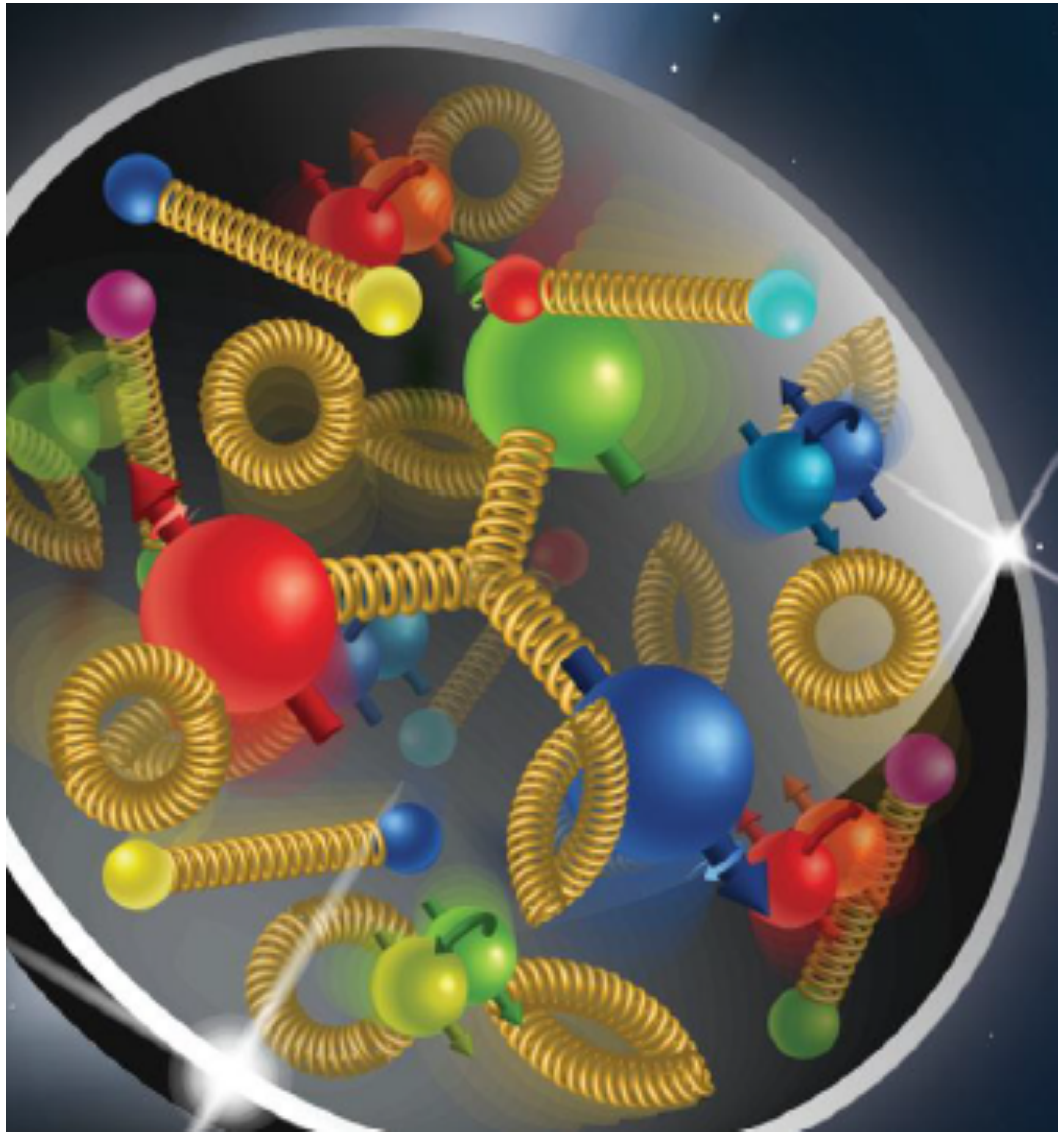
