Software & Computing Round Table

Jefferson Lab: Year in Review

Brad Sawatzky (JLab)

Greatly indebted to: Markus D., Bryan H., Graham H., Makoto A., David L., and the dozens of others who are doing the heavy lifting on these efforts!

JLab Data Center Updates

- Storage capabilities for varying IO profiles including:
  - → 6PB in two Lustre Parallel filesystems
  - → 5PB NFS and XRootD storage on ZFS
  - → A tape library capable of over 100PB of long term storage
  - → Burst disk storage for tape buffering at >10 gbps

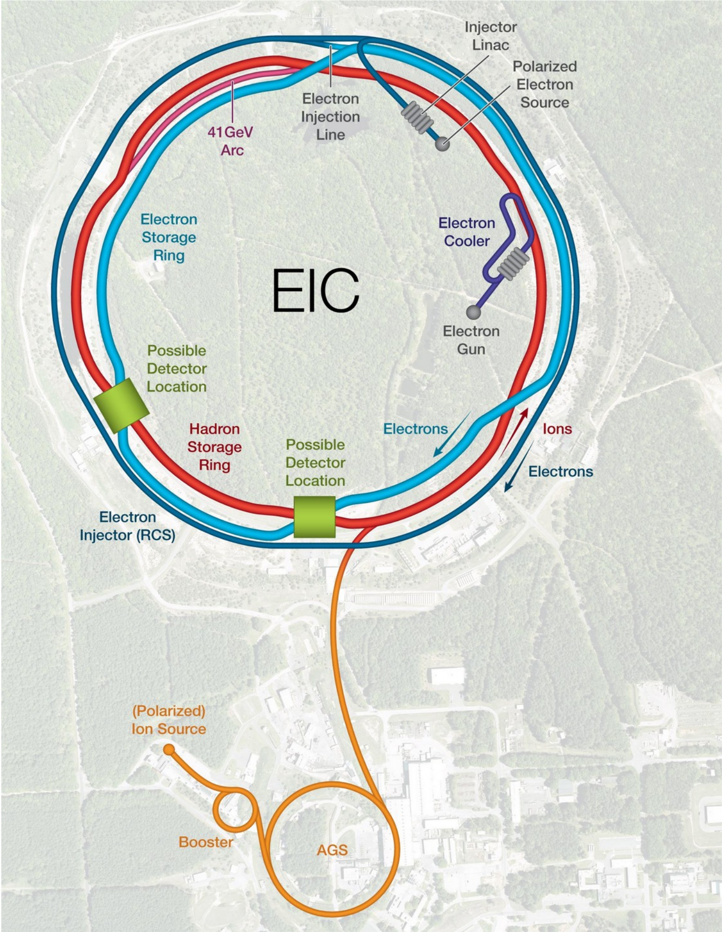
- Compute Resources
  - → GPU-enabled Machine Learning worker nodes
    - » 8 NVIDIA A100 GPUs (80GB)
    - » 56 NVIDIA T4 GPUs
  - → Equivalent of 9000 cores of AMD Milan CPUs
    - » Mix of Intel and AMD CPUs (mult. gen.)
  - → LQCD compute clusters including
    - » >400 KNL nodes cluster (OPA fabric)
    - » 256 NVIDIA 2080 GPU cluster (OPA fabric)
    - » 64 AMD MI100s GPU cluster (IB EDR fabric)

JLab Data Center Updates (Cont...)

- Built an additional hot aisle containment system for capacity in the data center.
- Completed the 100Gig scientific computing network upgrade (with the ESNet connection as the last mile)
- Open Science Grid engagement
  - → Established a multi-VO SciToken infrastructure for User Authorization which is being tested now with XRootD storage. This includes VO management using InCommon tools and can support federated identities.
  - → Built out our Open Science Grid Infrastructure to support multiple VOs and remove single points of failure in servers and services.
- Developing plans to improve and encourage use of containers and supporting infrastructure
- Deploying a new high-uptime ESX/VMware cluster to support Hall operations
- Increased our raw data ingest capability to support 3 halls at high rates with data duplication and data integrity verification (GlueX scale)
- SWIF version 2 for workflow management in full production, retired legacy tools.
- DTN (data transfer node) upgrades for Globus and XRootD (100Gig capable)
- Began construction of a Testbed for system development and research projects (LDRD, FOA, etc)
- Projects to improve farm utilization and stability (CPU and IO performance)
  - → Flagging underperforming jobs (low CPU or memory use, inefficient IO profiles on Lustre)
  - → New Tape IO strategies using disk caching to avoid that can cause IO bottlenecks to tape. (small files, file replication, hot spots in the user-managed cache)
  - → ML methodologies for data center optimization begun (LDRD2303)

Working Towards the EIC

- EIC: world's first collider of:
  - →Polarized electrons and
    - » polarized protons
    - » light ions (d, He, ..)
    - » heavy ions (up to Ur)
- Allow precision study of nucleon/nucleus at scale of sea quarks and gluons
- Jefferson Lab and BNL are host laboratories for EIC experimental program



Our Vision for Software & Computing at the EIC

“The purpose of computing is insight, not numbers.” Richard Hamming (1962)

Software & computing are an integral part of our research:

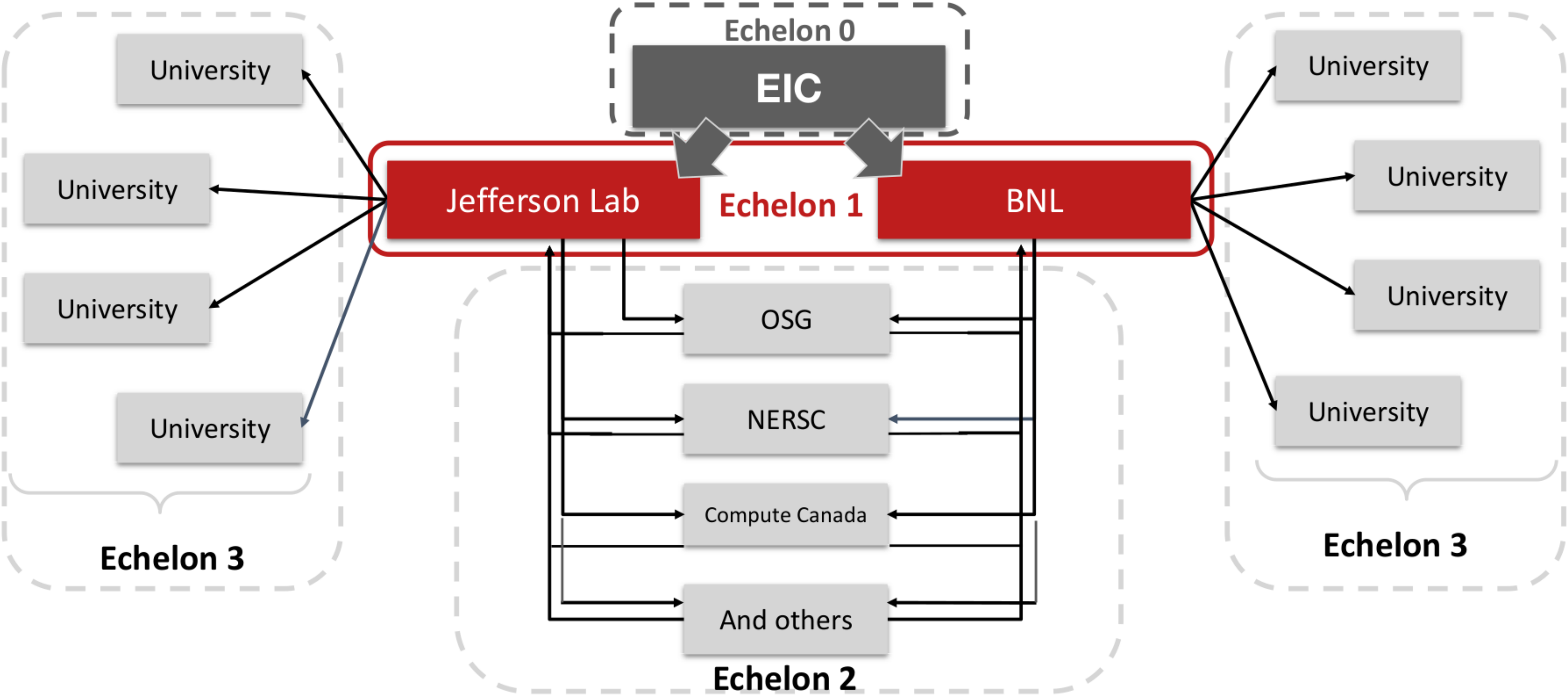
Survey among NP Ph.D. students and postdocs in preparation of ”Future Trends in NP Computing”

