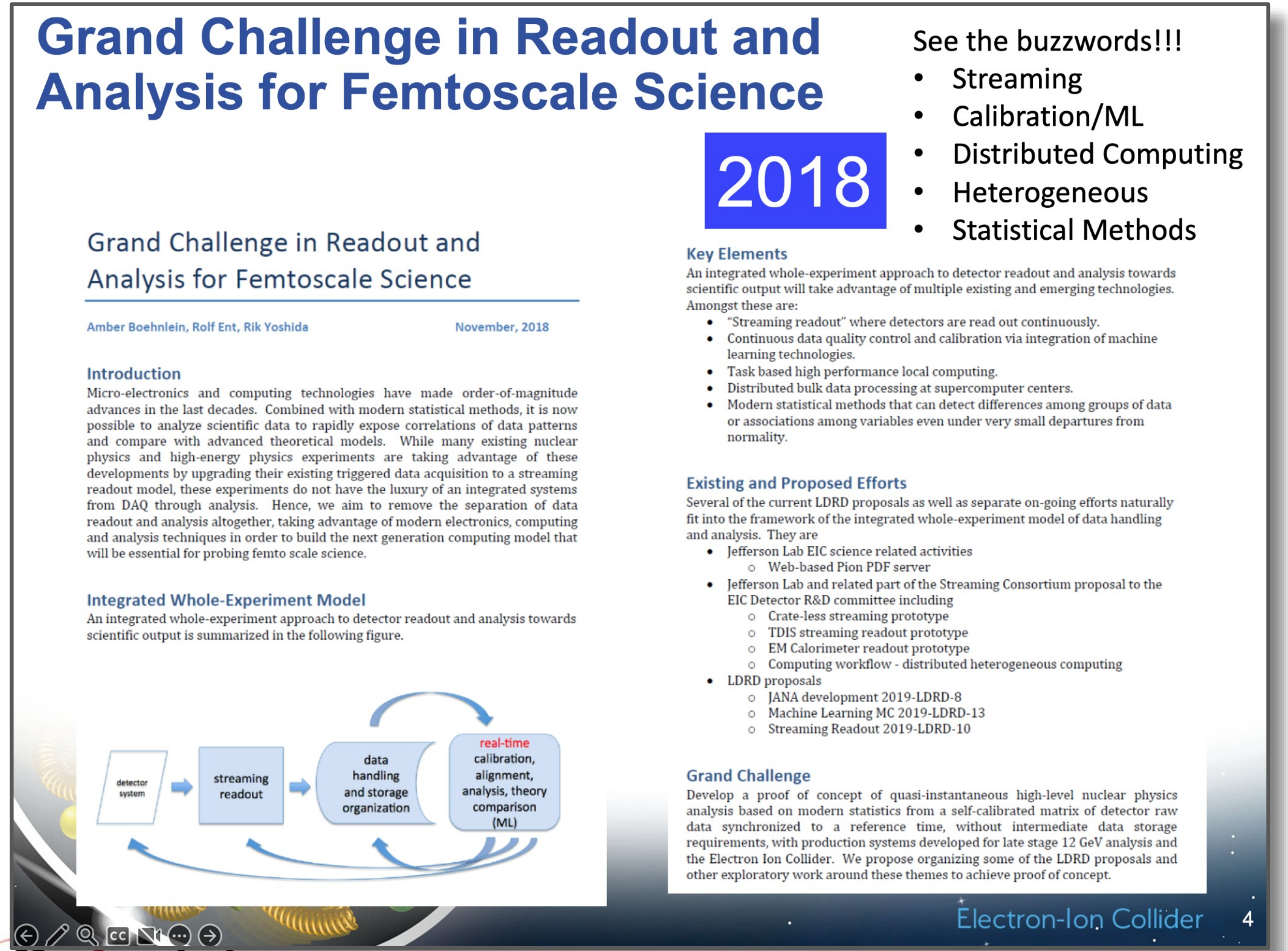
Streaming Readout DAQ: Challenges and Opportunities in Hall C

Brad Sawatzky (CST)
David Lawrence (EPSCI)

Streaming Grand Challenge

slide from Jan. 5, 2024 presentation by Rolf Ent



Grand Challenge in Readout and Analysis for Femtoscale Science

Amber Boehnlein, Rolf Ent, Rik Yoshida — November, 2018

Introduction

Micro-electronics and computing technologies have made order-of-magnitude advances in the last decades. Combined with modern statistical methods, it is now possible to analyze scientific data to rapidly expose correlations of data patterns and compare with advanced theoretical models. While many existing nuclear physics and high-energy physics experiments are taking advantage of these developments by upgrading their existing triggered data acquisition to a streaming readout model, these experiments do not have the luxury of an integrated systems from DAQ through analysis. Hence, we aim to remove the separation of data readout and analysis altogether, taking advantage of modern electronics, computing and analysis techniques in order to build the next generation computing model that will be essential for probing femto scale science.

Integrated Whole-Experiment Model

An integrated whole-experiment approach to detector readout and analysis towards scientific output is summarized in the following figure.

Key Elements

An integrated whole-experiment approach to detector readout and analysis towards scientific output will take advantage of multiple existing and emerging technologies. Amongst these are:

- "Streaming readout" where detectors are read out continuously.
- Continuous data quality control and calibration via integration of machine learning technologies.
- Task based high performance local computing.
- Distributed bulk data processing at supercomputer centers.
- Modern statistical methods that can detect differences among groups of data or associations among variables even under very small departures from normality.

Existing and Proposed Efforts

Several of the current LDRD proposals as well as separate on-going efforts naturally fit into the framework of the integrated whole-experiment model of data handling and analysis. They are

- Jefferson Lab EIC science related activities
  - Web-based Pion PDF server
- Jefferson Lab and related part of the Streaming Consortium proposal to the EIC Detector R&D committee including
  - Crate-less streaming prototype
  - TDIS streaming readout prototype
  - EM Calorimeter readout prototype
  - Computing workflow - distributed heterogeneous computing
- LDRD proposals
  - JANA development 2019-LDRD-8
  - Machine Learning MC 2019-LDRD-13
  - Streaming Readout 2019-LDRD-10

Grand Challenge

Develop a proof of concept of quasi-instantaneous high-level nuclear physics analysis based on modern statistics from a self-calibrated matrix of detector raw data synchronized to a reference time, without intermediate data storage requirements, with production systems developed for late stage 12 GeV analysis and the Electron Ion Collider. We propose organizing some of the LDRD proposals and other exploratory work around these themes to achieve proof of concept.