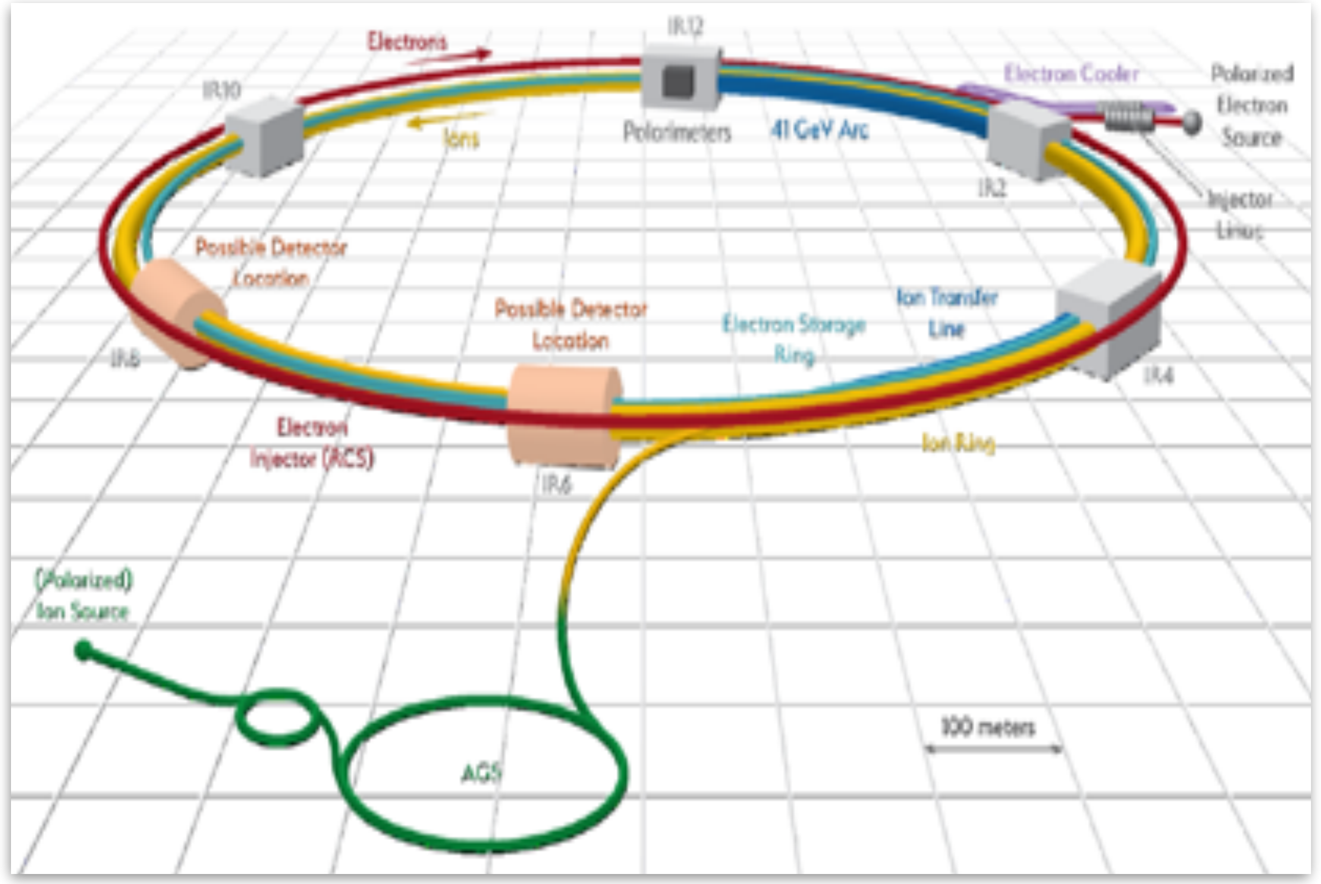