- **Goal** We work with our large Users Organization (over 1600 scientists from over 275 institutions) on data-intensive challenges and AI/ML and would like to ensure that also for the EIC scientists of all levels worldwide can participate in EIC analysis actively.
- **User-Centered Design**: To achieve this goal, we must develop simulation and analysis software using modern and advanced technologies while hiding that complexity and engage the wider community in the development.

Rapid turnaround of data for the physics analysis and to start the work on publications:

- **Goal**: Analysis-ready data from the DAQ system.
- **Compute-detector integration** with AI at the DAQ and analysis level.

Distributed Computing Model

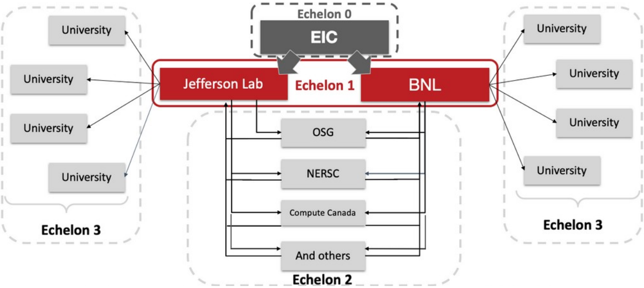


Nearly all storage (raw data, reconstructed data, simulated data) is stored across **Echelon 1** sites.

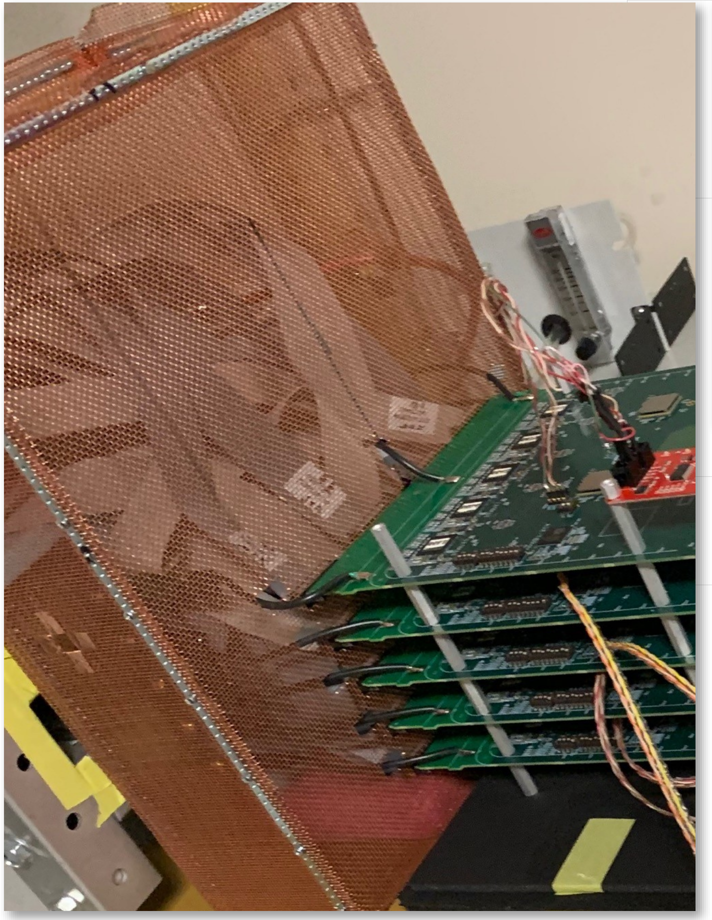
JLab-Supported Research Efforts

- Resilient and reconfigurable data transport layer
  - → EJFAT → IRIAD (new)
- AI/ML supported system management
  - → Data Center Reliability (new)
  - → AIEC, HYDRA
- Study challenges integrating research infrastructure across distributed sites
  - → JIRIAF (new)
- SRO and Data Analysis R&D
  - → JANA2
  - → ERSAP
- Exploiting AI/ML and heterogenous architectures
  - → PHASM
- Address Challenges of Autonomous Control and Experimentation
  - → INDRA-ASTRA

- Federated approach required
  - → Implementing login/auth for access to shared computing resources
  - → Deploy containerization and distributed software shares (cvmfs, etc)
- Data Management
  - → Rolling out XRootD and Rucio to provide a common data management framework and support distributed computing



Address Challenges of Autonomous Control and Experimentation



INDRA-ASTRA

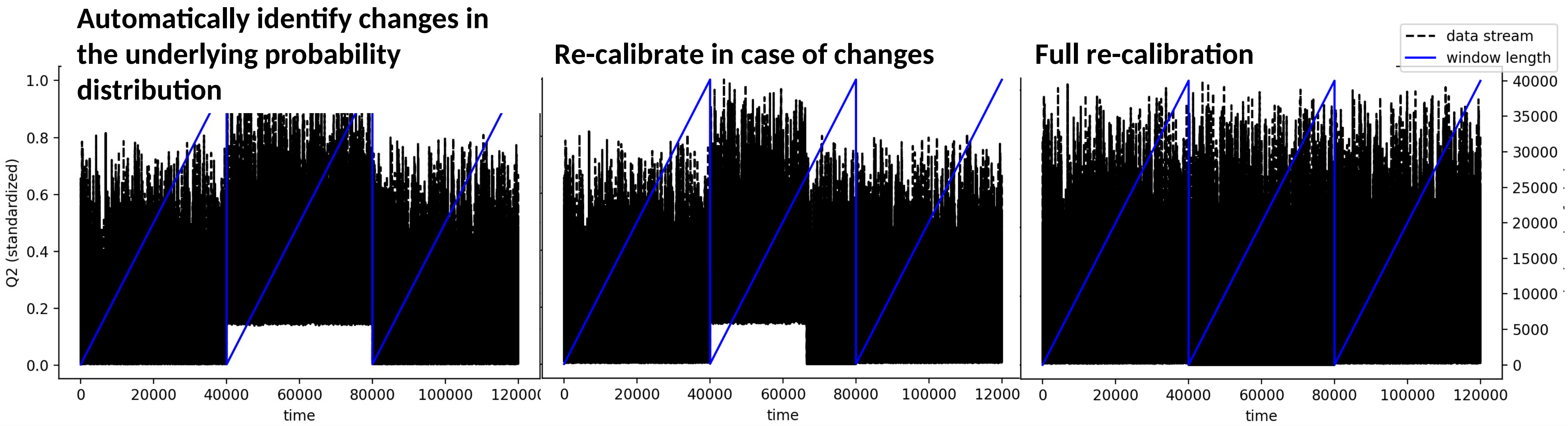
Develop a prototype for a fully automated, responsive detector system as a first step towards a fully automated, self-conscious experiment.