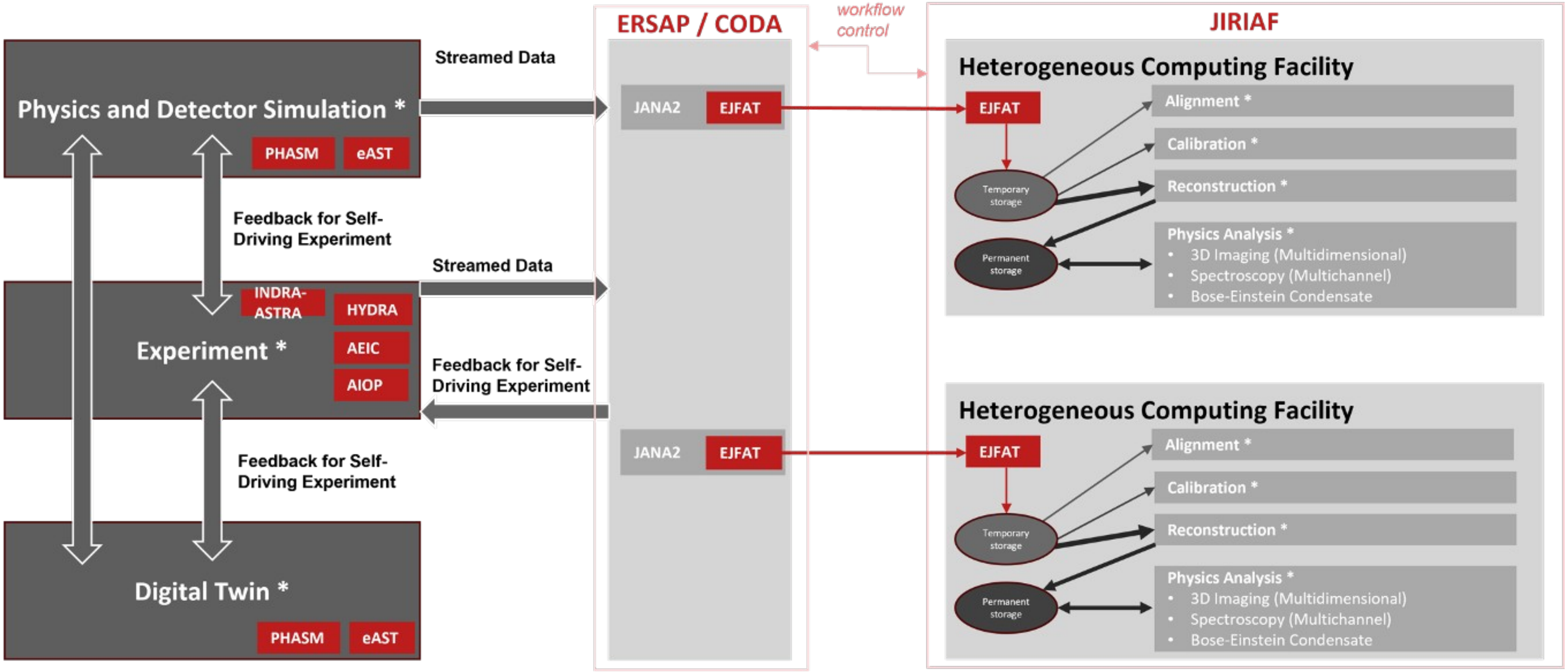
EPSCI
Experimental Physics Software and Computing Infrastructure

The Streaming Grand Challenge began in 2018.

Significant progress has been made since then on several fronts that include deployment of SRO-capable fast electronics, firmware development, and software (ERSAP, JANA2, InstaRec [CLAS12 ML track-recon.],...).

Phase II of the SRO GC is now beginning.

Streaming Grand Challenge Phase II



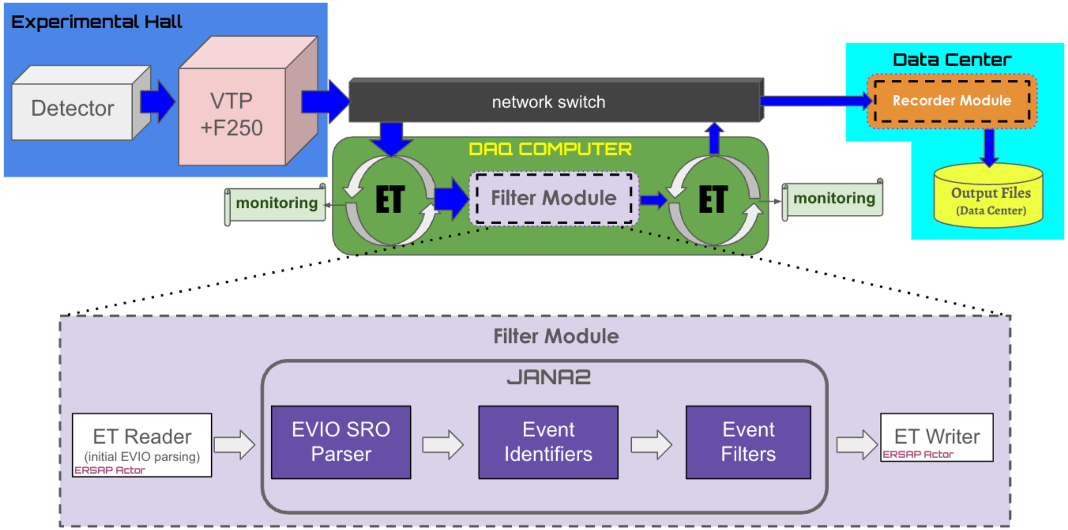
* AI/ML

Diagram from early 2024 by Rolf Ent, Markus Diefenthaler, Brad Sawatzky, and David Lawrence

GCII: Standard Test Setup

Experimental Physics Software and Computing Infrastructure

Hanjie L. *- Hall-C*
Dmitry R. *- EIC*
Brad S. *- SciComp+ENP*
David L. *- EPSCI*



