SRO Prospects And Preparation at FRIB

Giordano Cerizza
cerizza@frib.msu.edu

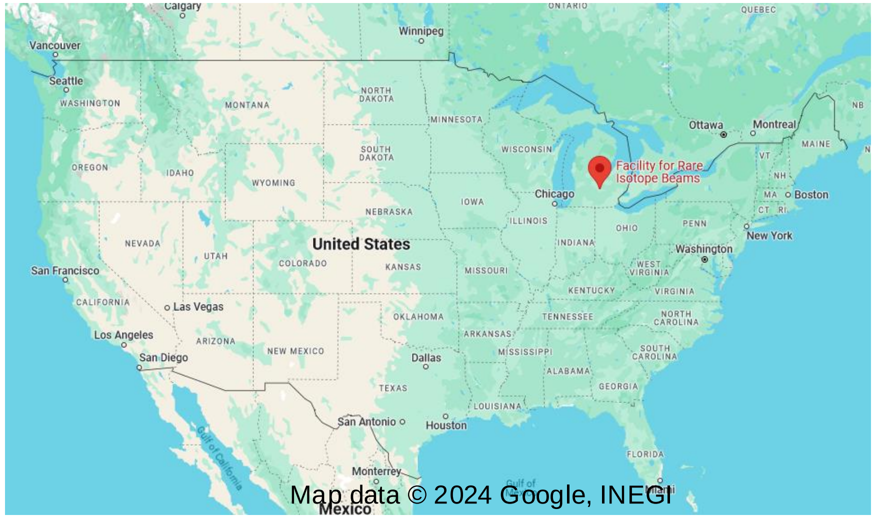
Streaming Readout Workshop XII

2-4 December 2024

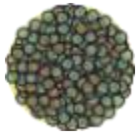
This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics and used resources of the Facility for Rare Isotope Beams (FRIB) Operations, which is a DOE Office of Science User Facility under Award Number DE-SC0023633.

The Facility for Rare Isotope Beams (FRIB)

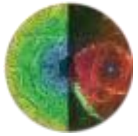
- FRIB is a scientific user facility funded by the US Department of Energy Office of Science (DOE-SC), Michigan State University (MSU), and the State of Michigan
  - Open to researchers from around the world
  - User Organization: 1800 members (125 colleges and universities, 13 national labs, 53 countries)
  - Experimental program began in May 2022
- Key feature of FRIB: high-power LINAC, 400 kW at 200 MeV/u for U-238
- Beams from oxygen to uranium
- FRIB also provides stopped and low-energy re-accelerated beams (< 12 MeV/u)
- **Access to atomic nuclei across the nuclear chart at a range of energies**



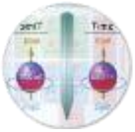
The FRIB Experimental Program



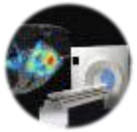
- Nuclear structure
  - How does subatomic matter arrange itself and how does it evolve?
  - What combinations of neutrons and protons form bound atomic nuclei?



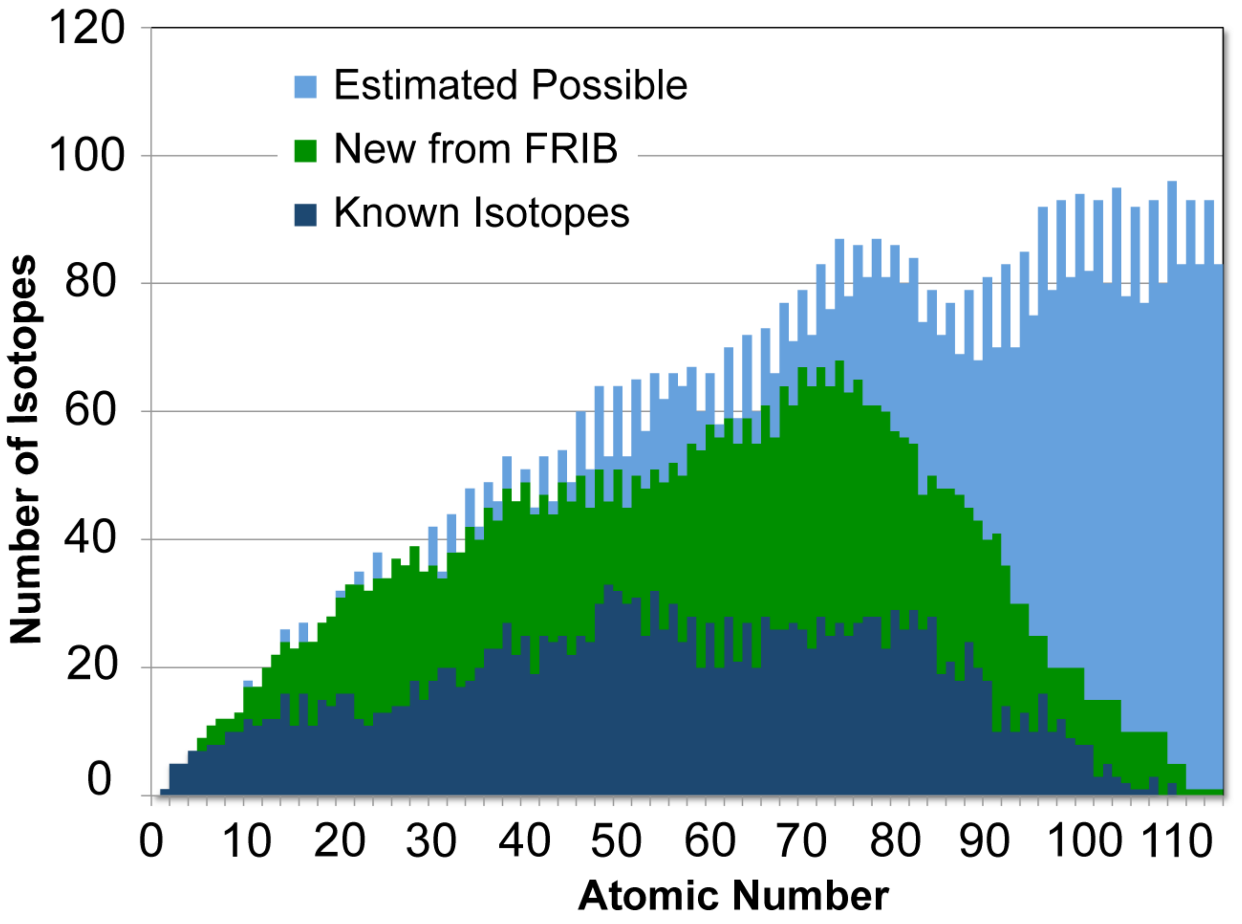
- Nuclear astrophysics
  - How are the chemical elements created?
  - What is the nature of matter at extreme temperatures and densities?



