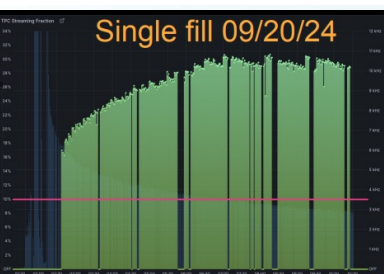
▶ Japan: nestDAQ, artemis (offline-online SW)

Shinsuke, Tomonori, Kotaro

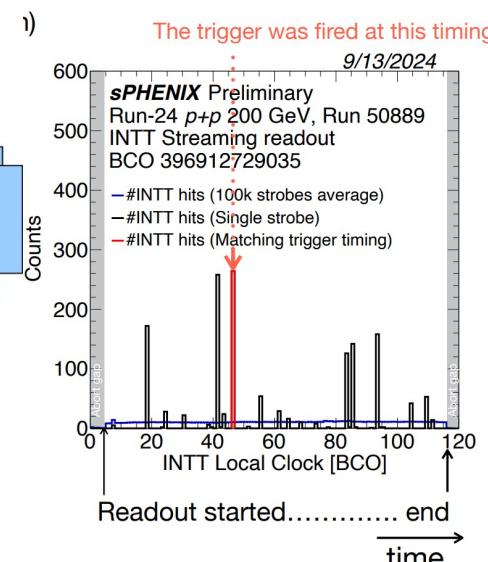
NestDAQ configuration



This year, a full complement consisted of 61 RCDAQ instances



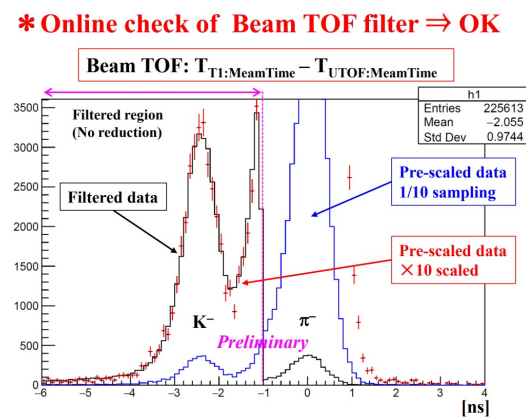
Streaming fraction increases as collision rate decreases within a fill -> works around disk bandwidth limitations



Test bench photos

Standard nuclear physics experiment scale (Much compact!)

- >4,000 channel readout w/ various detectors
- Data flow ~180 MB/s (~1.4 Gbps) (averaged in 4.2 sec. cycle)
- Beam extraction (2.0 sec.): ~260 MB/s (~2.1 Gbps)



Real-time processing with AI/ML in SRO

10

▶ Real-time reconstruction

- ▶ Autonomous selection of physics events (Cameron Dean)
- ▶ Data filtering on GPU and FPGA for the dRICH detector (Luca Pontisso, Cristian Rossi)
- ▶ Fast ML on FPGA for Particle Identification and Tracking (Sergey Furletov)
- ▶ Deep(er)RICH - Deep Reconstruction of Imaging Cherenkov Detectors (James Giroux)
- ▶ Data processing acceleration for the Belle II experiment (Qi-Dong Zhou)

▶ Data compression

- ▶ Real-Time data reduction with Artificial Intelligence for SRO (Fabio Rossi)
- ▶ Neural Compression for sPHENIX sparse TPC Data (Yihui Ray Ren)
- ▶ Decision tree autoencoder on FPGA (Tae Min Hong)

▶ Experiment control

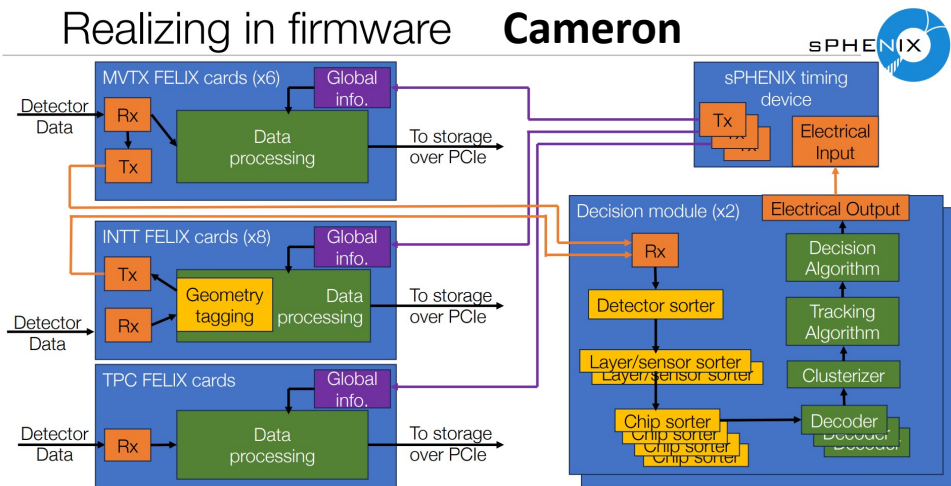
- ▶ AI for Experimental Control (Torri Jeske)