Software Specification Document
for JLab Standard Streaming Readout Detector Beam Test
Tool Set (JSROB)

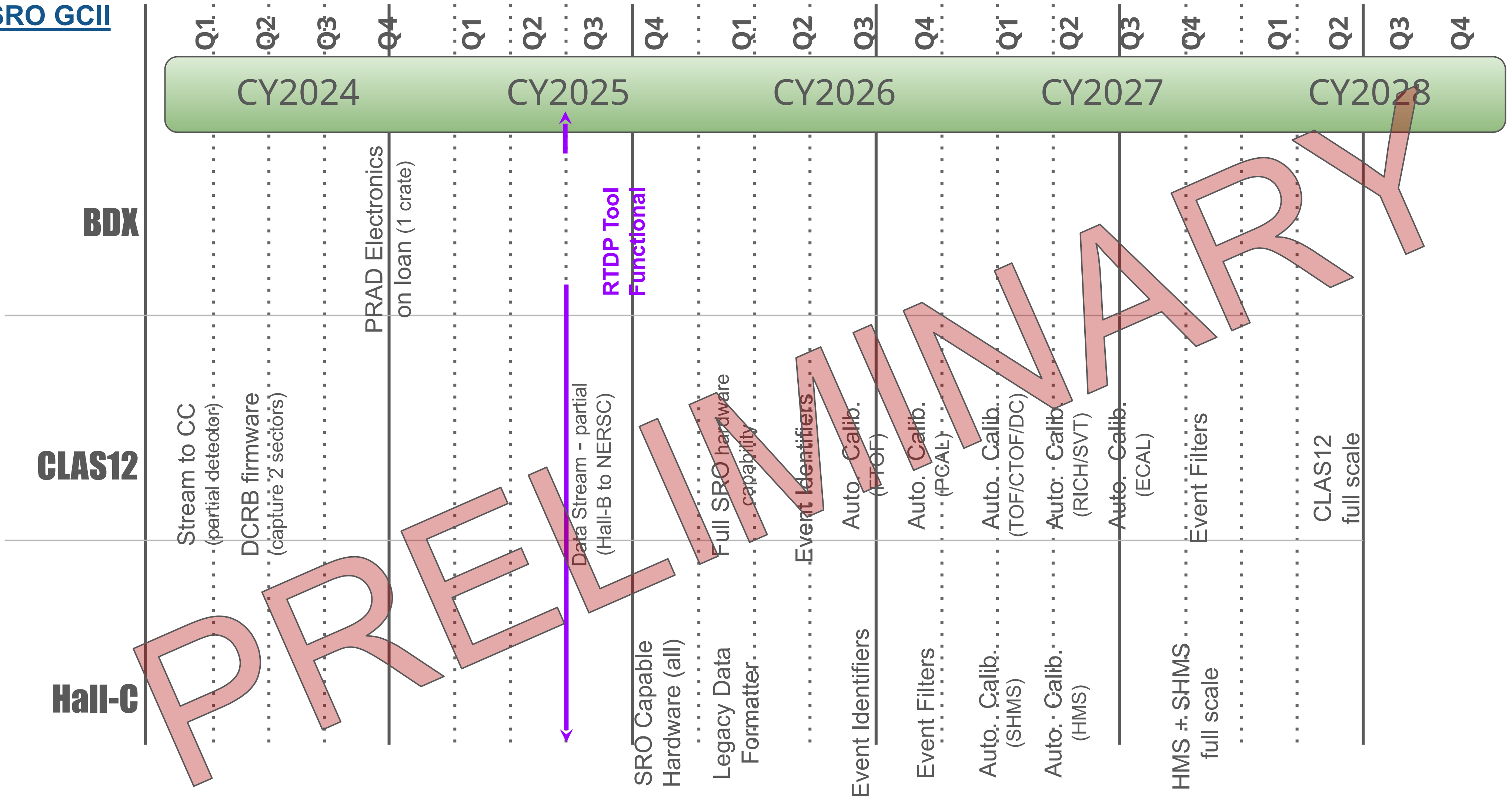
David Lawrence, Brad Sawatzky, Hanjie Liu, Dmitry Romanov

May/June 2024

1 Introduction

This document outlines the software and hardware specifications for a tool set intended for use in detector beam tests where small prototype detectors are brought in and tested with beam at Jefferson Lab. The tool set will support a Streaming Readout (SRO) Data Acquisition (DAQ) system, incorporating event identifier, event filter, monitoring, and event recording components. Users will have well defined places to insert custom algorithms for their detector(s). Basic monitoring with visualization will be provided with options for users to easily add their own histograms. Integration with the JLab SciComp farm will be included to couple streams directly to the HTC resource and the tape archive.

SRO GCII



CY2024
CY2025
CY2026
CY2027
CY2028
Q1 Q2 Q3 Q4
BDX
CLAS12
Hall-C
PRAD Electronics on loan (1 crate)
RTDP Tool Functional
Stream to CC (partial detector)
DCRB firmware (capture 2 sectors)
Data Stream - partial (Hall-B to NERSC)
Full SRO hardware capability
Event Identifiers
Auto. Calib. (TOF)
Auto. Calib. (PCAL)
Auto. Calib. (TOF/CTOF/DC)
Auto. Calib. (RICH/SVT)
Auto. Calib. (ECAL)
Event Filters
CLAS12 full scale
SRO Capable Hardware (all)
Legacy Data Formatter
Event Identifiers
Event Filters
Auto. Calib. (SHMS)
Auto. Calib. (HMS)
HMS + SHMS full scale
PRELIMINARY