R&D integrated with streaming readout and AI/ML efforts at Jefferson Lab; near real-time processing of data in JANA2

Status

- Developed method for autonomous calibration of DIS experiments using baseline calibrations and autonomous change detection.
- Developed ADWIN2 and multiscale method for autonomous change detection.
- Versatile multiscale method can be used to increase reliability of data and find and fix issues on time.

INDRA-ASTRA: Automated data-quality monitoring and calibrations



Adapted and extended ADaptive WINdowing technique (ADWIN)

- detecting distribution changes, concept drift, or anomalies in data streams
- detection with established guarantees on the rates of false positives and false negatives
- keeps a sliding window with the most recently read data
- whenever two sufficiently large sub-windows of have sufficiently different means, then it is likely the corresponding expected values are different, and the older portion of the window is dropped
- prediction intervals without any prior assumption on the underlying distribution of data samples

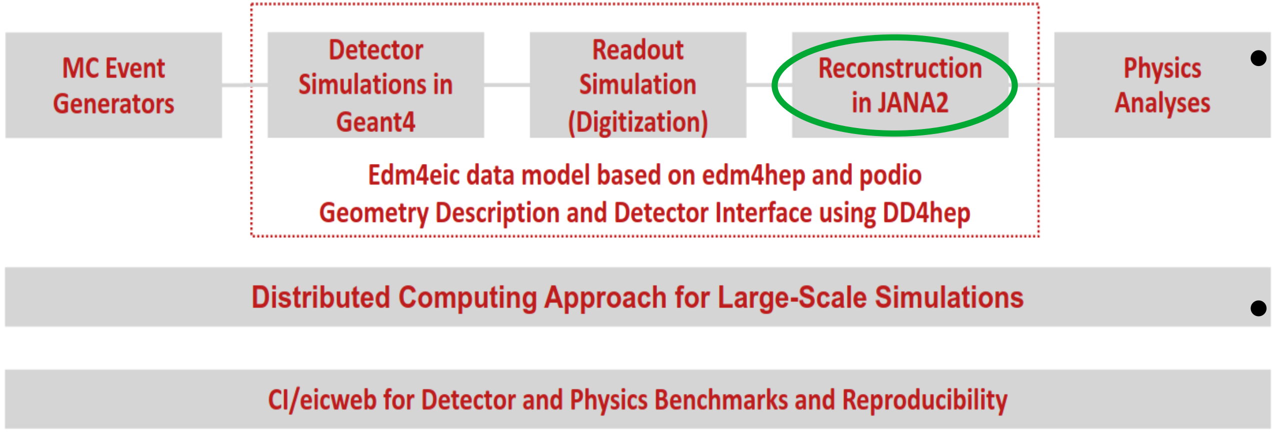
Selection of JANA2 for EIC ePIC Detector

EPIC Software & Computing

Clear mandate: Commonality; one software stack!

The design of the **modular simulation, reconstruction, and analysis toolkit** for the development of the EPIC detector and science program is based on the Statement of Software Principles and a decision-making process involving the whole community. The decisions have been well reviewed in the **EIC Software Infrastructure Review**.

Simulation, Reconstruction, and Analysis



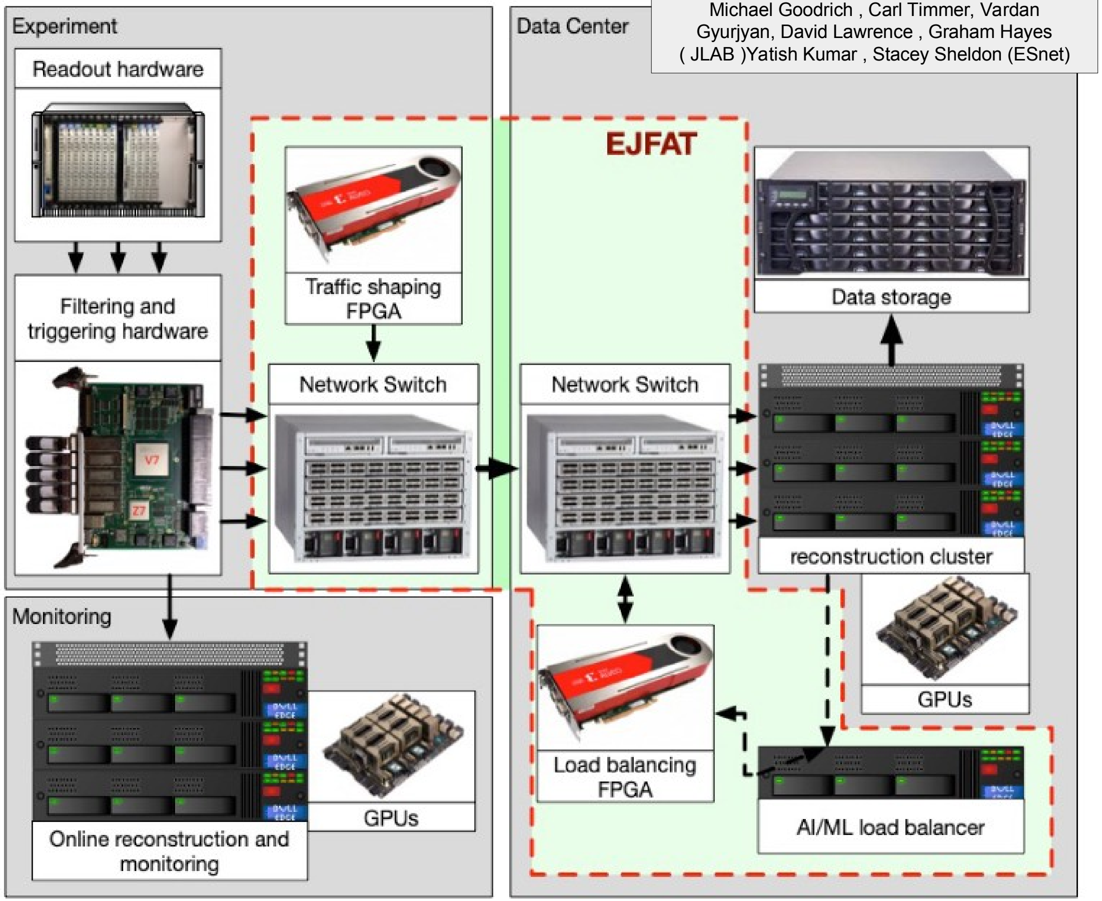
EPIC Collaboration Meeting, January 8, 2023. 2

- JANA is a multi-threaded reconstruction framework project with nearly 2 decades of experience behind it, written in optimized low-level C++.
- JANA2 is a rewrite that incorporates modern CS practices. Aims to be clean, lightweight, and user-friendly. “Paying off technical debt”
- JANA2 introduces an innovative dataflow parallelism paradigm under the hood, which has implications for streaming readout and heterogeneous hardware utilization
- Projects already using JANA/JANA2
  → GlueX, INDRA-ASTRA, BDX, TriDAS (+ERSAP) + JANA2 Streaming DAQ

ESnet JLab FPGA Accelerated Transport (EJFAT)