Real-time processing with AI/ML in SRO

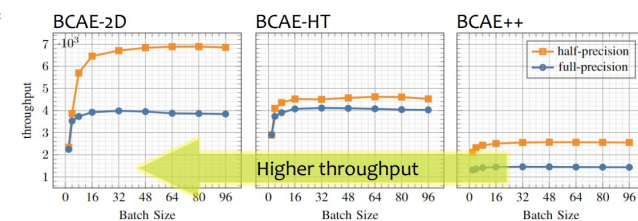
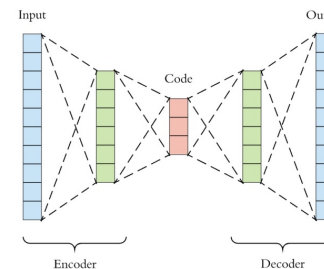
Stream sPHENIX-MVTX and INTT to FPGAs and determine if HF event is present through topology

Realizing in firmware Cameron



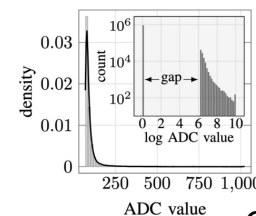
Neural Auto-Encoder

- A typical Auto-Encoder uses an Encoder network to compress the data into "code"; and a Decoder network to reconstruct the original input.



- The voxel distribution:
 - long-tailed (skewed)
 - sparse (many zero values)
 - zero-suppressed (discontinued)
 - 10-bit integer (saturated)

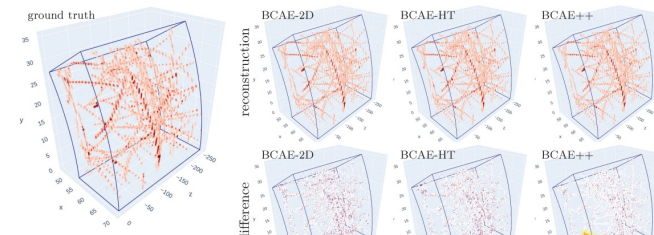
Yihui Ray



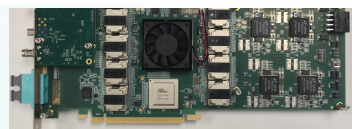
Very Challenging for a regular auto-encoder!

Brookhaven National Laboratory SRO XII, University of Tokyo, Dec. 2-4, 2024. Presenter: Yihui Ren (BNL)

Compression ~ 30



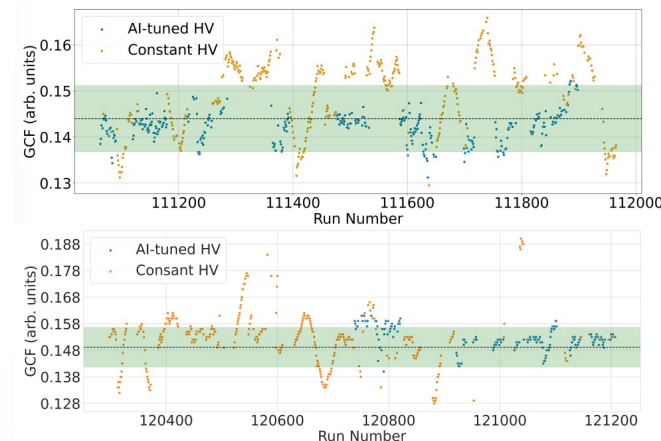
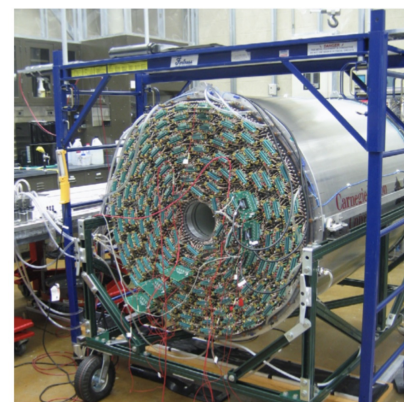
GNN-based tracking and decision algorithm ported in FELIX-712