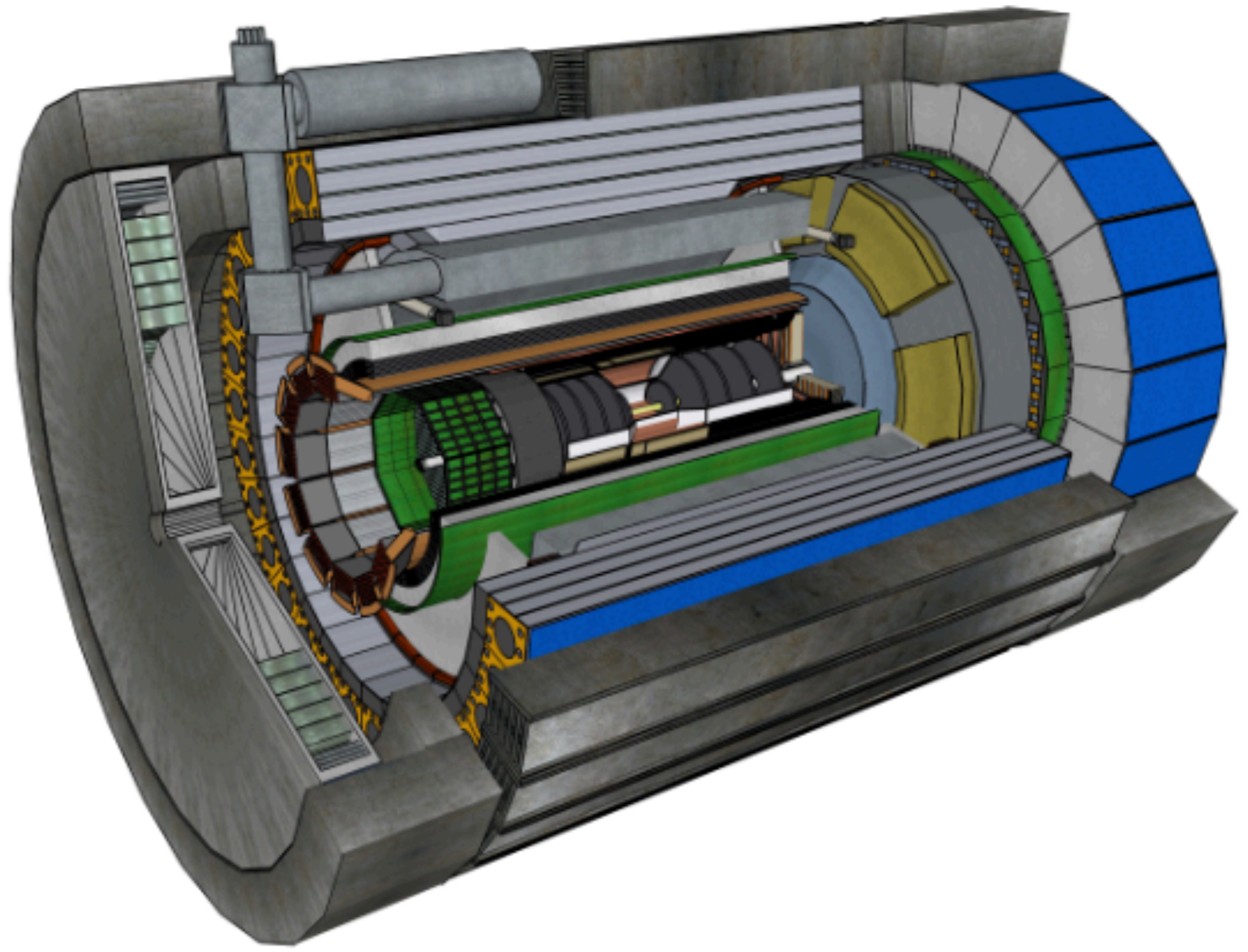