- Fundamental symmetries
  - Why is there more matter than antimatter?
  - Are neutrinos their own antiparticles?



- Societal applications and benefits
  - Medicine, energy, material sciences, environment, workforce development, etc.



Erler et al., Nature **486** (2012) 509-512.

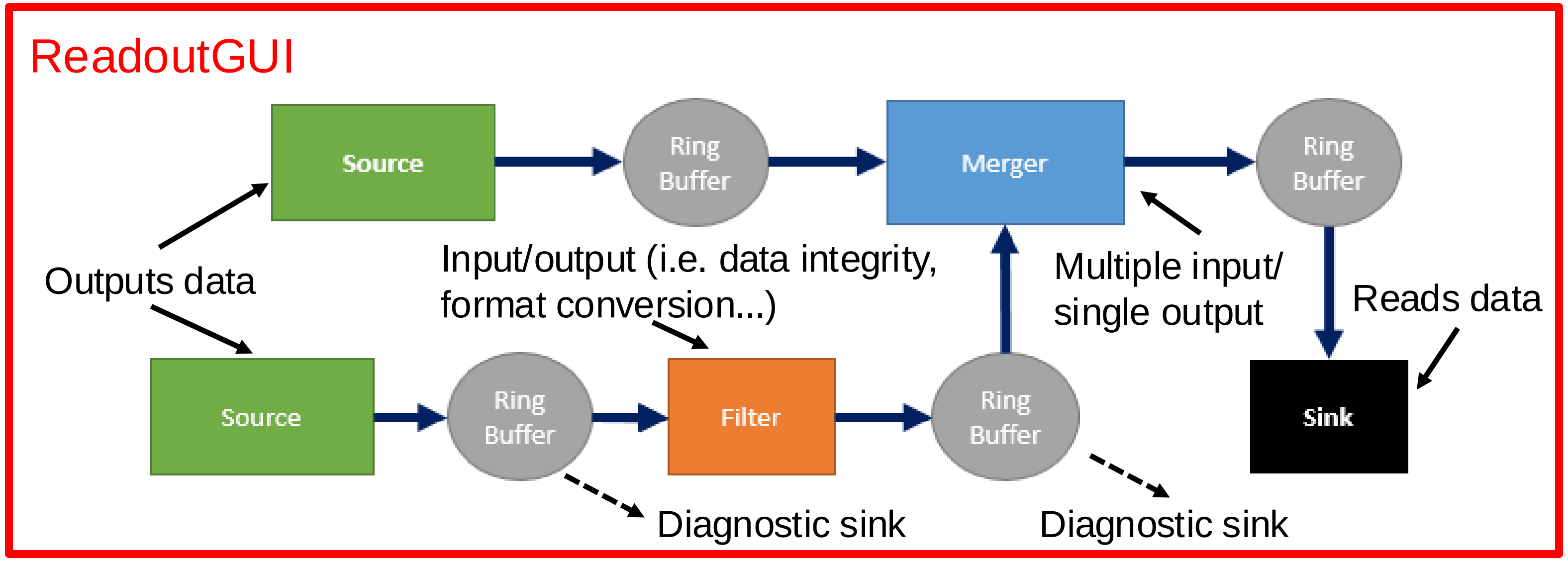
Status of FRIB in 2024

- New instrument developments
  - GRETA (at FRIB in 2025)
  - HRS (expected commissioning phase around 2030)
  - SECAR (in commissioning phase)
  - New detectors for higher rates, better resolution…
- R&D and integration of new data acquisition (DAQ) systems
  - Adoption of new DAQ systems (i.e. CERN SRS for tracking detectors with VMM3a ASICS) – success!
  - DOE R&D project for digitizers with WR capabilities – ongoing!
- Integration of DAQ systems with computing resources to reduce time to enable discoveries and improve decision-making during experiments with AI/ML

How about SRO at FRIB?

FRIBDAQ Readout Architecture

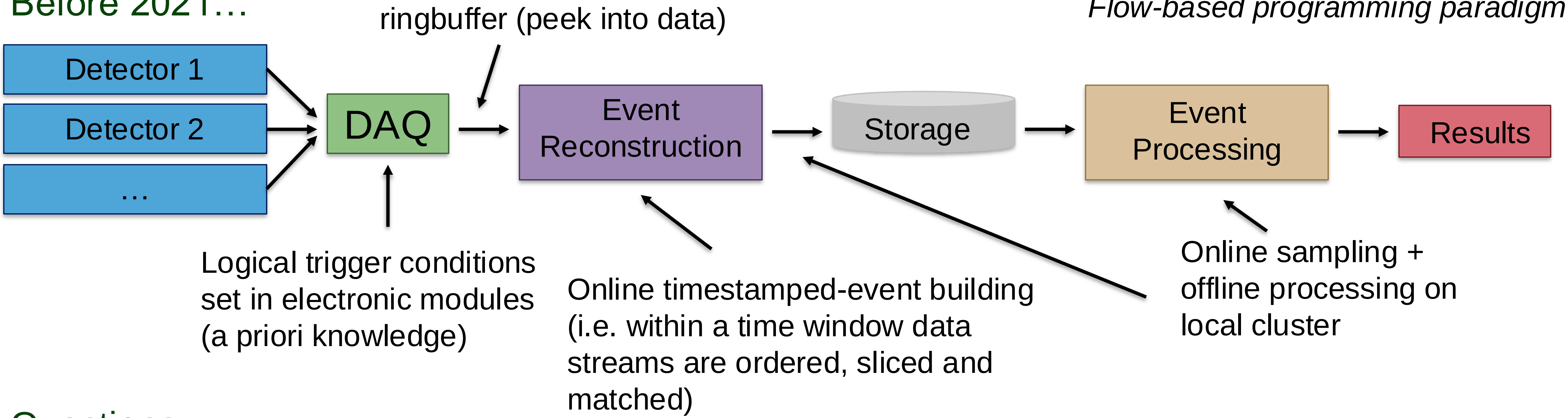
- Software suite (C++, Tcl – being replaced by PyQT5, Python) that provides a flexible and extensible framework for handling the data flow produced by nuclear physics experiments
- Support for a wide variety of electronics: VME, CAMAC, CAEN and XIA digitizers, GET, SRS
- Manage the data stream by breaking it into smaller pieces
- ReadoutGUI is the “conductor” of the system
  - Pipeline management, data recording, consolidation of data to output, …