Torri deploy an AI system to autonomously adjust detector controls during data acquisition

Jefferson Lab EPSC Carnegie Mellon University

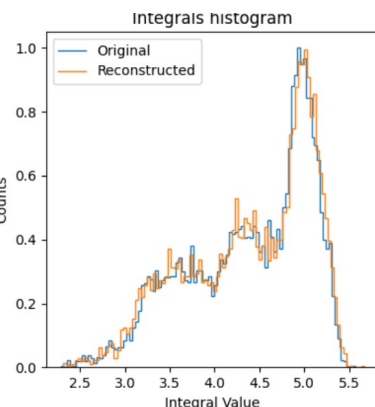
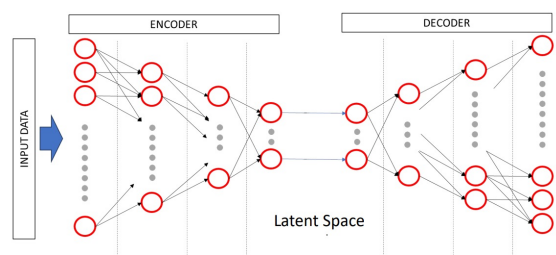
GlueX Central Drift Chamber



Fabio

Data reduction algorithm: Autoencoder

Machine Learning Algorithm
Dimensionality reduction
Unsupervised learning
Artificial Neural Network
Composed of two function:
- encoding
- decoding
FULLY CONNECTED AUTOENCODER WITH DENSE LAYER



ASICs for SRO

- ▶ ePIC ASICs and ASICs from OMEGA and INFN
- ▶ SRO ASICs and systems from CAEN, NALU, ALPAH CORE
- ▶ Streaming capable ASICs with AI/ML features

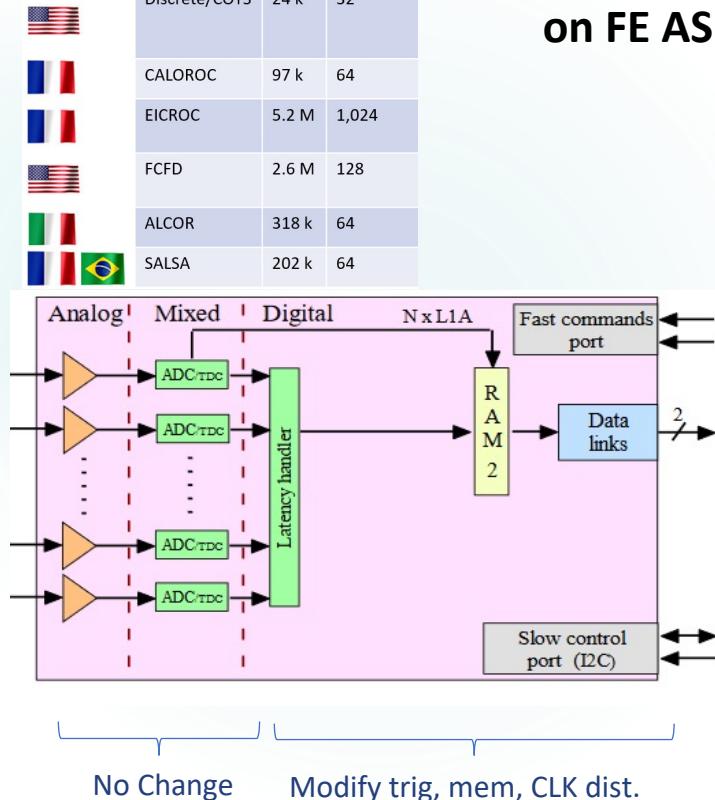
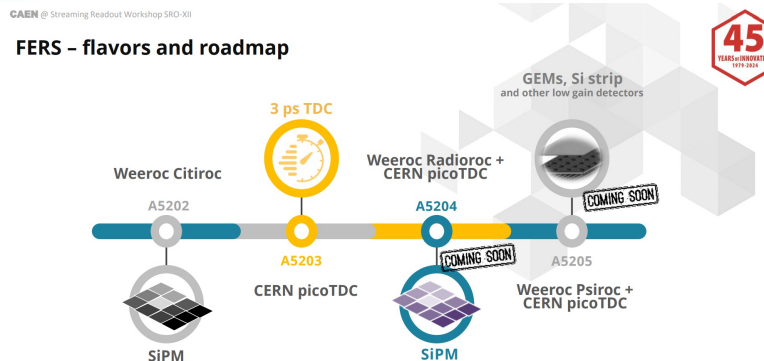
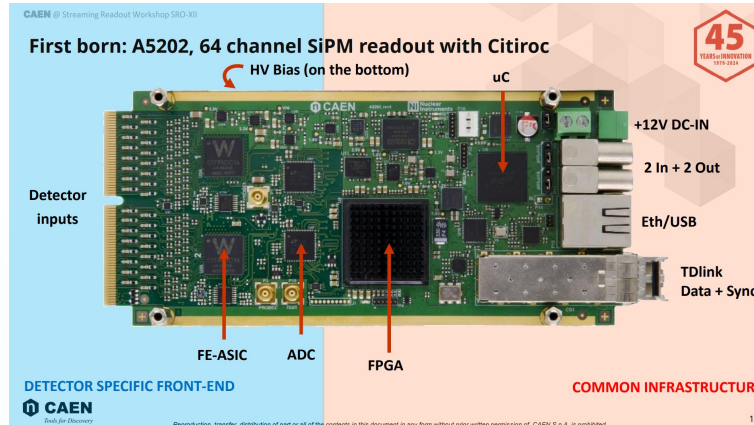
Fernando, Christophe, Angelo