- Prototype dynamic steering of streaming data
  - → Data format contains metadata describing content
  - → Using standard IP based network, all traffic is directed to an FPGA device
  - → Firmware modifies IP packet headers to reroute data based on data type, and what kind of destination it should stream to
  - → Route data based on the kind of data it is, where it came from, and where it needs to go
  - → Reroute data in-flight when system architecture reconfigures
  - → Remote data sources can be agnostic of the destination hardware configuration
- Initial implementation has static translation tables to implement data distribution schemes
  - → Round robin, sorting by time or source
- Goal is dynamic, intelligent steering of data

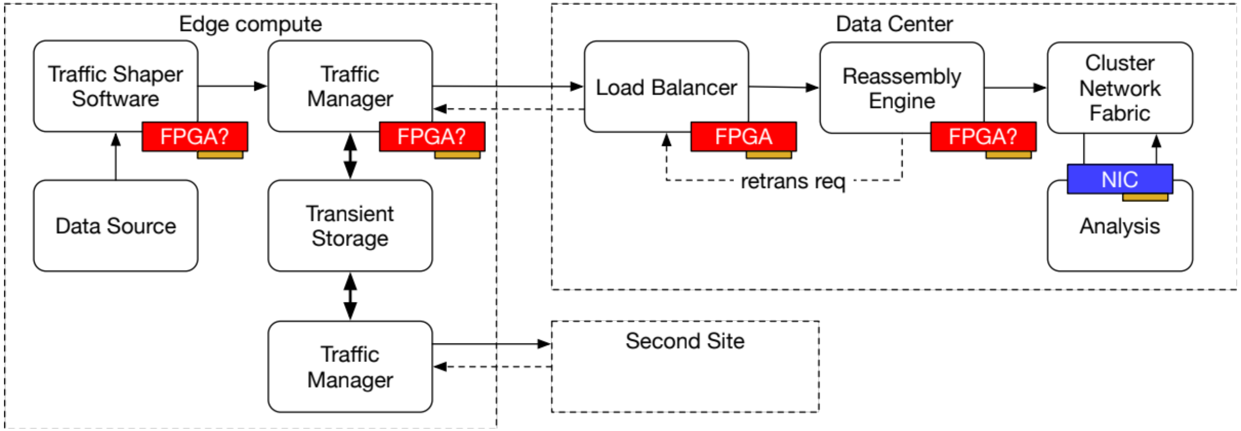
Michael Goodrich , Carl Timmer, Vardan Gyurjyan, David Lawrence , Graham Hayes (JLAB)Yatish Kumar , Stacey Sheldon (ESnet)



Integrated Research Infrastructure Architecture Demonstrator (IRIAD)

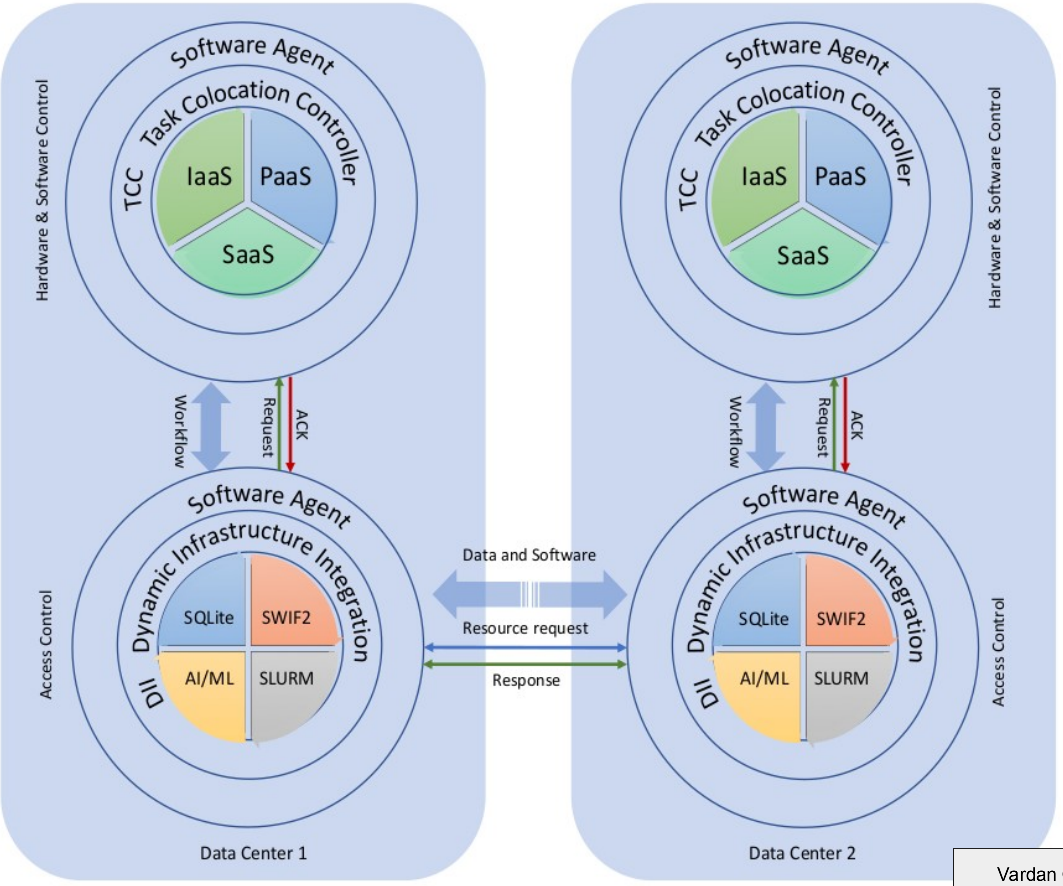
Graham Heyes (PI), Amitoj Singh (Co-PI)

- In FY22 the ASCR Funded EJFAT project developed an FPGA based network traffic load balancer to dynamically steer streaming data into a data center.
- IRIAD builds on EJFAT to deliver a demonstration system with the following features:
  → Generic interface to an experiment at the edge compute and to analysis in the data center.
  → Data delivery into the memory of heterogenous hardware as though the IRIAD system did not lie between data source and analysis
  → Transient storage to smooth data flow and allow for retransmission on error
  → The ability to load balance or failover between two data center sites
  → Resilience and fault tolerance within the data center
  → Guaranteed data delivery and processing for time critical use cases
- IRIAD will demonstrate core technology for a HPDF to handle time-critical science use cases and capability to stream data between a detector and data center at geographically diverse sites (ie. EIC/BNL and JLab)



JLab Integrating Research Infrastructure Across Facilities (JIRIAF)

- Study the challenges of heterogeneous, distributed, and opportunistic compute resource provisioning from several participating data centers that will be presented to and end-user as a single, unified computing infrastructure.
- Develop policies and requirements for computational workflows that can effectively utilize opportunistic resources
- Appraise capabilities to combine geographically diverse computing facilities into an integrated science infrastructure.
- Evaluate an infrastructure that dynamically integrates temporarily unallocated or idled compute resources from various providers.