FRIB Approach to SRO

- Before 2021…

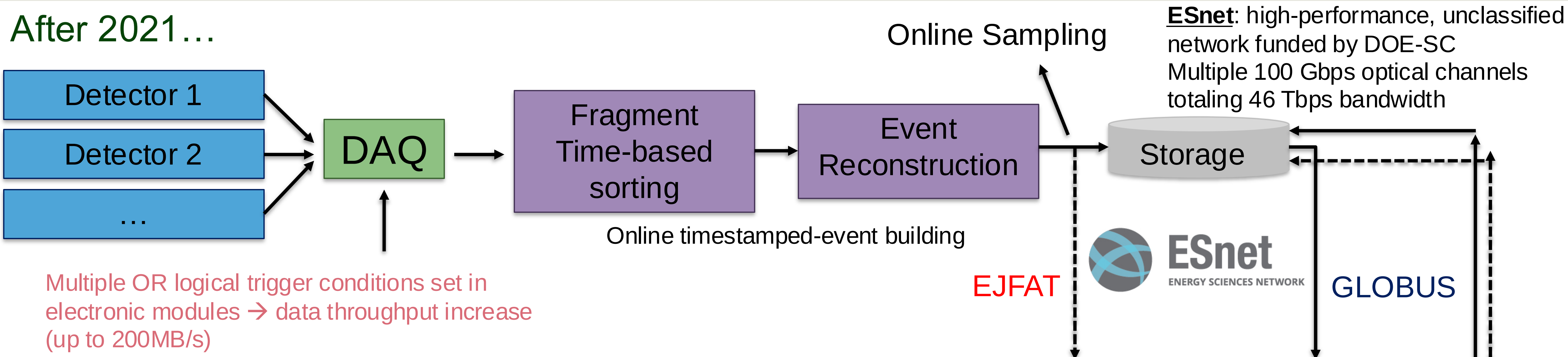
Flow-based programming paradigm

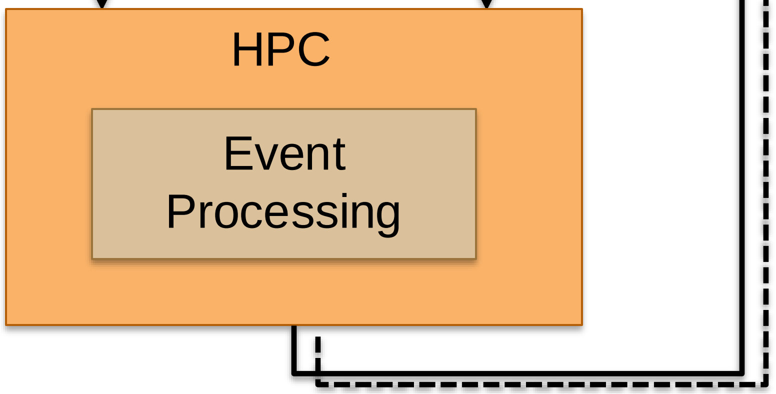


- Questions:
  - Can SRO electronics and global timing synchronization improve data taking and data analysis efficiencies with very rare isotope beams? (FYI 1h beam is ~20k$)
  - Where should we put efforts considering the limited resources in hardware R&D?
  - How do we manage the continuous increase in data rates?

FRIB Approach to SRO

- After 2021…

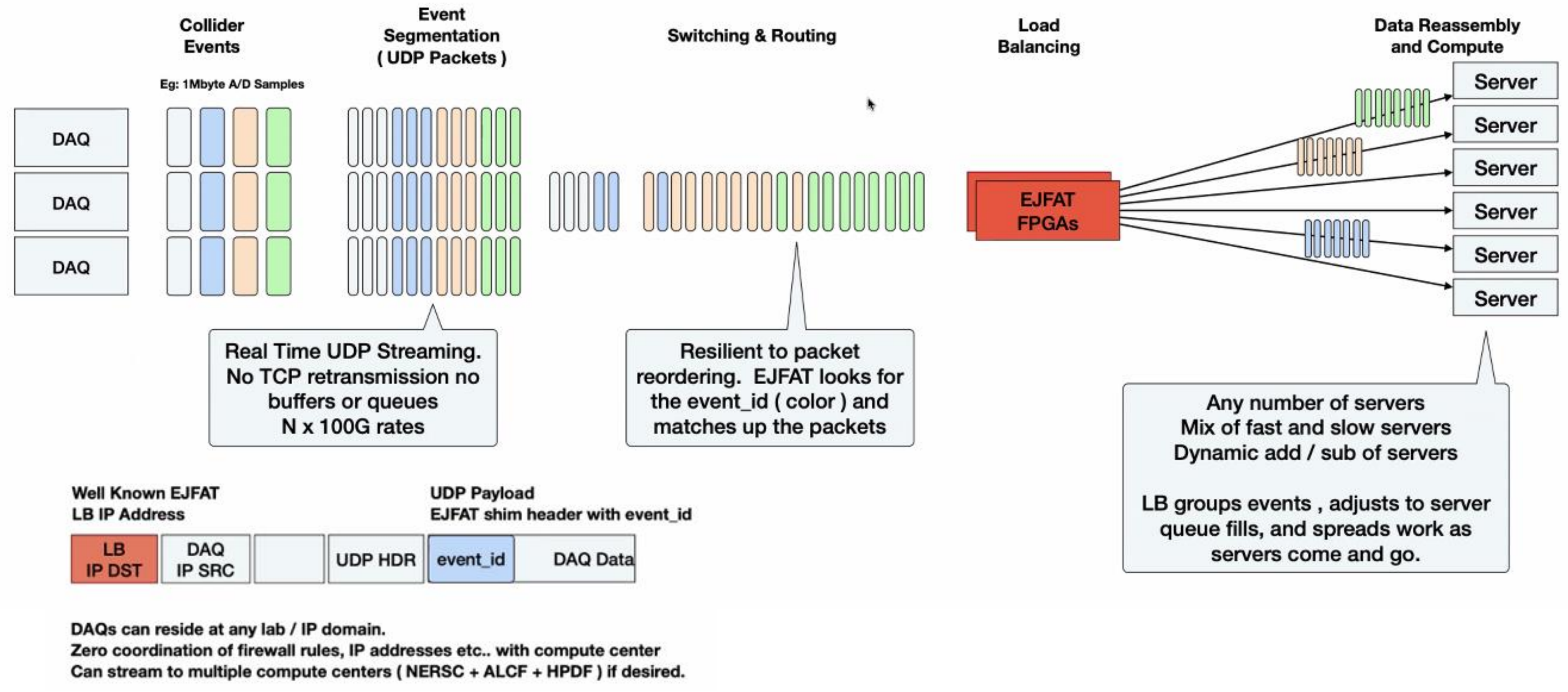




- How did we achieve this?
  - Our software DAQ and Analysis libraries runs in containers
    - » Minimum requirement for portability
  - Parallel event processing
  - Taking advantage of DOE resources (new paradigm - local cluster ONLY for readout out and fast diagnostic)
- Did it work? Spoiler: yes ☺