Giovanni, Luca, Esko

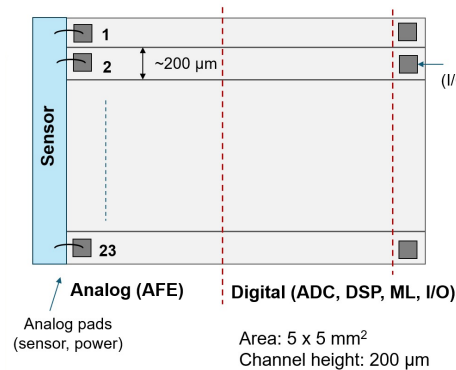
Soumyajit

	#Ch	#Ch/Unit
Discrete/COTS	24 k	32
CALOROC	97 k	64
EICROC	5.2 M	1,024
FCFD	2.6 M	128
ALCOR	318 k	64
SALSA	202 k	64

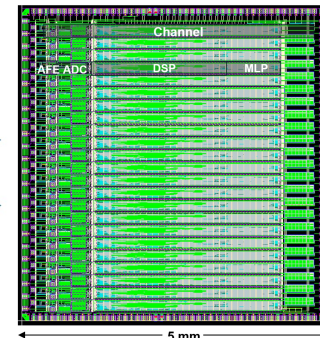
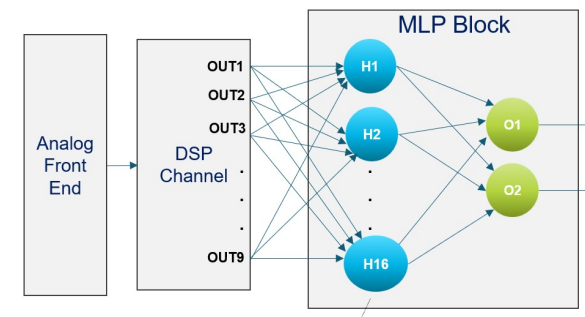
CAEN-FERS: all-in-one readout system based on FE ASICs



Overview of the proposed readout ASIC



- Simplified AFE design (charge-sensitive amp + anti-aliasing filter)
- Per-channel 12-bit ADC generates inputs for on-chip DSP and ML
 - Fully-differential, includes calibration
- Per-channel DSP:
 - Programmable shaping filter (FIR)
 - Waveform alignment / snipping
 - Baseline removal
- Per-channel machine learning (ML):
 - Artificial neural network (ANN): multi-layer perceptron (MLP) with programmable weights
 - Performs classification or regression
- Additional blocks:
 - Low-speed I2C programming and testability interface
 - Output serializer



Data aggregation

FELIX card (BNL)

Hao

Status of FLX-182

- Hardware functionalities are fully validated
 - Total 28 (24+4) links @ 25 Gb/s are available for data transmission
 - PCIe Gen4 performance
 - 2x Gen4x8 endpoints, theoretical payload bandwidth 120.47Gb/s for each endpoint
 - 2 x8 endpoints: 2x 113.2 Gb/s, 94% of theoretical bandwidth
 - 1 x8 endpoint: 1x 118 Gb/s, 97.9% of theoretical bandwidth
- Different flavors of FELIX firmware have been implemented, and functionality demonstrated
- 50+ FLX-182 cards have been produced for different HEP and NP experiments
 - ATLAS Phase-II Upgrade, ALICE at CERN, and CERN DRD7 hardware platform
 - ePIC at EIC
 - sPHENIX at RHIC
 - CBM/RE21 at FAIR