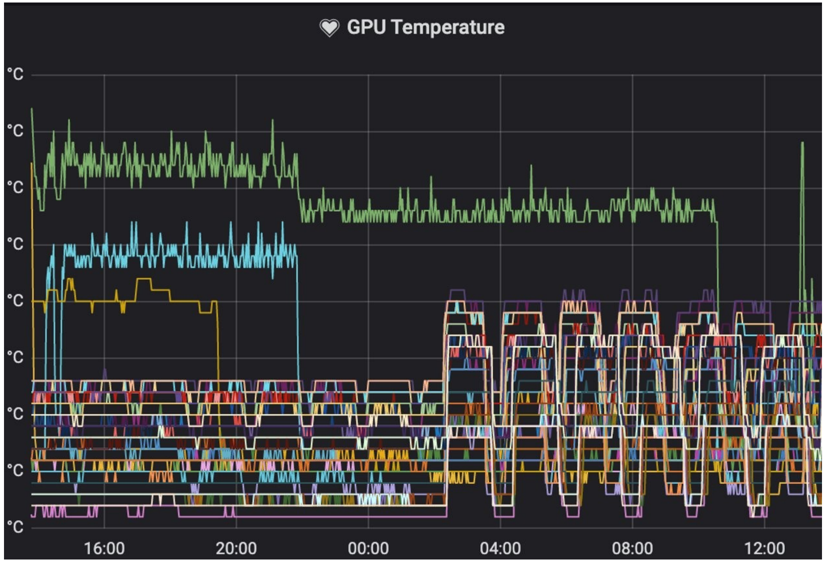


Vardan Gyurjyan
Christopher Larrieu

Adaptive Strategies for Optimal Computing Availability

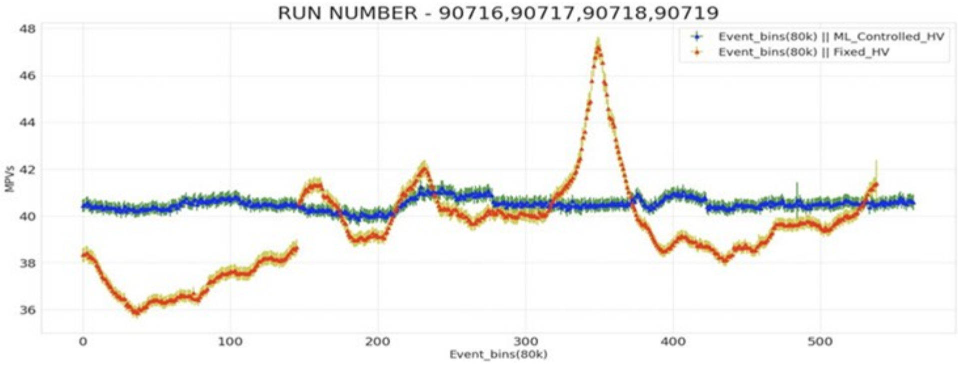
- A vision for making data centers operate more efficiently using AI/ML to
  - → Avoid resource contention
  - → Perform dynamic resource allocation
  - → Facilitate real-time, streaming work
- Knowledge Base curated from storage, compute, network, and environmental sources.
- Phase 1 (this LDRD): building the Digital Twin, and performing anomaly detection and prognostication
  - → Building a test environment
  - → Model refinement with live cluster data
  - → Testbed environment for vetting changes
  - → Initial focus: Single Node HTC-style (farm) jobs

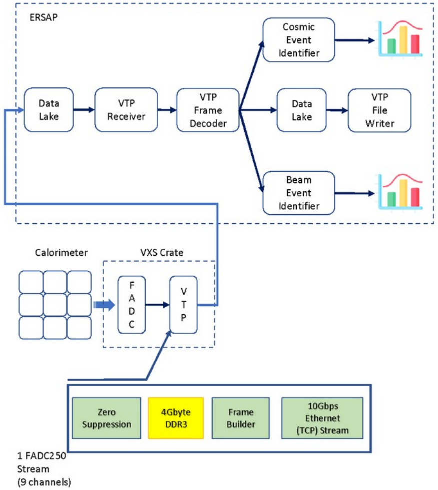
- Bryan Hess, PI
- Malachi Schram, Co-PI
- *A Multidisciplinary Collaboration between Data Science and Scientific Computing Operations*

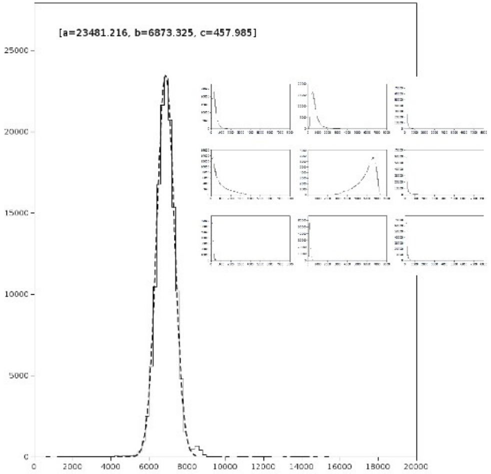


Ongoing Projects: AIEC, ERSAP, PHASM

- AIEC – AI for Experiment Calibration/Controls
  - → Operational: GlueX CDC detector
- ERSAP – Environment for Real-time Streaming, Acquisition and Processing
  - → Deployed EIC prototype calorimeter beam test application at DESY
  - → Rapidly prototyped and debugged DESY data-stream processing code based on shift taker request
  - → Successfully identified beam leptons in near real time
- PHASM – "Parallel Hardware viA Surrogate Models"
  - → Explore creating AI-based surrogate models to adapt legacy code to modern hardware (ie. GPUs, FPGAs)

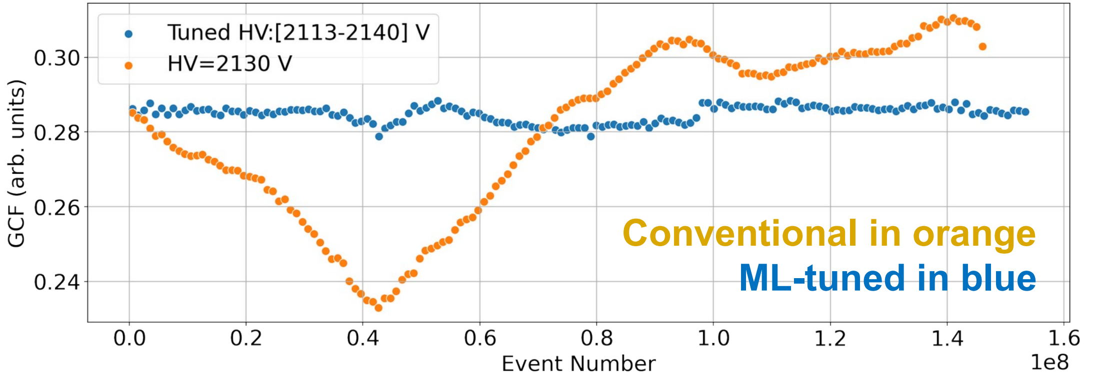


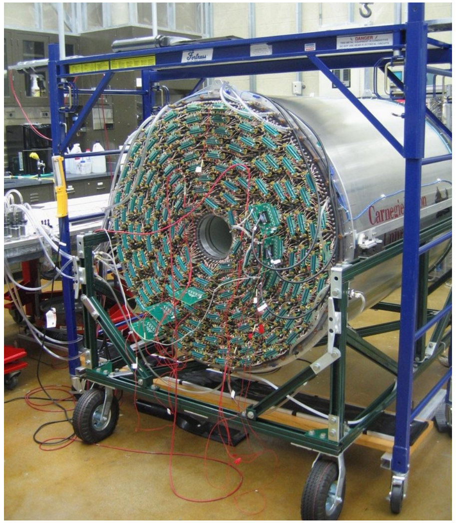


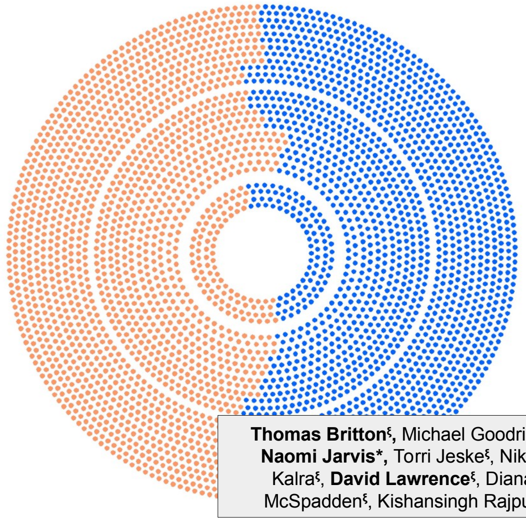


AI Experimental Calibration and Control (AEIC)