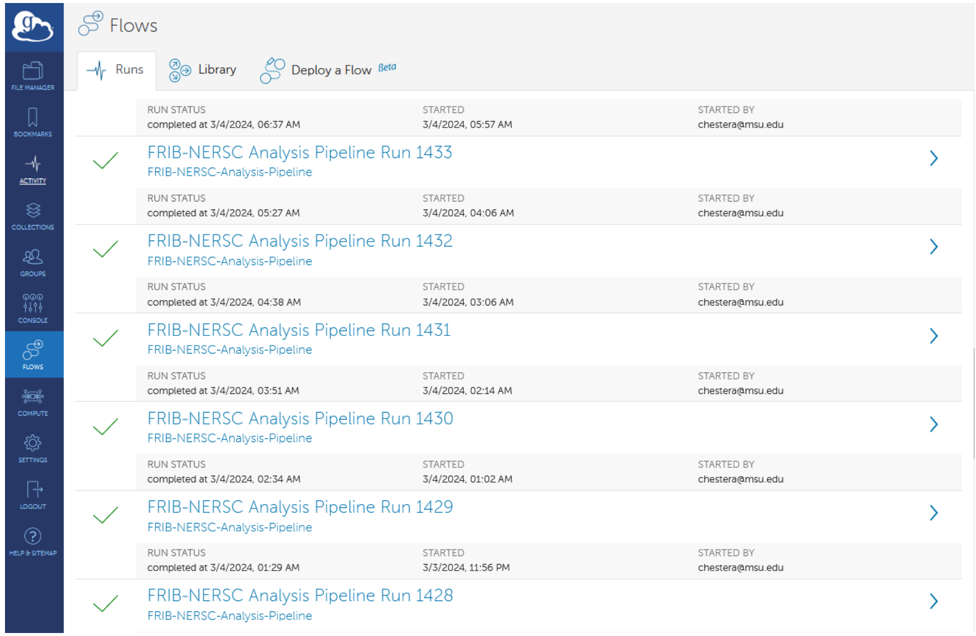
EJFAT

- ESnet-JLAB FPGA Accelerated Transport (EJFAT) Load balancer



Automated Data Processing With Globus Flows

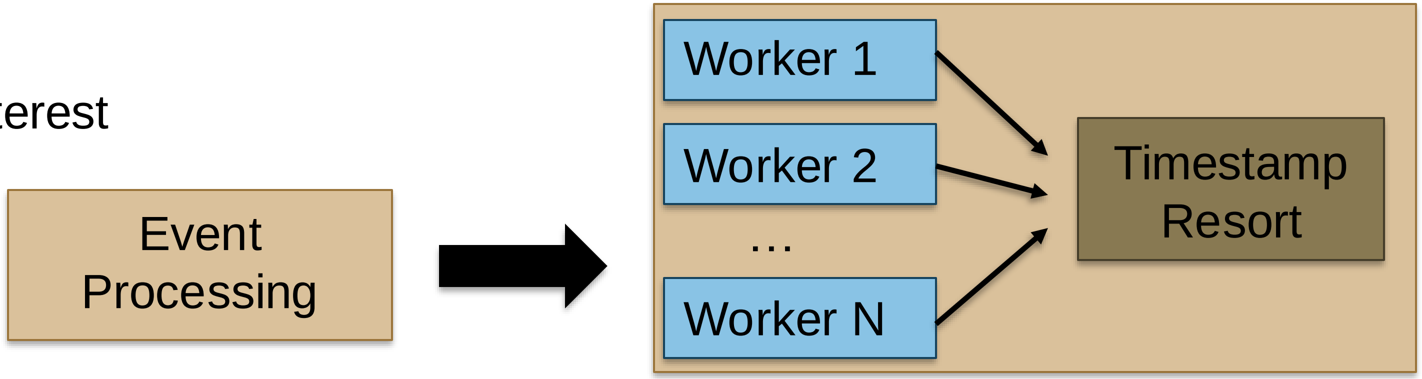
- “Secure, managed automation of complex workflows at scale”
  - Common tasks implemented via hosted “action providers” (e.g. data transfer)
  - Globus Compute provides a “function as a service” (FaaS) platform for remote execution of user code
  - Error handling
- Action providers and other operations can be assembled into a workflow
- Python-based SDKs for Globus and Globus Compute
  - Register applications, configure inputs, run, manage results
- Web-based monitoring
- Data processing workflow:
  1. Acquire data at FRIB
  2. Transfer raw data to NERSC via ESnet
  3. Process data remotely at NERSC
  4. Transfer processed data back to FRIB via ESnet



https://docs.globus.org/api/flows/

FRIBDAQ Upgrades for High-Data Rates

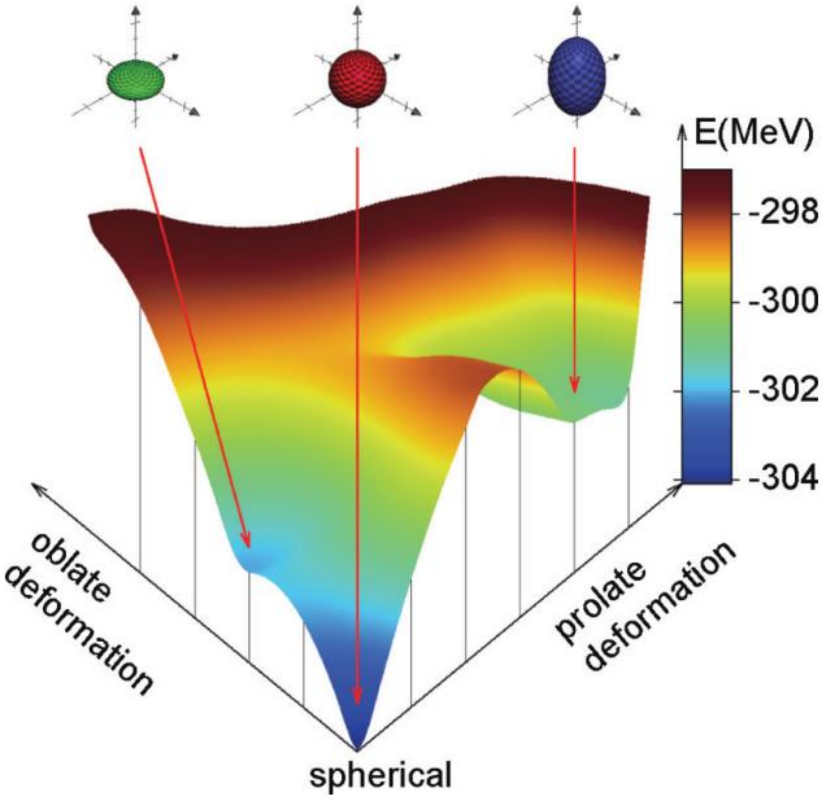
- The parallel Event Processing framework:
  - Based on several communication patterns (fan in, fan out, router-dealer, pipeline)
  - Has ability to run the code in safe-threaded or MPI modes
  - Has built-in frameworks for common operations that allow users to plug in sequential code chunks to do the work (i.e. loading user-built code into shared objects – fit/ML models)
  - Ability to edit events (iovec) where each worker contains:
    - » EventFilter (software trigger) – filtering background/bad events
    - » Appender – appending user code info
    - » Editor – reducing data volume
    - » Classifier – tagging and selecting events of interest