January 2024 ePIC Collaboration Meeting

9–13 Jan 2024

Argonne NATIONAL LABORATORY

# *WG Overview*
# *ePIC Streaming Computing Model*

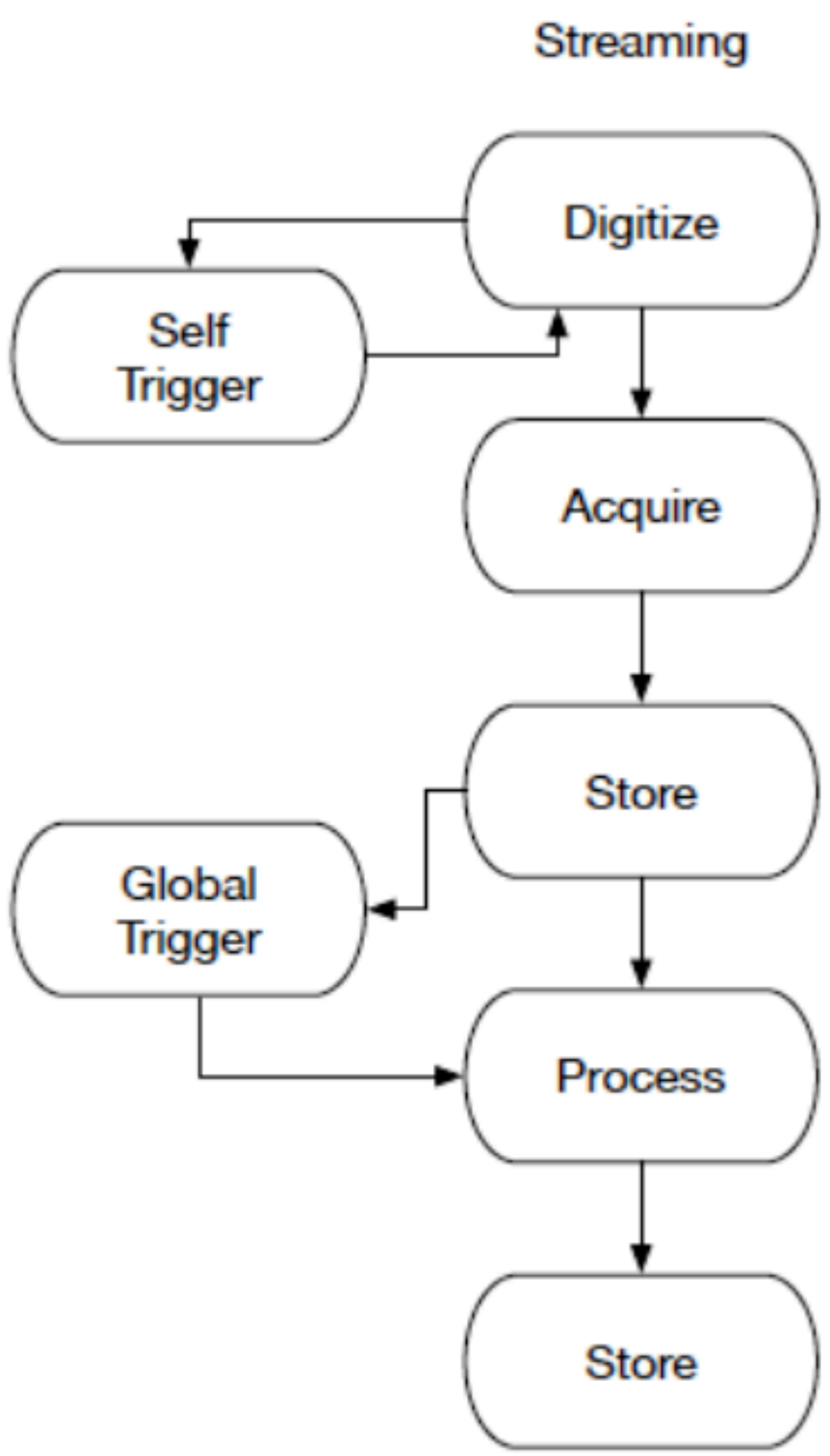
M.Battaglieri (INFN), M.Diefenthaler (JLab), J.Huang (BNL), J.Landgraf (BNL), T.Wenaus (BNL)

ePIC Brookhaven National Laboratory Jefferson Lab e-JLab12 INFN

# Streaming RO

## Streaming read out (SRO)

✳ A HIT MANAGER receives hits from FEE, order them and ship to the software defined trigger




✳ All channels continuously measured and hits streamed to a HIT manager (minimal local processing) with a time-stamp

✳ Software defined trigger re-aligns in time the whole detector hits applying a selection algorithm to the time-slice

- the concept of 'event' is lost
- time-stamp is provided by a synchronous common clock distributed to each FEE

**SRO DAQ**

- **Pros**
  - All channels can be part of the trigger
  - Sophisticated tagging/filtering algorithms
  - high-level programming languages
  - scalability
- **Drawbacks:**
  - we do not have the same experience as for TRIGGERED DAQ

## Why SRO is important?

✳ **High luminosity experiments**
- Write out the full DAQ bandwidth
- Reduce stored data size in a smart way (reducing time for off-line processing)

✳ **Shifting data tagging/filtering from the front-end (hw) to the back-end (sw)**
- Optimize real-time rare/exclusive channel selection
- Use of high-level programming languages
- Use of existing/ad-hoc CPU/GPU farms
- Use of available AI/ML tools
- (future) use of quantum-computing

✳ **Scaling**
- Easier to add new detectors in the DAQ pipeline
- Easier to scale
- Easier to upgrade

**Many NP and HEP experiments adopt a SRO DAQ**

- CERN: LHCb, ALICE, AMBER
- FAIR: CBM
- DESY: TPEX
- FRIBS: GRETA
- BNL: sPHENIX
- JLAB: SOLID, BDX, CLAS12, …

# ePIC Streaming Computing Model Working Group

## Compute-Detector Integration to Maximize Science

**Broad ePIC Science Program:**
- Plethora of observables, with less distinct topologies where every event is significant.
- High-precision measurements: Reducing systematic uncertainties of paramount importance.