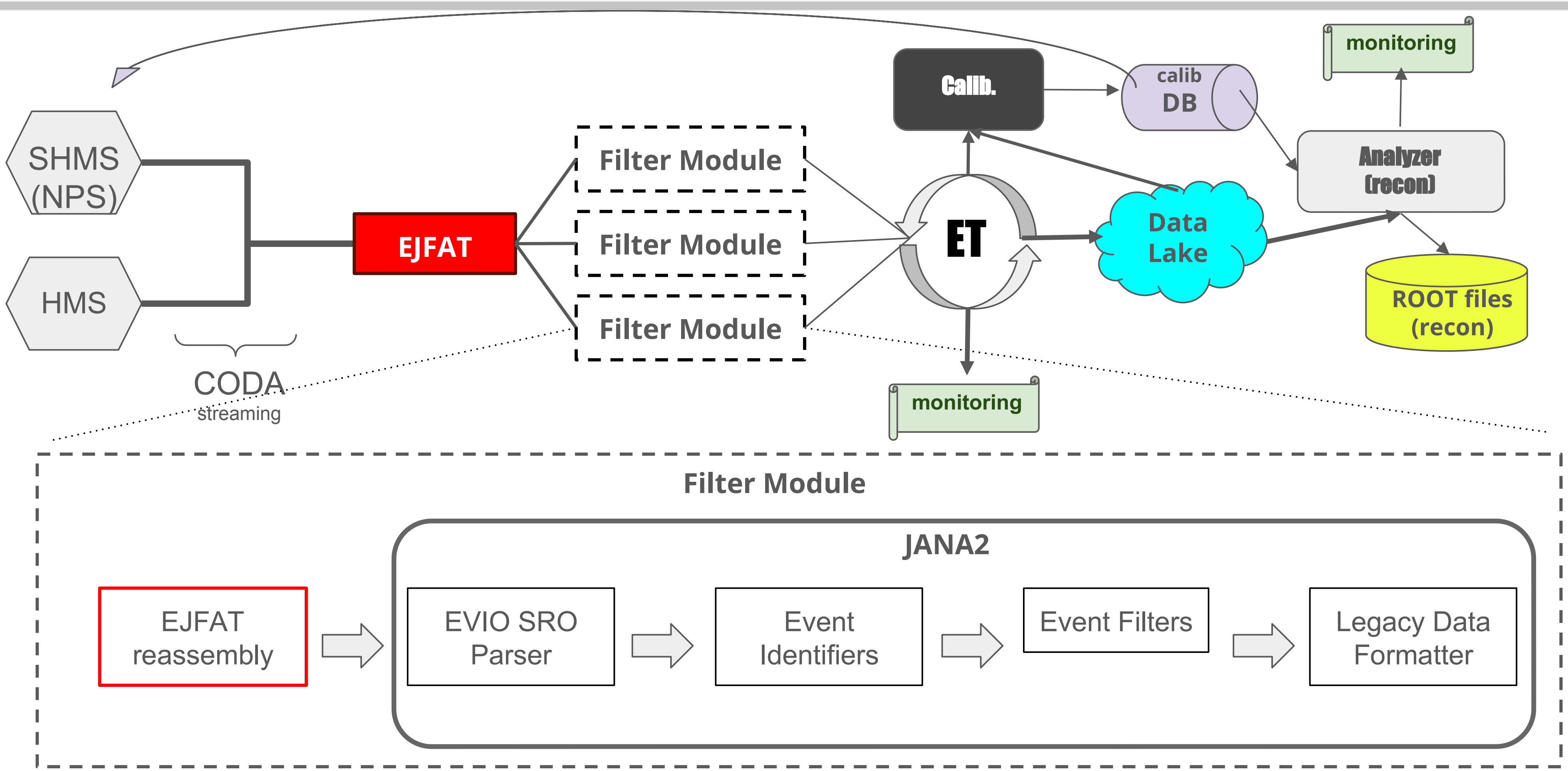
A Proposed Road to SRO in Hall C

- Advantages
  - → Baseline Hall C hardware is already SRO capable
    - » JLab F250s, VTPs, CAEN 1190s, JLab TI/TM, etc
  - → Hall C has well understood 'simple' detector packages
    - » HMS, SHMS
    - » NPS (schedule dependent)
  - → Relatively simple optics, track reconstruction, etc
  - → *Pre-existing, well understood NIM triggers exist and can be a big advantage here!*

NOTE: This dovetails with the Test Setup development mentioned earlier (Hanjie, et al.)

- Challenges
  - → Must have minimal impact on existing Hall C program
  - → Identify collaborators in Hall (both staff and in experimental collab.)
  - → 'hcana/Podd' software is *not* SRO capable
    - » Full analyzer rewrite in new framework difficult/time consuming, requires experts in JANA2 and Hall C systems
    - » Solution:
      - – Generate JANA2 'shim' software to filter SRO stream into legacy Hall C EVIO format on the fly
      - – Extend JANA2 filter layer capabilities in stages while demonstrating different aspects of Grand Challenge along the way

Nominal Hall C / SRO Data Flow



A Proposed Road to SRO in Hall C

- Stage I
  (Have consensus/plan by Fall 2024; allocate people for dev in 2025)
  - → Identify collaborators/support across divisions
    - EPSCI, Hall Staff, DAQ/FE group
  - → Identify necessary changes to Hall C DAQ config to support this SRO model
    - » Can we get away with SW/config change only?
    - » Is existing payload module firmware compatible with our SRO needs?
      - – ie. CAEN1190 needs bench testing for SRO mode?
      - – TI/TM firmware update likely
      - – F250 firmware updates?
      - – VTP availability, firmware
      - – Do we need to swap between firmware modes, etc.
- Stage II (2025 → ...)
  - → Develop the JANA2 shim software to repackage SRO CODA stream into legacy EVIO format for Podd/hcana
    - » This is a component of the Hall C test stand setup underway
    - » Packet-capture some cosmics data in ‘SRO mode’ from HMS, SHMS during a down period
  - → Resurrect ET support in HCana
  - → Demonstrate software path works on the bench
    - » SRO replay → Shim → HCana
- Stage III (2026?)
  - → Collect live data w/ beam in Hall C for a shift adjacent to a pre-scheduled program interruption (beam studies, etc)
    - » HMS/SHMS H(e,e’p) elastics would be ideal, but even single-arm data off carbon foil would be very educational
  - → Packet capture at minimum
    - » could also test “full” SRO pipeline as stretch goal
- Stage IV (2026 → 2027)
  - → Debug and selectively build out SRO capabilities of JANA2 layer
    - » Pick 1 or 2 modest but achievable goals to accomplish in JANA2 layer
      - – *Pre-trigger ID to match NIM equivalents (initial goal)*
      - – *refined* t_0 extraction from hodoscopes
        - » input for Tracking (then vertex recon.)?
      - – Cluster ID, energy from Calorim.
      - – PID from PreSh/Sh and/or Cerenkov
    - » Cross-compare with HCana output at “NIM layer” in realtime, and in aggregate (histos)
  - → Investigate SRO-based real-time detector calibration
    - » Cherenkov gains and/or scintillator MIPs could be simple testbed for this
      - – Manually adjust the HVs and have JANA2 system push them back in real time
- Stage V (Fall 2027)
  - → Take a back-to-back runs in SRO mode and ‘Legacy’ mode in Hall C
  - → Directly compare outputs and evaluate success