A Real Example: E0 Transition in Decay Spectroscopy

- $0^+ \rightarrow 0^+$ transitions are a probe of nuclear radii and shape deformation
- Tools to understand shape coexistence as ρ^2(E0) transition probabilities depend on underlying nuclear wave functions
- β^- decay is used to populate excited stats which may decay by E0 transition
  - Internal conversion (IC) or internal pair formation (IPF) are the main modes
  - The lifetime of the $0^+ \rightarrow 0^+$ transition provides information about the difference in the radius of the charge distribution of the two states
- To explain better I'll take the example of the β-decay $^{80}Ga \rightarrow {}^{80}Ge$

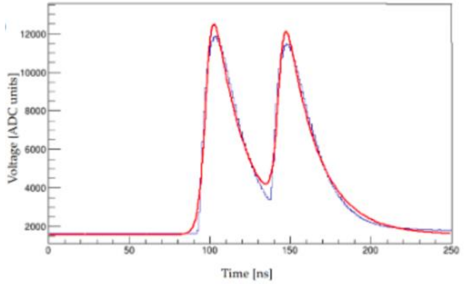
Deformation of 0^+ states in ^{68}Ni

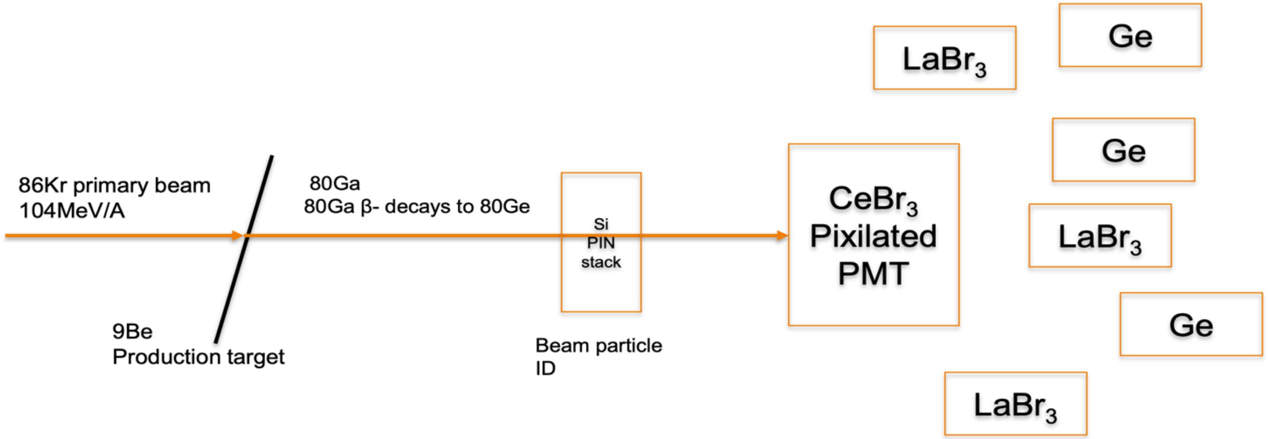


S. Suchyta et al. 2014

E0 Transition in Decay Spectroscopy: Setup

- ^{80}Ga decays to ^{80}Ge via β-decays
  - This decay is detected in the $CeBr_3$ detector and populates several energy level of ^{80}Ge
- Of interest are the decays that populate 0_2^+
  - This eventually de-excited to 0_1^+ emitting a γ-ray (detected in the $LaBr_3$) and a conversion electron (detected in the $CeBr_3$)
  - Similar de-excitations have half lives of 50ns
- How to measure it?
  - Digitize the pulses in the $CeBr_3$
  - Sum signal at 500MHz
  - Pixels at 250MHz
  - Trace lengths of a few microseconds





- IC/IPF have characteristic double and triple pulses with Δt→0 (very hard to distinguish)
- Data rate is dominated by ~200traces/event CeBr3 (~3kHz trigger rate, ~200MB/s)
- Event processing:
  - PIN detector to discard heavy ion
  - Analyze traces with analytic function to determine the number of traces
  - Find correlation between implantation events with decay events

Trace Analysis

- We approach the problem in two ways
  - Levenberg-Marquardt curve fitting (LMFit) on a) entire dataset b) double pulse candidates, predetermined by a binary classifier such as a SVM or random forest (M. Kaymak et al. 2019)
    - This method is computational complex (requires calculation of the Jacobian matrix, J^TJ and sometimes its inverse which is applied to each event separately) and fails for $\Delta t \rightarrow 0$
  - Deep learning model based on auto-encoders with an additional parameter estimation branch that combines PCA and parameter estimation (end-to-end) by using the gradient vector only

For both we use the same empirical model of the pulse as function of position, rise time, amplitude, decay rate + offsets