**Streaming Readout Capability Due to Moderate Signal Rate:**
- **Capture every collision signal**, including background.
- Event selection using all available detector data for **holistic reconstruction**:
  - **Eliminate trigger bias** and provide accurate estimation of uncertainties during event selection.
- Streaming background estimates ideal to **reduce background** and related systematic uncertainties.

| | EIC | RHIC | LHC → HL-LHC |
|---|---|---|---|
| Collision species | $\vec{e}+\vec{p}, \vec{e}+A$ | $\vec{p}+\vec{p}/A, A+A$ | $p+p/A, A+A$ |
| Top x-N C.M. energy | 140 GeV | 510 GeV | 13 TeV |
| Peak x-N luminosity | $10^{34}$ cm$^{-2}$ s$^{-1}$ | $10^{32}$ cm$^{-2}$ s$^{-1}$ | $10^{34}$ → $10^{35}$ cm$^{-2}$ s$^{-1}$ |
| **x-N cross section** | **50 µb** | **40 mb** | **80 mb** |
| Top collision rate | 500 kHz | 10 MHz | 1-6 GHz |
| $dN_{ch}/d\eta$ | 0.1-Few | ~3 | ~6 |
| **Charged particle rate** | **4M $N_{ch}$/s** | **60M $N_{ch}$/s** | **30G+ $N_{ch}$/s** |

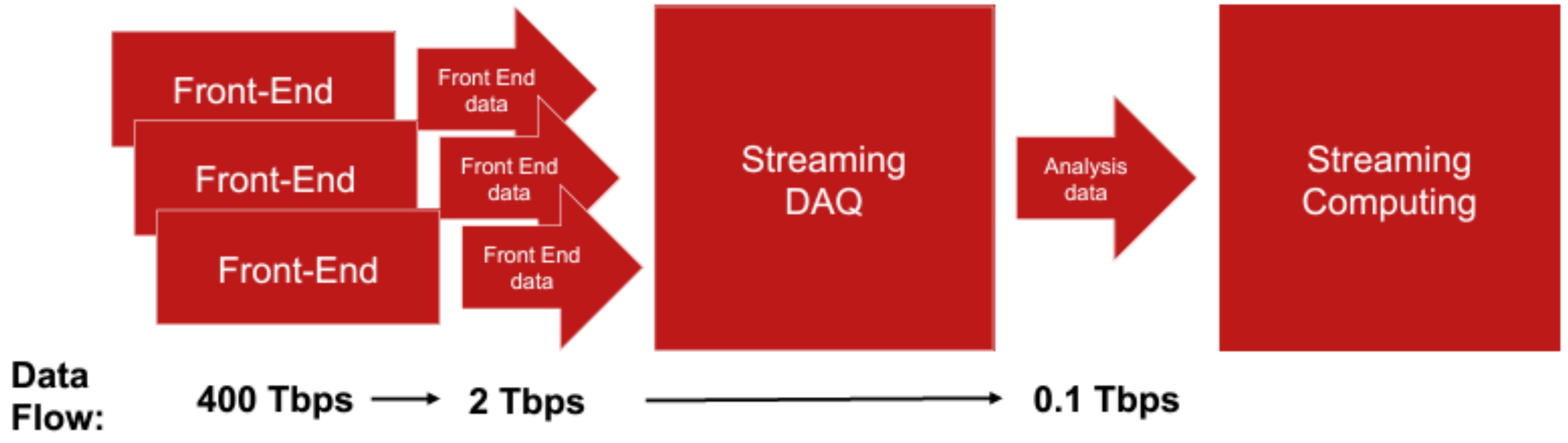
Streaming Readout XI, December 2, 2023. 4 ePIC

## Why SRO is important for ePIC?

✳ **Broad ePIC Science program**
- Unbiased data set for unplanned physics
- Selective (św) triggers to focus on for planned physics

## Compute-Detector Integration to Accelerate Science

- **Problem** Data for physics analyses and the resulting publications available after O(1year) due to complexity of NP experiments (and their organization).
  - Alignment and calibration of detector as well as reconstruction and validation of events time-consuming.
- **Goal Rapid turnaround of 2-3 weeks for data for physics analyses.**
- **Solution** Compute-detector integration using:
  - AI for autonomous alignment and calibration as well as reconstruction and validation for rapid processing,
  - Streaming readout for continuous data flow of the full detector information,
  - Heterogeneous computing for acceleration.