JANA2 'shim/filter' Requirements

- Goal 1: Convert SRO data stream to legacy event-structure evio format
  - → Use Hall C NIM pre-triggers in pre-defined F250 channels as bootstrap
    - » NIM pre-trigger defines an event is present, and provides t_0 time reference
    - » Identify hits around the NIM pre-trigger time into 'legacy' ROC/event/bank structure
  - → Attach JANA2 output stream to ER/ET process to:
    - » write data to disk for offline analysis, and
    - » feed hcana and populate histograms in real time via ET layer
- Goal 2: Real-time monitoring data from JANA2 Shim layer
  - → provide selected (software) scalar data extracted from SRO stream to cross check w/ existing Hall C hardware scalar data in real-time
  - → Note: scalar "servers" in Hall C runs in parallel to CODA DAQ
    - » Will not require special JANA2 processing.
    - » Scalar data does not have to be in SRO stream (optional)
- Goal 3: Have JANA2 framework identify events on its own
  - → Can use NIM signals as 'hint', and/or cross check as devel progresses
  - → First 'production' goal is to demonstrate ~1:1 match between NIM and JANA2 event ID
    - » Refine from there...

Thoughts on Application to NPS Program

- Advantages
  - → near-realtime detector monitoring and automated calibration
    - » Detector stability is critical for enhanced analysis downstream!
  - → improved and automated anomaly identification
  - → enhanced online data selection / rejection
    - » can operate at event-level and/or detector level (ie. ROI)
  - → improved signal to background for data on disk
    - » accelerate time to science if you can minimize trash and offline (re-)calibrations
  - → AI/ML algorithms can be extremely fast and accurate – an excellent match for near-realtime applications
- Do *not* have to go “all in” with full SRO to start
  - → ie. could employ a loose trigger and implement a “Level-2” farm with access to complete detector information
- Note: SRO will be the standard for the next generation of experiments at JLab, EIC and elsewhere
  - → We must develop experience and expertise within our Collaborations and community to take advantage

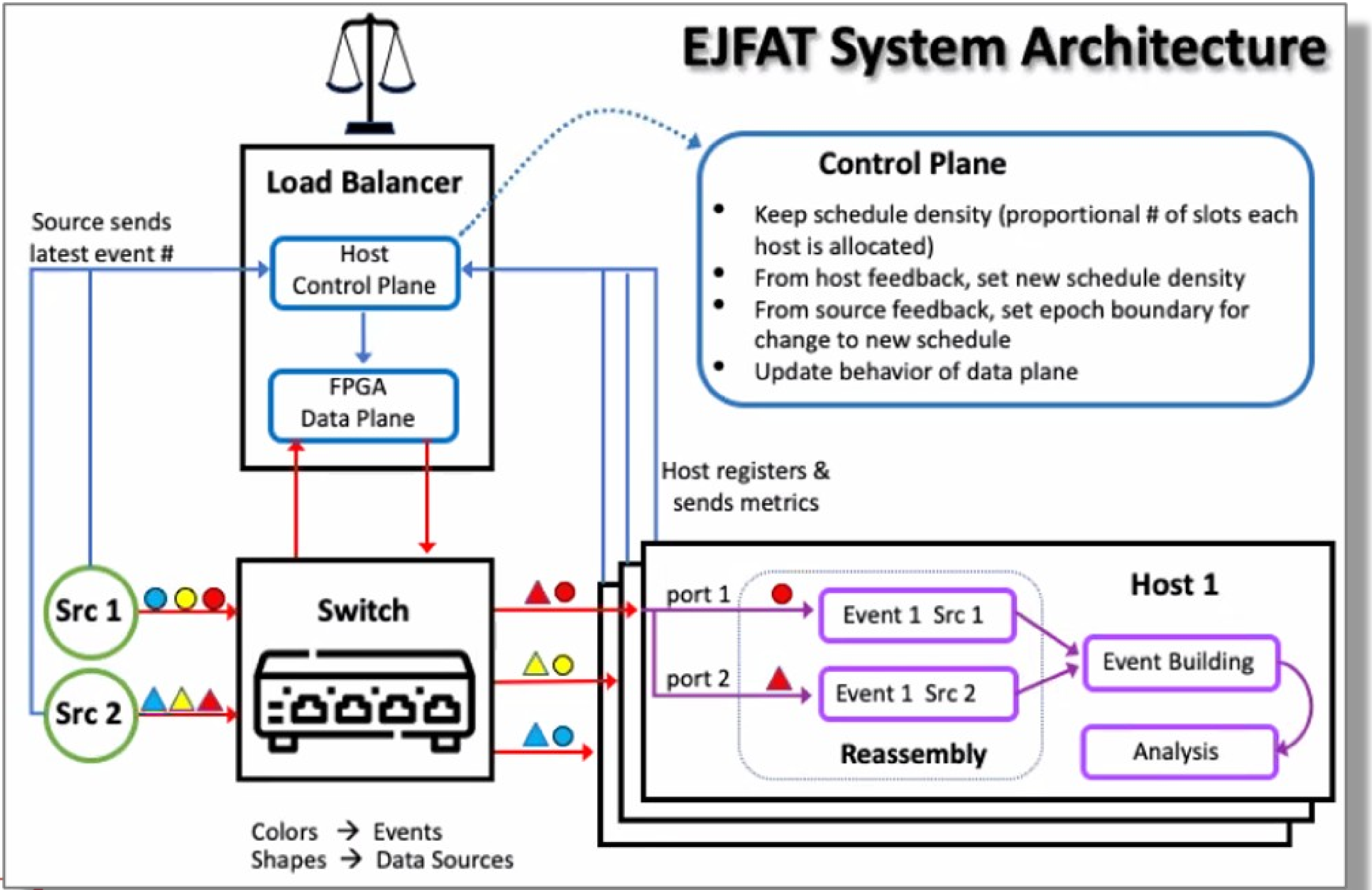
- Challenges
  - → A lot of work to do before any SRO system is production ready
    - » Many pieces still need to be merged into a coherent system
  - → Detector design, readout, and control hardware must be developed with SRO considerations in mind
    - » Hall C, NPS is compatible as-is
    - » “Understood” detector system is a big plus
  - → Suitable analysis software and front-end firmware need to be developed
    - » hcana/podd is not a ideal match for this or future SRO needs
      - – “hybrid” work-around?
        JANA2 filter → hcana
  - → AI/ML algorithms are often challenging to train and validate
  - → Compute and network resources need to be assessed well in advance