Work by postdoc Sadeghi B. - paper in writing

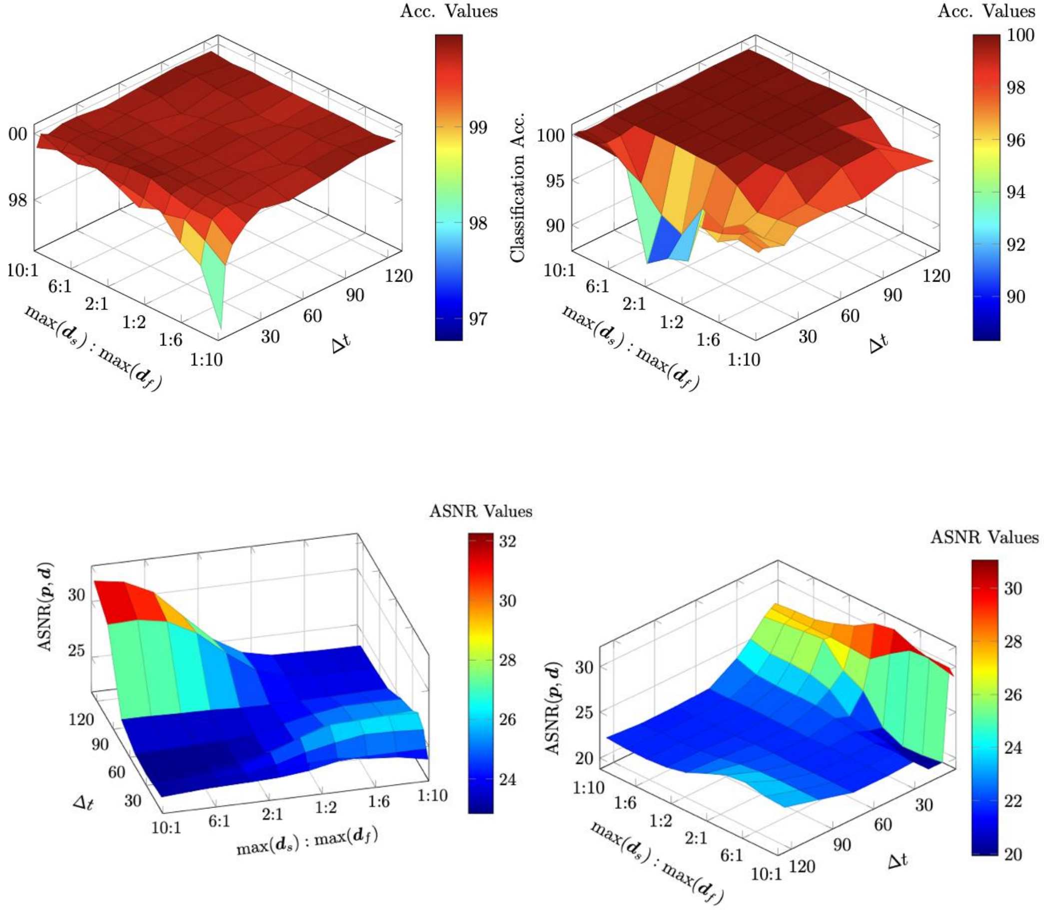
Trace Analysis: Results

	SAEFit	LMFit
$\mathrm{ASNR}(\boldsymbol{p}, \boldsymbol{s})$: single pulse fitting	23.53 dB	21.52 dB
$\mathrm{ASNR}(\boldsymbol{p}, \boldsymbol{d})$: double pulse fitting	24.70 dB	23.88 dB
$\mathrm{ASNR}(\boldsymbol{p}_f, \boldsymbol{d}_f)$: first part of $\boldsymbol{d}$ fitting	22.16 dB	13.76 dB
$\mathrm{ASNR}(\boldsymbol{p}_s, \boldsymbol{d}_s)$: second part of $\boldsymbol{d}$ fitting	22.02 dB	13.32 dB
$\mathrm{RMSE}(\Delta t)$: estimation error for $t_2 - t_1$	**2.16** sampling period	22.21 sampling period
$\mathrm{RMSE}(\max(\boldsymbol{p}_f))/(\max(\boldsymbol{d}))$	0.029	0.253
$\mathrm{RMSE}(\max(\boldsymbol{p}_s))/(\max(\boldsymbol{d}))$	0.034	0.257
Inference run-time	**5** seconds	8,128 seconds
Training run-time	7,200 seconds	187 seconds
Classification accuracy	**99.69%**	$\leq$ 96.78%

*traning set 245k traces

confusion matrix	predicted single		predicted double	
	SAEFit	LMFit	SAEFit	LMFit
actual single	**99.69%**	$\leq$ 96.75%	**0.31%**	$\geq$ 3.25%
actual double	**0.31%**	$\geq$ 3.20%	**99.69%**	$\leq$ 96.80%

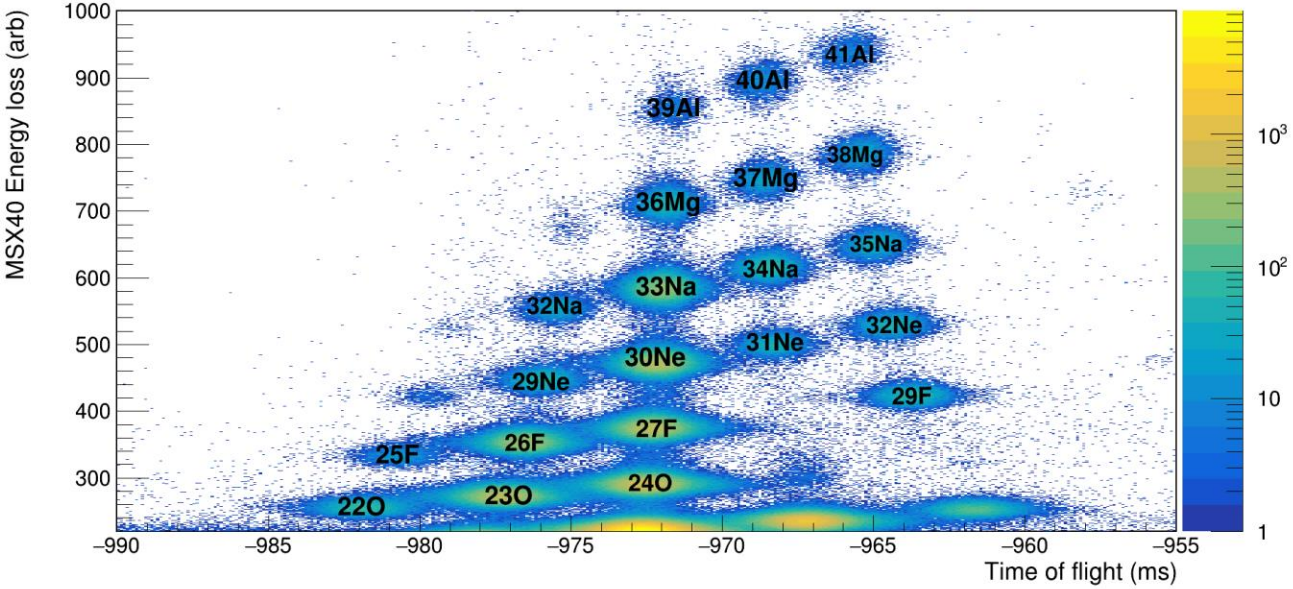
*results on 70k traces



Data Processing Workflow and Results

- Some results of first production test Spring24
  - 140 hours of experiment, 11.4TB raw data, 2.5TB processed data (80% reduction)
  - One hour of recorded data fully processed and back in ~10 minutes (not user-optimized code)
  - Average utilization of ~1500 CPU
  - Total of ~700 compute hr for full experiment

Particle identification plot generated in near real time from data processed at NERSC (courtesy of R. Lubna)