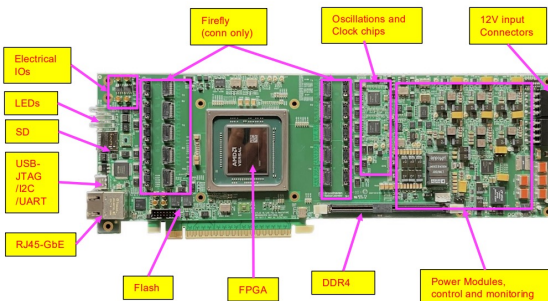


Picture of FLX-182 cards

FLX-155

Main features of FLX-155

- AMD/Xilinx Versal Premium FPGA: XCVP1552-2MSEVSA3340
- PCIe Gen4 x16 / PCIe Gen5 2x8
- 56 FireFly optical links
 - Compatible with various options
 - Default configuration for ATLAS
 - 48 data links up to 25 Gb/s
 - 4 links for LTI
 - 4 links for 100GbE
- Electrical IOs
- 1 DDR4 Mini-UDIMM
- USB-JTAG/USB-UART
- SD3.0/QSPI
- GbE
- White Rabbit



Picture of assembled FLX-155



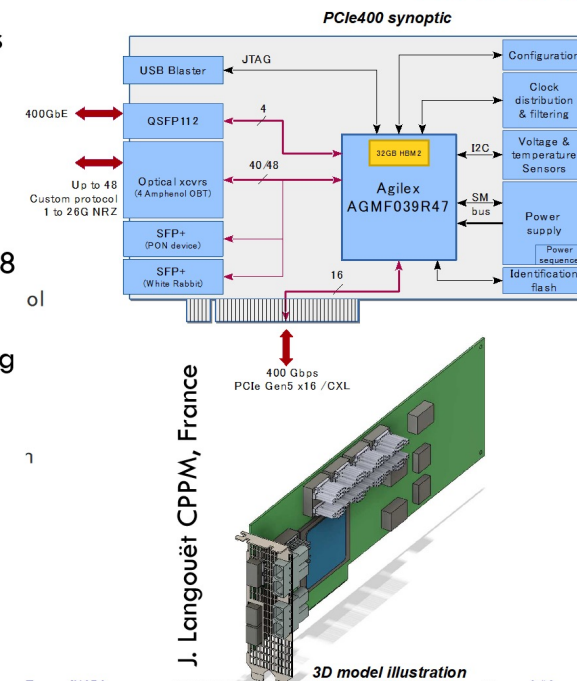
FLX-155 revision will be available for user in 2025 fall,

PCIe400

Niko

The PCIe400 - an attempt to fit all

- Gen5 PCIe interface – 500 Gbit/s raw / ~ 450 Gbit/s effective
- Alternatively 400 Gbit/s Ethernet raw / ~ 350 Gbit/s effective
- 32 GB of HBM2e memory / 410 GB/ peak memory bandwidth / 368 GB/s effective
- Up to 48 SERDES / links - excluding PCIe
- One of the most advanced FPGAs today



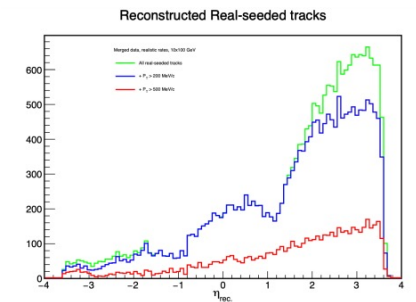
Next steps for the PCIe400

- First version in production → next testing, testing, testing and debugging → followed by revised (final) version
- Collection of numbers for intermediate upgrade of readout during LHC LS3 (2026 – 2029) and production of cards until 2027

For the Future

▶ ePIC collaboration meeting

- ▶ "ePIC Data: From Detector Readout to Analysis", parallel session at the ePIC Collaboration meeting, January 20–24, 2025, Frascati, Rome.
- ▶ Discussions among detector teams, software & computing teams, DAQ teams, and PWGs are super-demanding and critical for the future developments



▶ Next edition of SRO workshop in 2025

- ▶ Catania in Italy (M. Battaglieri & M. Bondì)