- Sensitive detectors need to be calibrated to obtain optimal resolution
  - → Calibrations cause a delay between data collection and analysis (weeks-months)
  - → Multiple iterations are needed to converge to final set of constants
  - → Development platform
    - » GlueX Central Drift Chamber
    - » Gaussian process w/ 3 features:
      - – atm. pressure, gas temp, HV board current draw
- Main Goal:
  - → Dynamically adjust the controls of a sensitive detector to reduce/eliminate need for calibration







Thomas Britton[§], Michael Goodrich[§], **Naomi Jarvis***, Torri Jeske[§], Nikhil Kalra[§], **David Lawrence**[§], Diana McSpadden[§], Kishansingh Rajput[§]

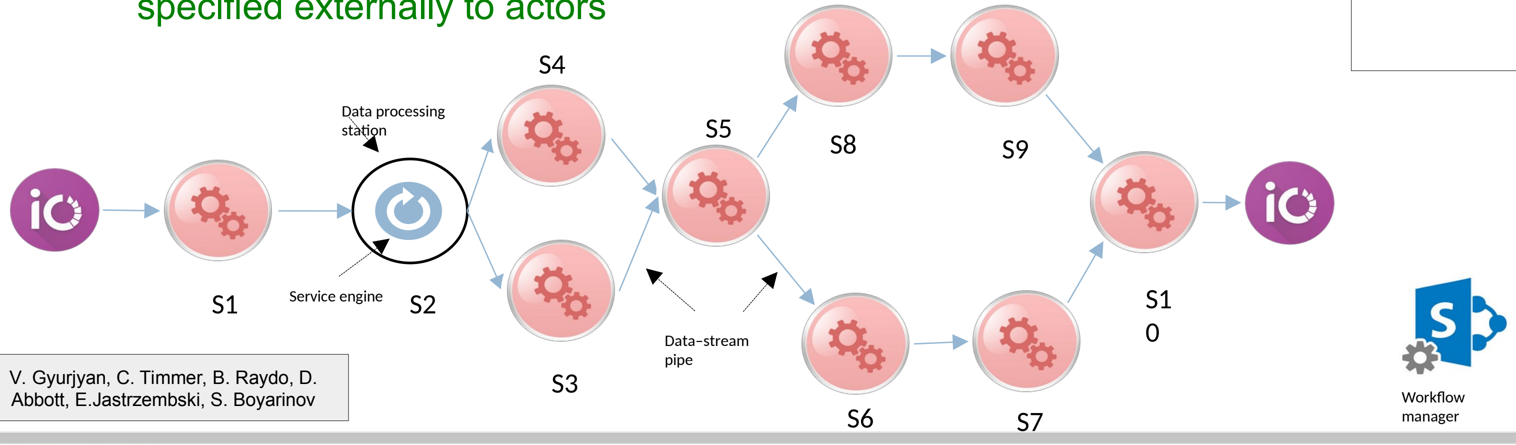
Future Plans (AEIC)

- Investigating system incorporate an uncertainty quantification into the process
  → develop a system that can say “I don’t know” and will collect more data to “learn”
- Successfully proposed as capstone project for UVA Data Science Department Master’s program
  → Started meeting with 4 students. They will start work in earnest in Spring Semester
- Focus will shift to 2880 element E/M calorimeter
  → Sophisticated LED monitoring system can be used as input

Milestones	
FYQtr	**Description**
FY23Q2	• Review historical data. • Prepare data sets including FCAL-LED skim files and existing calibrations. • Develop model(s) to extract relative calibration constants from LED events.
FY23Q3	• Map the relation between gain and HV • Implement policy for detecting and dealing with problem PMTs.
FY23Q4	• Integrate automated system for AI/ML control of FCAL PMT HV • Automatic generation of calibration constants into operations

ERSAP – Environment for Real-time Streaming, Acquisition and Processing

- ERSAP is a software LEGO system
  - » Looking to serve needs of future experiments at JLAB and EIC
  - » Key focus is supporting a streaming readout (SRO) model
- → Reactive event driven actor, data-stream pipe, and orchestrator
- → Stream of data quanta, flowing through directed graph of actors
- → Application is a network of independent “black-box” actors (including user-provided algorithms)
- → Data move across actors, not instructions
- → Actors communicate with each other by exchanging the data quanta across connections by message passing, where connections are specified externally to actors



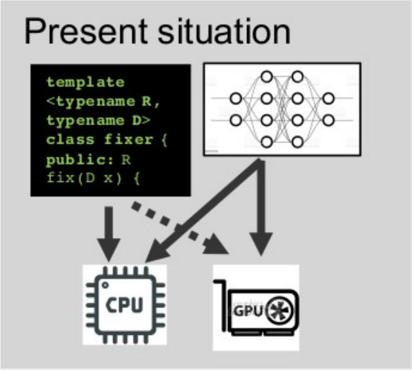
V. Gyurjyan, C. Timmer, B. Raydo, D. Abbott, E.Jastrzembski, S. Boyarinov

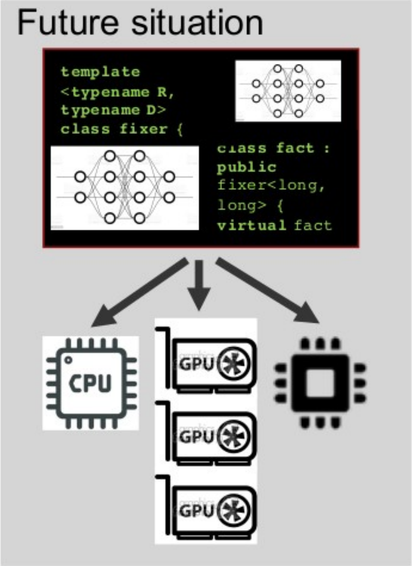
- Deployed EIC prototype calorimeter beam test application at DESY
- Rapidly prototyped and debugged DESY data-stream processing code based on shift taker request
- Successfully identified beam leptons in near real time

Parallel Hardware viA Surrogate Models (PHASM)