<u>Spring25 second production run test goals:</u>

- Test the ML classification on a real experiment – “online” event tagging
- Proxy data to multiple HPC facilities
- Possibly first test of EJFAT@FRIB → data transfer time from minutes to ms

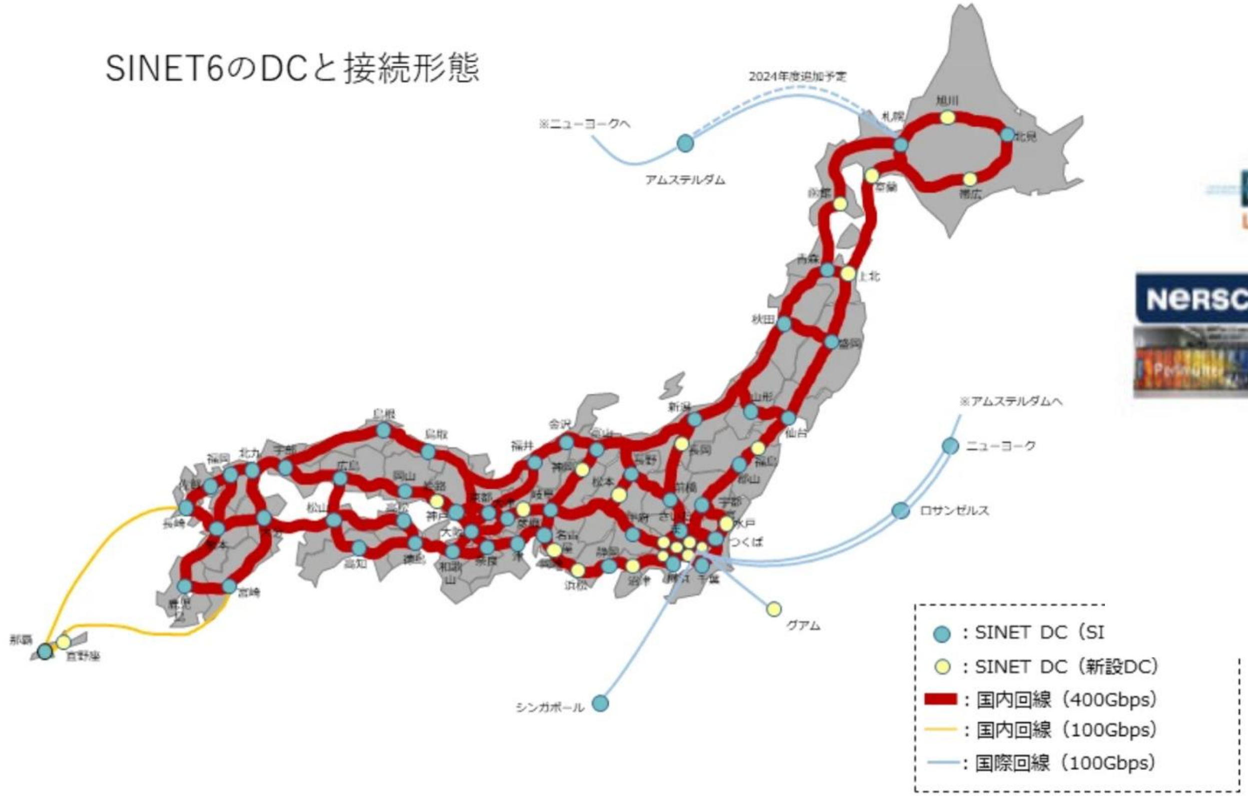
What Did We Learn?

- By having ESnet connection we can offload the heavy computational needs to much more performant clusters
  - Pros: less money spent at FRIB for storage, nodes (cost of ownership shrank)
  - Cons: obtaining resources is not a trivial thing (i.e. no real mechanism in place for it at DOE level)
- We can extract result faster than we can take data empowering the users in making decisions to adjust and optimize detector configurations in near-(on)line for very rare isotope beams
- Our DAQ software framework is growing in the “right” direction. Other key parts have been identified for upgrades (parallelization of event building with event correlation capabilities, heartbeat system to fast monitoring of the components)
- ML/AI is a key part of the FRIB future for beam optics optimizations, fragment separation, online decision making, and offline analysis
- We can aggregate data rate up to 200-300 MB/s on multiple VME/PCIe backplane but…

What Are We Missing?

- …it's becoming inefficient (i.e. expensive very fast) → looking at SRO
- We have limited resources on SRO hardware R&D. One of the identified solution was participating in the SPADI-A because of the matching timescales and similar interests/goals
- We haven't identified (yet) a solution for global timing synchronization for the experimental program (we aim at <30ps timing resolution for TOF measurements)
  - White Rabbit as "commercial solution" is under study
  - Active R&D with RIKEN for the CTS module (see Baba-san's poster)
  - Open to other options (MIKUMARI,…)
- Commercial and HEP solutions are often "ok" but far from being optimal for the needs of the (Low Energy) Nuclear Physics community → collaborations are becoming keys for the future our scientific programs

US-Japan Collaboration: Food for Thoughts



SINET: Science Information NETwork (サイネット)

国立情報学研究所(NII)

ESnet: high-performance, unclassified network funded by DOE-SC
Multiple 100 Gbps optical channels totaling 46 Tbps bandwidth

How to connect these two high-speed networks?