Summary / Take-aways

- SRO is the future of DAQ and experimental programs across all fields
  - → Advantages far outweigh the negatives
  - → We are (correctly!) cautious and move slowly, *but* we ignore modern tools and techniques at our peril. ("Better the devil we know…" is not wrong, but it is no limiting us.
  - → Like all new approaches, there will be pitfalls and problems to overcome before the advantages will shine.
    - » To take full advantage of the SRO / Integrated experiment workflow model requires detectors to be designed with monitoring and realtime feedback mechanisms as a goal, not an afterthought.
  - → *We need to get rolling on this process, develop the software, hardware, and hands-on expertise/experience to validate this new approach.*
  - → JLab/Hall C is a *uniquely* well suited environment for this program

EPSCI / FEDAQ Projects

ESnet/JLab FPGA Accelerated Transport



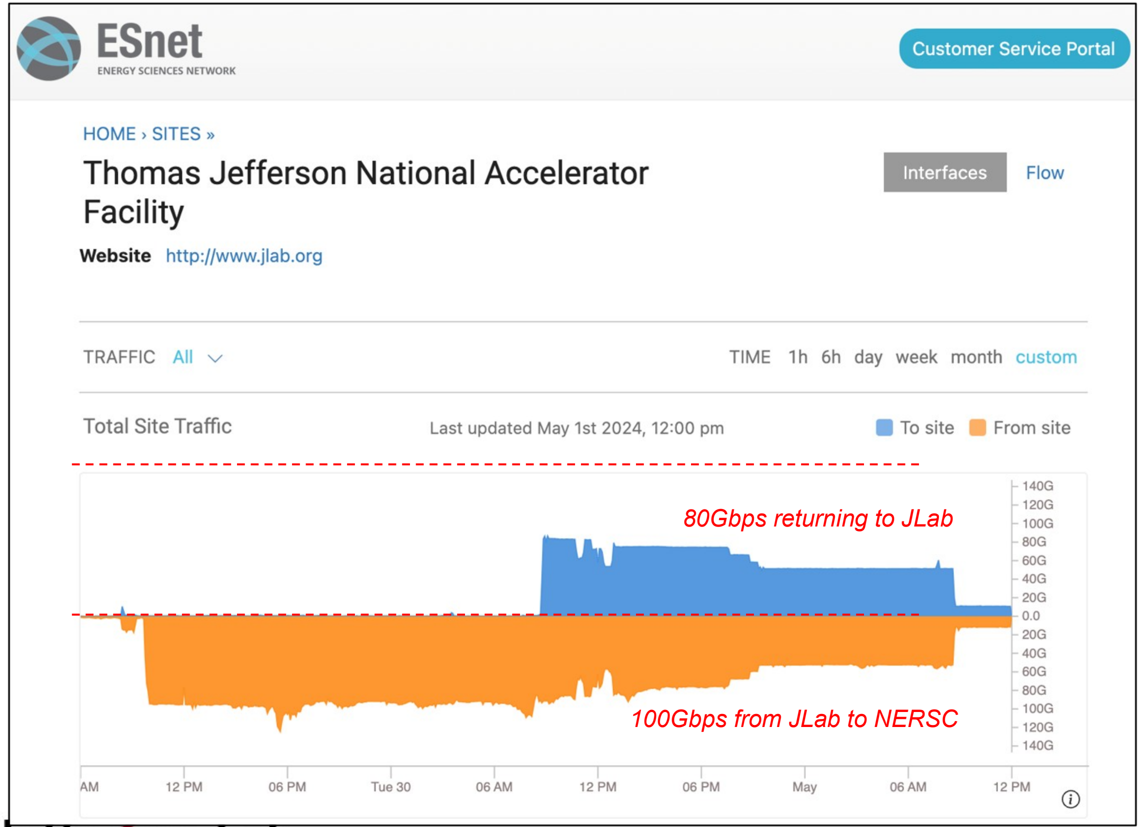
GCII: JIRIAF + EJFAT + NERSC Testing

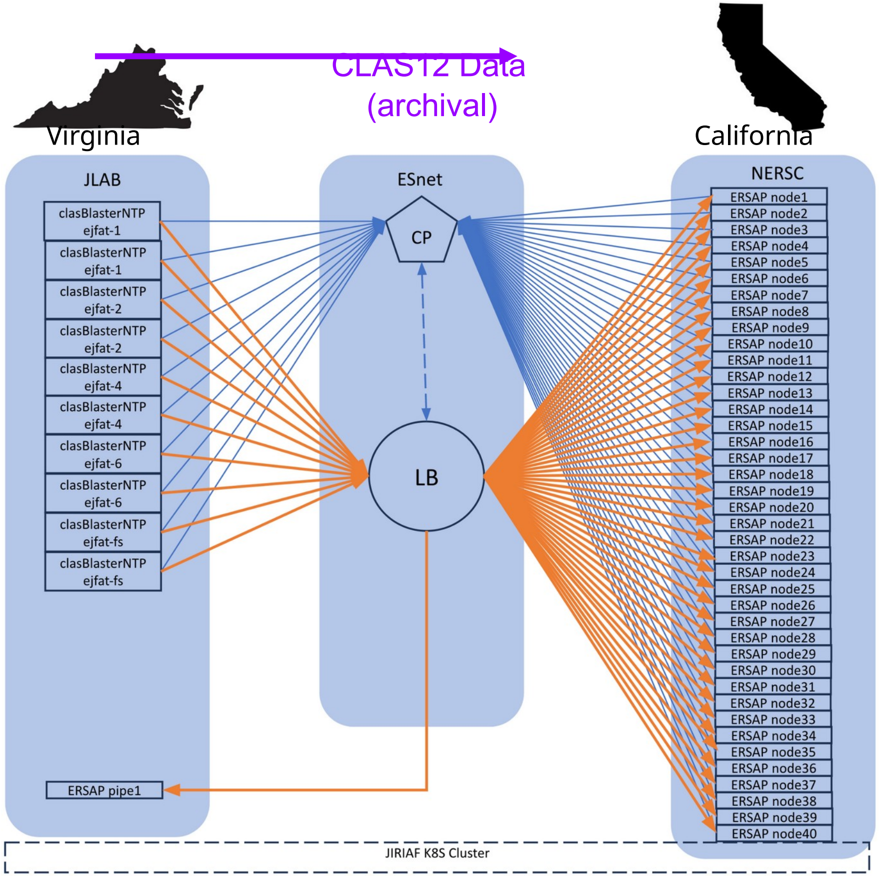
PI: Vardan Gyurjyan
LDRD Project
(funding FY23, FY24)