Nathan Brei
Xinxin Mei
David Lawrence

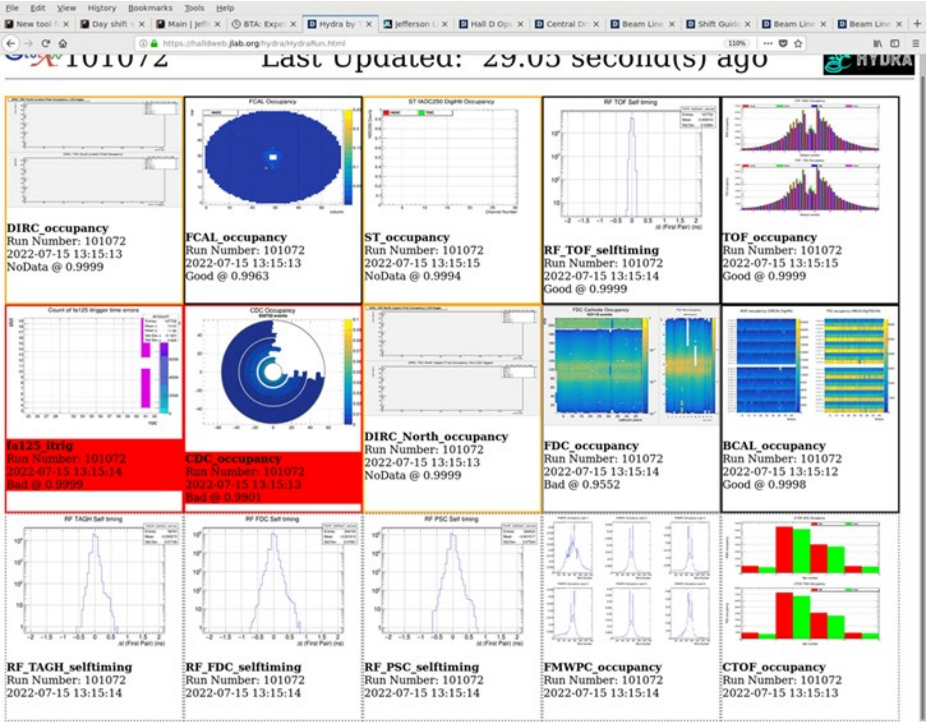
- What is PHASM
  - → LDRD project at JLab
  - → 1 year old; 2–3 developers
  - → Proof of concept
- Basic Idea
  - → Simplify training a neural net surrogate model to mimic and replace an arbitrary piece of existing numerical code.
  - → Systemize and formalize the process from analysis to deployment.
- Perspective Shift
  - → A neural net surrogate model of an algorithm is a *transformation* of that algorithm.
  - → Eventually, classical numerical methods and their data-driven analogues will be understood under a unified theory.





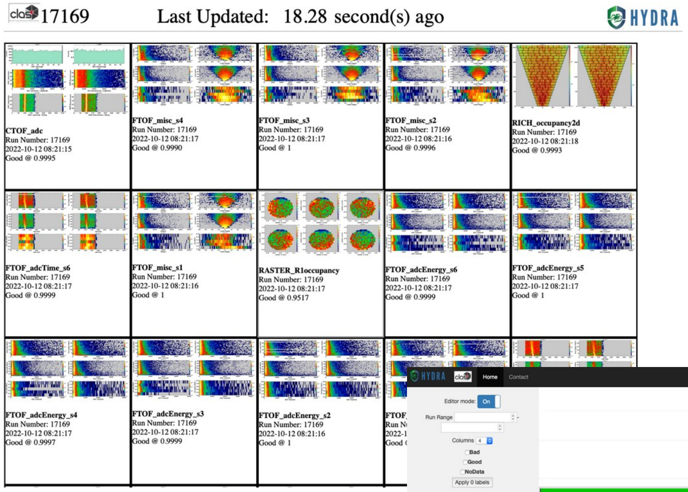
Ongoing Projects: HYDRA, G4 Development

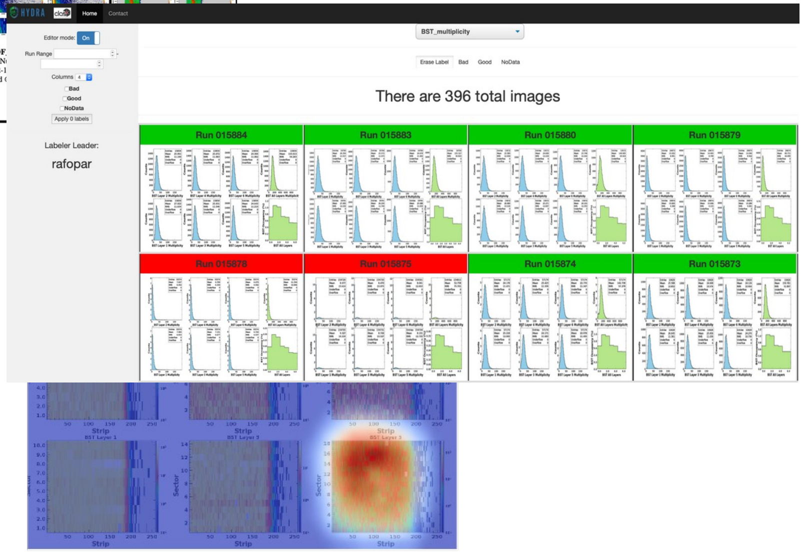
- HYDRA – AI based detector monitoring
  - → Now deployed in all JLab Halls (A, B, C & D)
    - » initially developed in Hall-D
  - → Significant advancements to Hall B deployment
- Geant4 / eAST R&D at JLab
  - → In May Makoto Asai was awarded Suwa Prize by the Foundation for High Energy Accelerator Science for his work on Geant4
  - → US-Japan collaboration established.
  - → SBIR funding with Tech-X was secured for the second year
  - → Collaboration with NASA on "Radiation transport in very thin material for Artemis project" initiated.
  - → Started on sub-event parallelization
  - → Ongoing work to support high concurrency hetrogeneous architectures

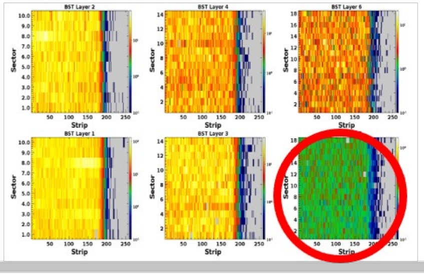


- AI based detector monitoring
  → Do the anomaly detection that is critical for smooth data taking, but is tedious and exhausting for humans
- Web-based UI
  → Tagging/Labeling
  → Anomaly monitoring
- Extensible Framework
  → All JLab Halls have deployed HYDRA
- Future development
  → Continue to streamline UI and generalize software deployment
  → Determine and flag data regions that triggered anomaly for users
    » ie. reveal *why* Hydra labeled something ‘bad’

Thomas Britton, David Lawrence, Torri Jeske, Kishan Rajput and Nathan Baltzell







eAST (eA Simulation Toolkit)

Scientific Achievement

A simulation toolkit based on the latest version of Geant4 that is supposed to be the common base of all the electron-ion collider detector simulation studies.

Significance and Impact

We promote the most proper and efficient use of the latest Geant4 functionalities in both physics and computing aspects for all the EIC-related simulation efforts and also, through this project, we ensure the timely development of Geant4 itself reflecting the needs of EIC.