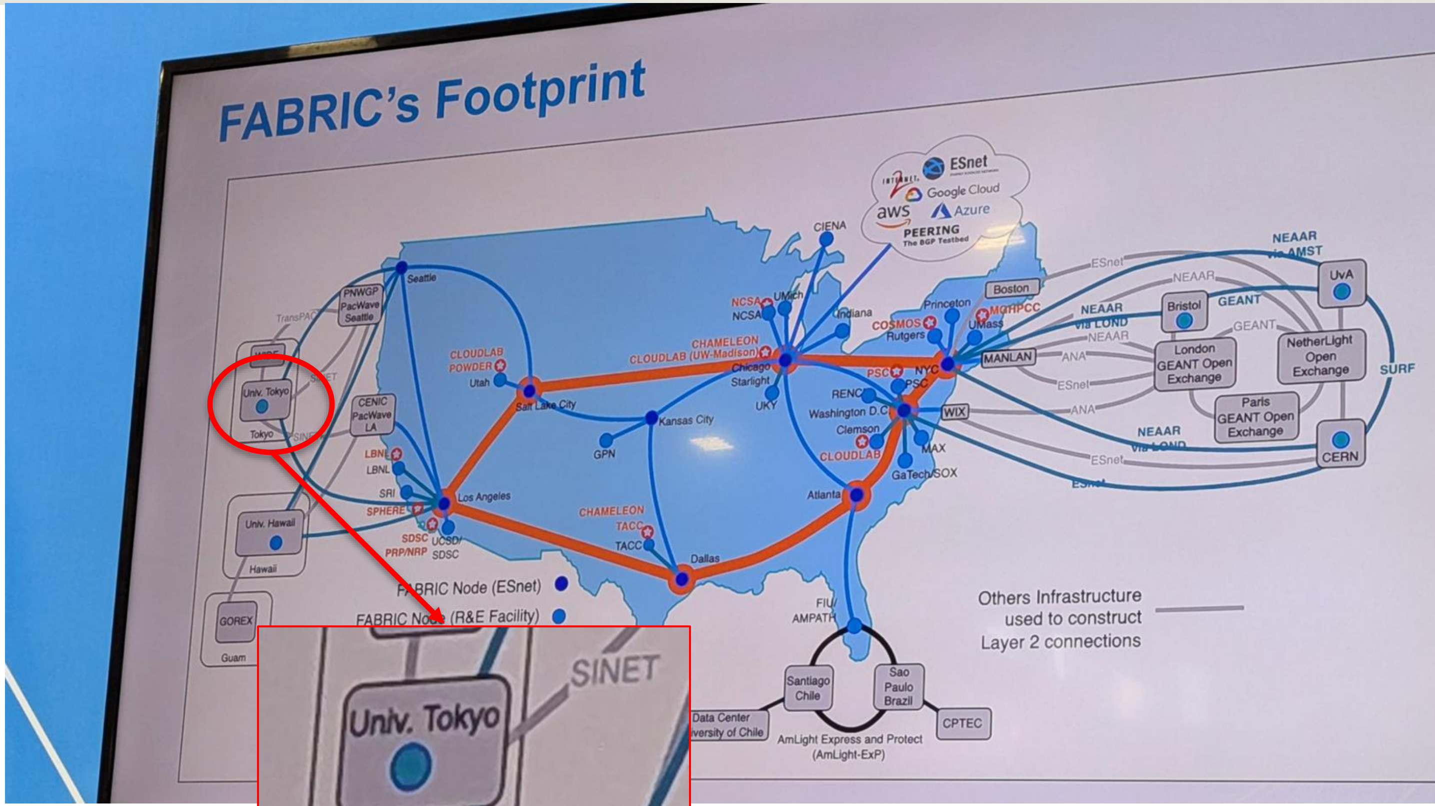
US-Japan Collaboration: Food for Thoughts

FABRIC

FABRIC (FABRIC is Adaptive Programmable Research Infrastructure for Computer Science and Science Applications) is an International infrastructure that enables cutting-edge experimentation and research at-scale in the areas of networking, cybersecurity, distributed computing, storage, virtual reality, 5G, machine learning, and science applications.

The FABRIC infrastructure is a distributed set of equipment at commercial collocation spaces, national labs and campuses.

CLAS12 at JLAB has been using FABRIC to support EJFAT



Pic from slide from SC24

fabric-testbed.net

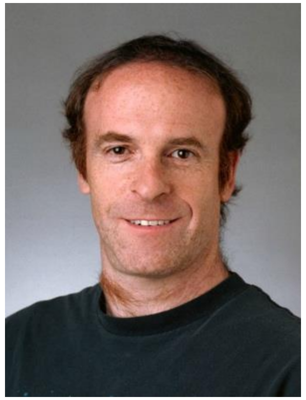
Summary

- FRIB is interested in collaborating with JLAB and SPADI-A for SRO R&D and solutions
- As of today, FRIB doesn't technically need SRO but there are use cases that could be a good testbed for the future (<3 years)
- The synergy between FRIB and DOE HPC facilities and projects (ESnet, EJFAT) is really paying off and it's moving forward
- ML/AI integration in the readout/analysis pipeline is becoming more important to optimize and speed up diagnostics and result extractions
- I want to thank my group for the amazing efforts

Aaron Chester

Genie Chang

Bashir Sadeghi

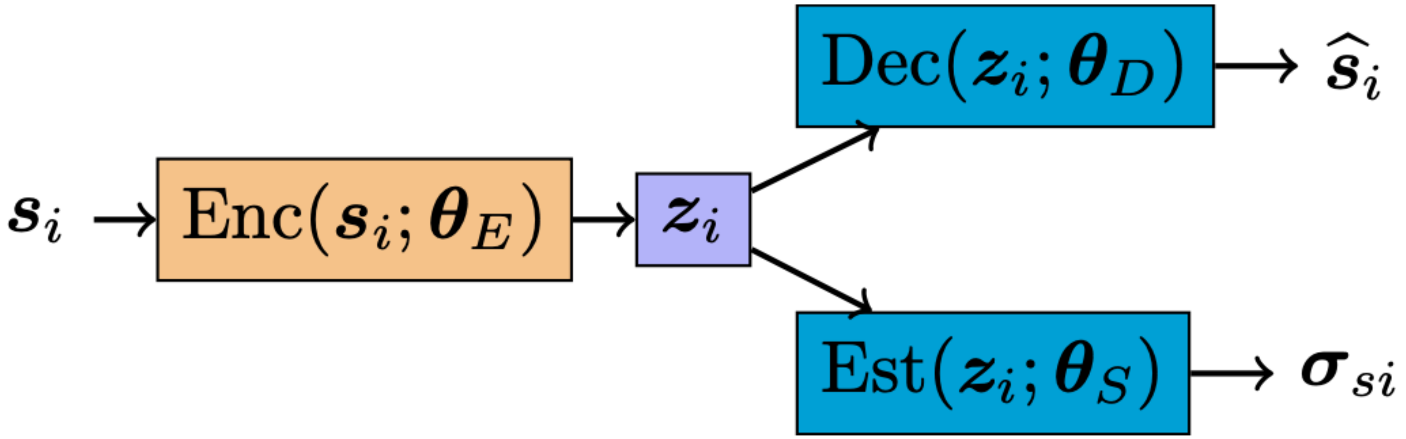


And a very young Ron Fox (retired)

Backup

Deep Learning Approach: Single Pulse

Single pulse parameter estimation:



$$\min_{\boldsymbol{y}_s, \boldsymbol{\theta}_E, \boldsymbol{\theta}_S, \boldsymbol{\theta}_D} \frac{1}{n_s} \sum_{i=1}^{n_s} \left\{ \|\boldsymbol{s}_i - \boldsymbol{p}_s(\boldsymbol{y}_s + \boldsymbol{\sigma}_{is})\|_2^2 + \lambda \|\boldsymbol{s}_i - \widehat{\boldsymbol{s}}_i\|_2^2 \right\}$$

where

$$\boldsymbol{y}_s := \{t_s, a_s, r_s, k_s, o_s\}$$

Classification and Fitting Single/Double Pulse