PI: Graham Heyes
ASCR Project
Project Lead: Michael Goodrich

Experimental Physics Software and Computing Infrastructure

*Stream exercise sending CLAS12 data from JLab to NERSC using the **JIRIAF** system*

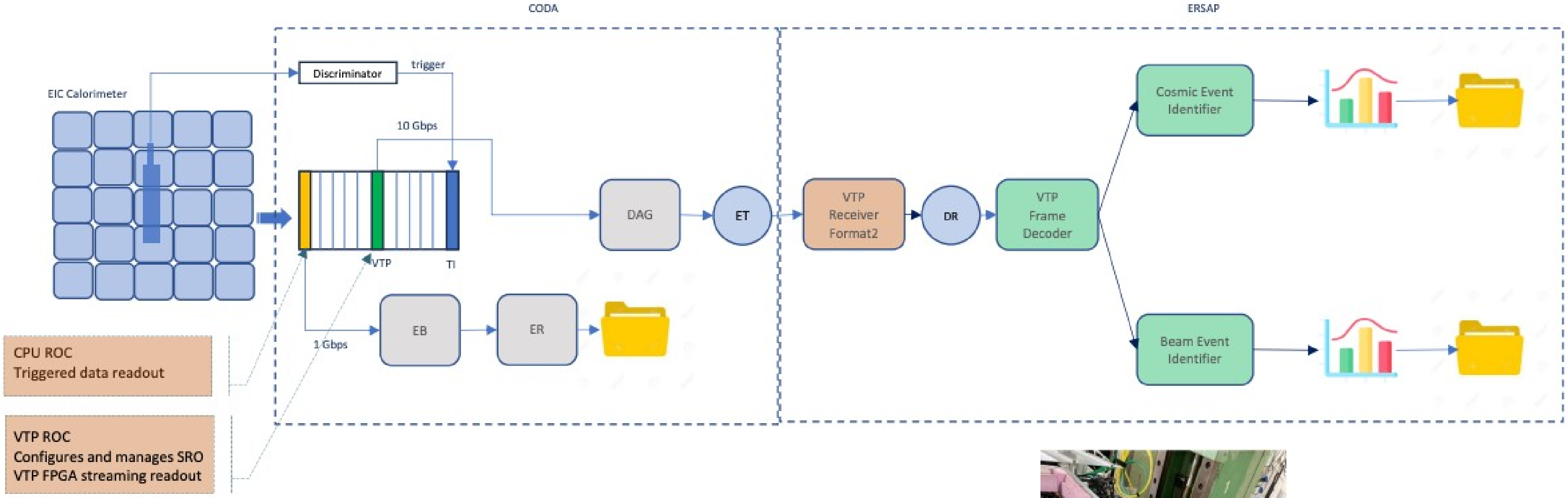




ERSAP - **E**nvironment for **R**eal-time **S**treaming, **A**cquisition and **P**rocessing

EPSCI
Experimental Physics Software and Computing Infrastructure

EIC prototype calorimeter SRO pipeline at DESY. CODA & ERSAP



ERSAP design is *event reactive actors*, networked by data pipelines.

- Compositional actors with conditional data routing at runtime.
- Flow-based programming paradigm

Triggered data are waveforms read out over the VME bus.
Stream data are integrated sums and times of all hits over a threshold in the calorimeter regardless of the trigger status.