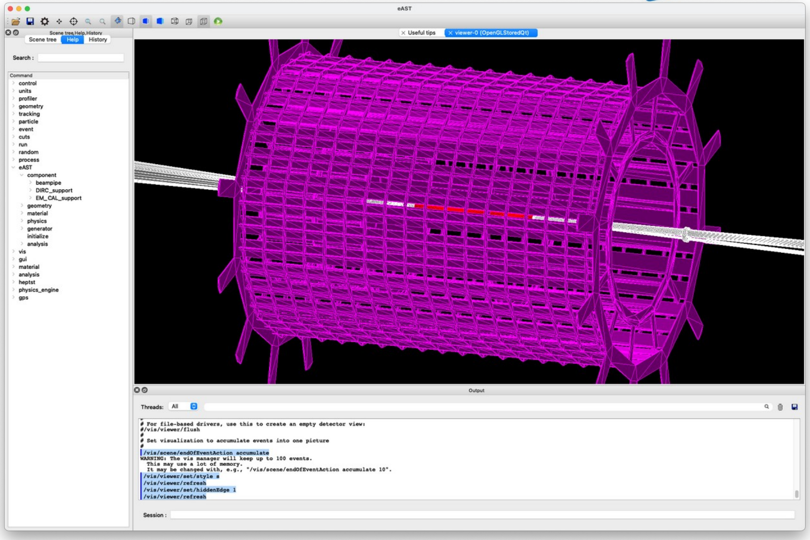
Research Details

– Implementing interfaces (e.g. CAD interface through GDML, HepMC3 interface with afterburner) and utilities (e.g. geometry sanity check routines) in modular manner easily controllable by macro commands and reusable with experiment's frameworks.

– Maintaining and benchmarking proper set of Geant4 physics models that are most appropriate for EIC.

– Offering a framework for heterogeneous computing architecture, in particular combination of CPU and GPU, through the novel sub-event parallelism.

– Offering a mechanism of controlling physics and geometry options per detector components (a.k.a. regions).



Screenshot of eAST GUI with geometry of Beam pipe, DIRC and EM_CAL supports @IP6 / Athena. Three GDML files of these structures are loaded separately and combined in eAST.

Reference:

eAST GitHub source code repository :
https://github.com/eic/east

eAST documents :
https://eic.github.io/east/content/manual.htı

Project Lead: Makoto Asai

eAST Status and Plans

Synchronizing with Geant4 releases

- Issues and extensions spotted through eAST development were ported back to Geant4 and included in Geant4 patch releases or new version of Geant4
- eAST is validated every time new Geant4 version (patch or regular) released.

Physics options for eAST

- A physics list that is most suitable for EIC physics resume is developed and maintained / tested.
- It optionally includes processes for low-energy neutron, electro- and gamma-nucleus, optical photon, as well as shower parameterization.
- In 2023, some additional event biasing options are planned be added.

"Pythonization"

- eAST is currently controlled by UI commands (interactively or through macro file).
- In 2023, we plan to offer Python layer to control eAST (and key functionalities of Geant4).

Interfacing to DD4Hep

- Geometry of eAST can currently be defined either by C++ coding or through GDML files
- DD4Hep is the choice of ePIC for geometry exchange format. In 2023, we plan to develop DD4Hep/eAST interface

Heterogeneous computing

- As Geant4 is moving toward sub-event parallelism with tasking mechanism, some EIC-specific physics models that are suitable to running on GPUs will be developed and tested.

Geant4 Status and Plans

Sub-event parallelism is coming

- Task-based event parallelism was released in Dec 2021, that allows the development of sub-event parallelism.
- 2022 : reorganization of Geant4 kernel to remove some loop-dependencies foreseen with sub-event parallelism
- 2023 : Phase-I sub-event parallelism
  - Master thread organizes sub-event tasks and merges the results.
  - Tasks still have the common physics processes and entire detector geometry
- 2024~ : Phase-II sub-event parallelism
  - Each task has only its unique physics processes and sees the limited geometry that are necessary for that particular task.
  - Some first physics models suitable for GPU

Physics validation for EIC (and NP in general)

- The Geant4 physics validation portal (https://geant-val.cern.ch/) does not have much of the validation tests suitable for EIC or Nuclear Physics in general.
- JLab will lead the effort of enriching test cases based on JLab (and NP in general) experiment results.
  - We need both data and experiment configuration.

Geant4 Technical Forum

- The next Geant4 Technical Forum (where Geant4 users and developers exchange information) is scheduled on February 2nd.
- https://geant4.web.cern.ch/collaboration/technical_forum
- We invite your participation!

QuantOm: Advancing the Experimental-Theoretical Workflow

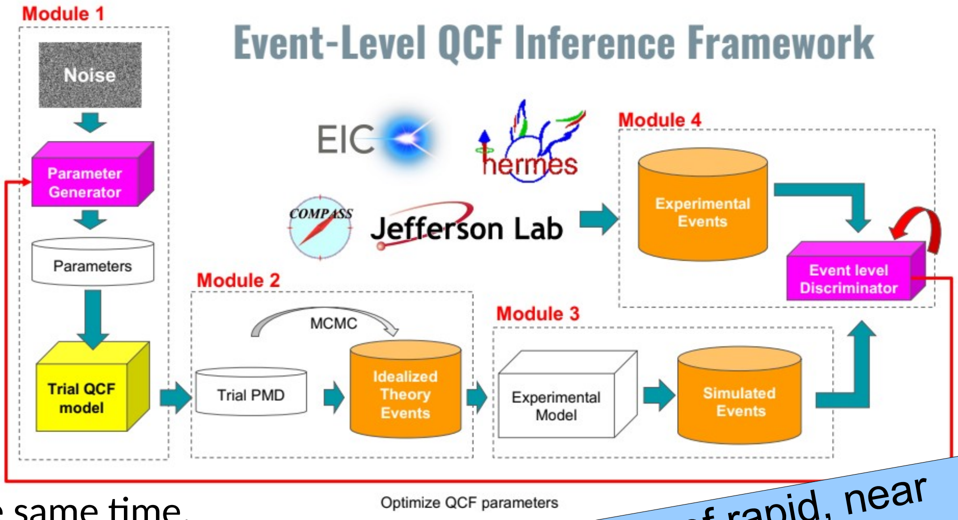
Funded via SciDAC
(Scientific Discovery Through
Advanced Computing)

- **Developing a workflow on the event level:**
  - The extraction of PDFs, TMDs, and GPDs is a multidimensional data challenge. We analyze high statistics data sets with strong correlations in five or more kinematics and with various final-state particles. Access to the data on event level allows theoreticians to studying these correlations directly.

Developing a joint experimental-theoretical workflow

  - Extracting PDFs, TMDs, or GPDs directly from the experiment allows experimentalists and theoreticians to work closely together. This not only removes the delay in providing the experimental measurement but truly enables joint experimental-theoretical wok.

- **Developing a HPC workflow:**
  - The extremely parallelized architecture allows to study the strong correlations in the data in an unprecedented manner, while maximizing the experimental precision at the same time.
  - The accelerated hardware of the new HPC systems is ideal for AI/ML, allowing us to do the parallelized workflow at the event level in near real-time.
    - EIC will produce analysis-ready data in near real-time using streaming readout and AI/ML.



Supports a vision of rapid, near real-time data processing from DAQ to physics analysis output

Summary

- Software and Computing continues to be an extremely active area for the JLab community.
- AI/ML is broadly recognized as a critical-path tool in future research.
  → All JLab divisions (Physics, Theory, CST, Accelerator, …) have engagement and are pursuing R&D projects along these lines.
- Sci-Comp resources (both software and hardware) are growing to meet the demands of the JLab program, and with the developing EIC vision in clear view.
- Data rates within JLab ramping up rapidly and will continue to do so (GlueX, CLAS12, SBS, NPS, MOLLER, SOLID)
  → Provides both incentive and a functional need to evaluate the status quo and develop new workflows/technologies
  → New systems allow for new approaches in DAQ design and software tooling to be supported
  → Dovetails nicely with EIC planning and R&D

Thank You

Useful links:

JLab EPSCI Group

Software & Computing Round Table

Future Trends in Nuclear Physics Computing