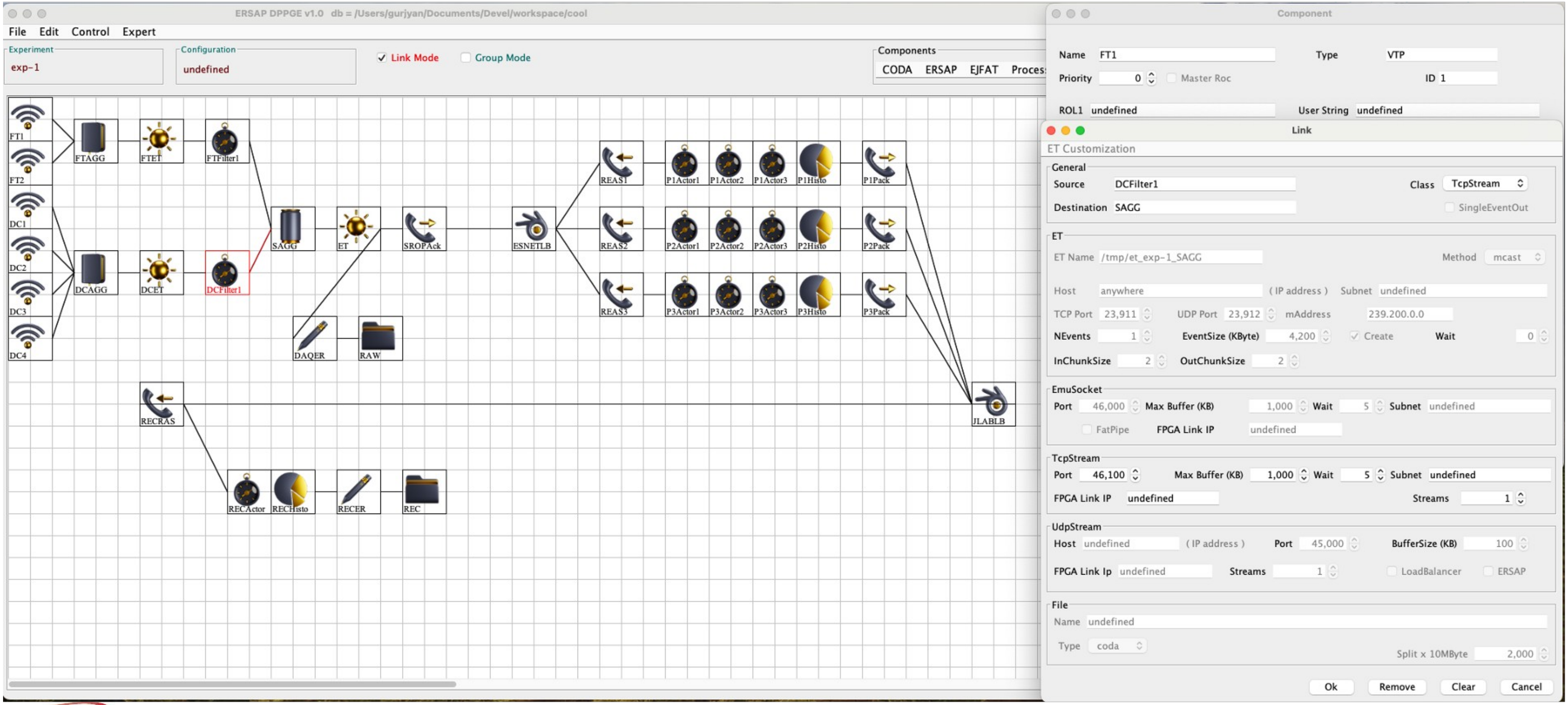
recent beam test at DESY included ERSAP environment for online processing

Funded through JLab SciOPS

GCII: DPPGE + DPOL (SRO Replacement for JCEdit + COOL)

Graphical Configuration of multi-process, distributed system

Vardan Gyurjyan



GC II Development Milestones
(*Very* Draft/Tentative)

SRO TestBed Milestones

- Step 1
  - → Identify / build testbed DAQ stand (Hanjie + Dmitri)
    - » Pre-existing setup should be compatible with the base system and testbed requirements
    - » Ensure hardware components are available for use on our timeline (ideally, HW should not be committed to other projects)
    - » Identify other external users or demands on the system and hardware?
- Step 2:
  - → Build / identify an appropriate detector (with cabling, HV, etc) to read out
    - » Simple is better in the early stages
    - » cosmics telescope, small PbG, PbWO4 array, etc.
- Step 3:
  - → Establish a working DAQ and characterize the system with a well understood software framework and CODA config
    - » “Hall C” setup; hcana
    - » fix hcana ET input as part of this?
    - » Want the JANA2 workflow ASAP (Dmitri has this in-hand?)

SRO TestBed Milestones

- Step 4a
  - → Develop an SRO workflow compatible with the testbed framework
    - » EJFAT/ERSAP/JANA2 filter modules?
      - – See also: Analyzer/Visualizer requirements
    - » Ideally (ultimately) this framework would find events and repackage into a 'HallC' event stream for further decoding and visualization?
    - » We should require/incentivize getting ET input working for hcana/podd again!
      - – What are intermediate steps?
        - » What already exists for SRO replay/visualization?
- Step 4b:
  - → Develop "streaming firmware" for modules
    - » what will need updating; is the firmware already available?
    - » will/can this firmware support "standard" readout modes too, or only SRO?

SRO Testbed Milestones

- Step 5:
  - → Deploy SRO framework in testbed environment and begin testing
    - » Should have or develop a set of “low-level” synthetic tests that can be run to verify operation of the firmware and software components.
    - » We need to think about how we will test/debug the individual components and block out some plans for this step.
    - » What does ‘good’ data streaming off the payload modules look like?
      - – Do we use netcat/wireshark and packet capture, or are there ‘simple’ decoders that can ID module headers, timestamps, etc.?
      - – (I often look at the evio2xml dump of a ‘raw’ CODA file in debugging stages – how do we do that with SRO?)
    - » What are similar tests to validate the EJFAT layer; JANA2 filter layer, etc?

SRO Testbed Milestones

- Step 6a:
  - →Plug in the testbed detector and reproduce what we saw in 'conventional' mode
    - » ADC histos, TDC histos, scaler/event rate comparisons
    - » 'cosmic yields' after cuts (ie. simple physics cross checks)
    - » Sanity check NIM trigger vs. JANA2 event ID
- Step 6b:
  - →Load / Stress testing
    - » Push the end-to-end system performance by adding modules with noise (ie. low thresholds)
    - » Use sources to generate 'real' backgrounds (detector dependent)
  - →Build-out and test in prep for a in-beam test of this system

SRO Testbed Milestones

- Step “7” :
  - → Debug, iterate, and stabilize
  - → Could start playing with Calibration feedback here (detector dependent)
    - » ie. simple gain-matching feedback on cosmics in a multi-channel hodoscope or calorimeter array could be a fun testbed / development environment
    - »

SRO Full-Scale Milestones

- Steps 8 + :
  - → Select a production Hall system for full-scale test (ie. Hall C?)
    - » (Re-)Evaluate required network, disk, CPU requirements for selected system. Can we deploy within the selected Hall's existing compute cluster, or are Farm resources required?
  - → Schedule some dedicated beam time in selected Hall (at least 1–2 calendar days; ideally separated by a gap)
    - » Should be done at least 1-year in advance
      - – Needs buy-in from Thia, Hall leadership, and Collaboration on the floor
    - » Will need to clearly document and test roll-out and back-out procedures
    - » Will require a long down (summer) for deployment and testing in the Hall environment
  - → Ideally do back-to-back tests on production system and compare results
    - » Run a luminosity (beam current) scan and look at deadtimes, absolute yields, etc
    - » Scheduling such that we can take advantage of an elastics setup or some other over-constrained physics config would be advantageous