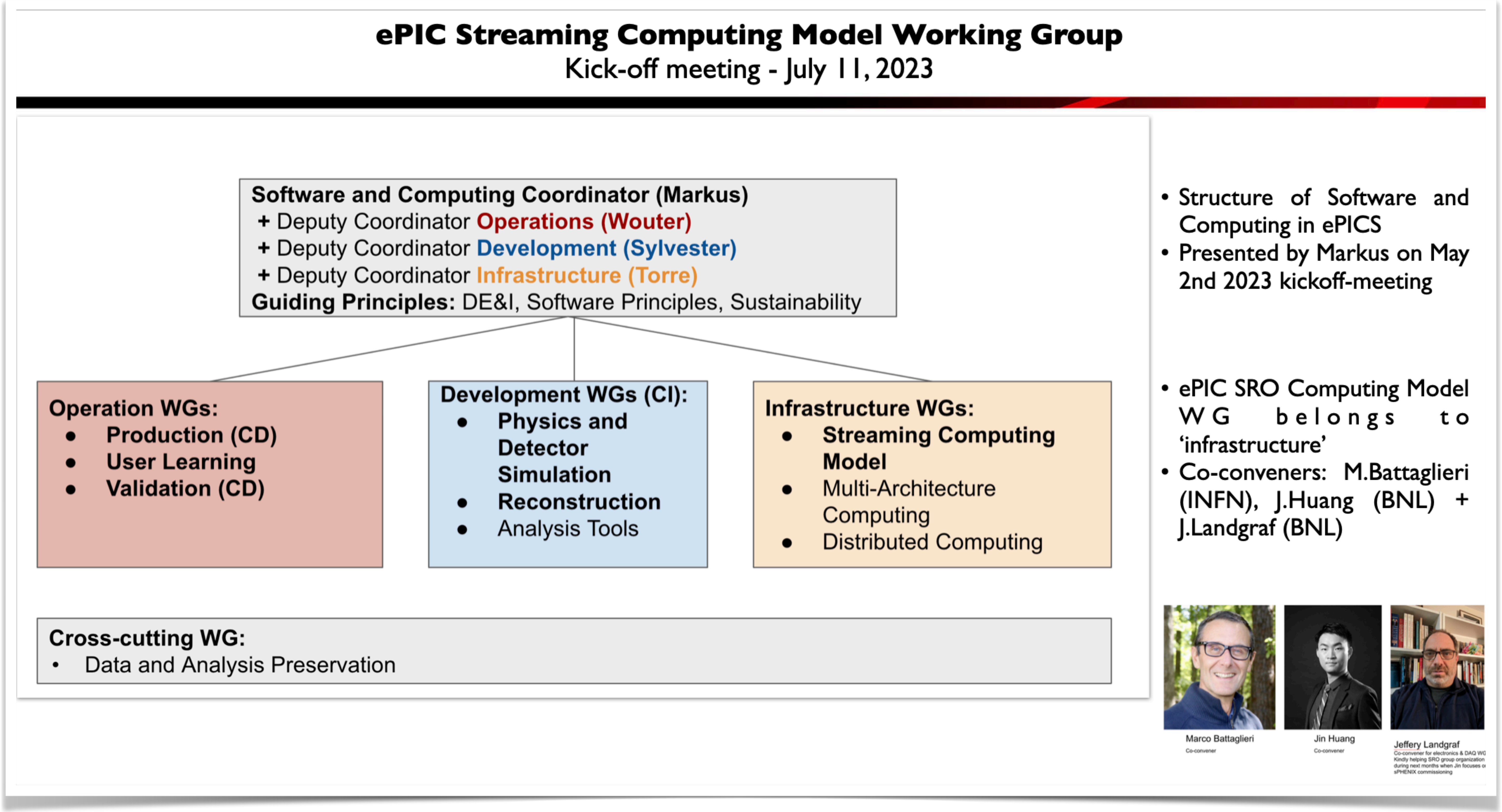



Streaming Readout XI, December 2, 2023. 5 ePIC

✳ **Expedite ePIC Science outcome**
- Computing/detector integration for a rapid turnaround

Presented by Markus at APS/SRO-XI in Kona (HI) in December

# ePIC Streaming Computing Model Working Group




## WG activity

- WG activity started in July 2023, regular Zoom (bi-and) weekly meetings (Tue 9:00 AM ET) + dedicated 1-2 days workshops)
- Each meeting focused on a specific topic
- Invited talks of world-experts
- … a lot of discussion

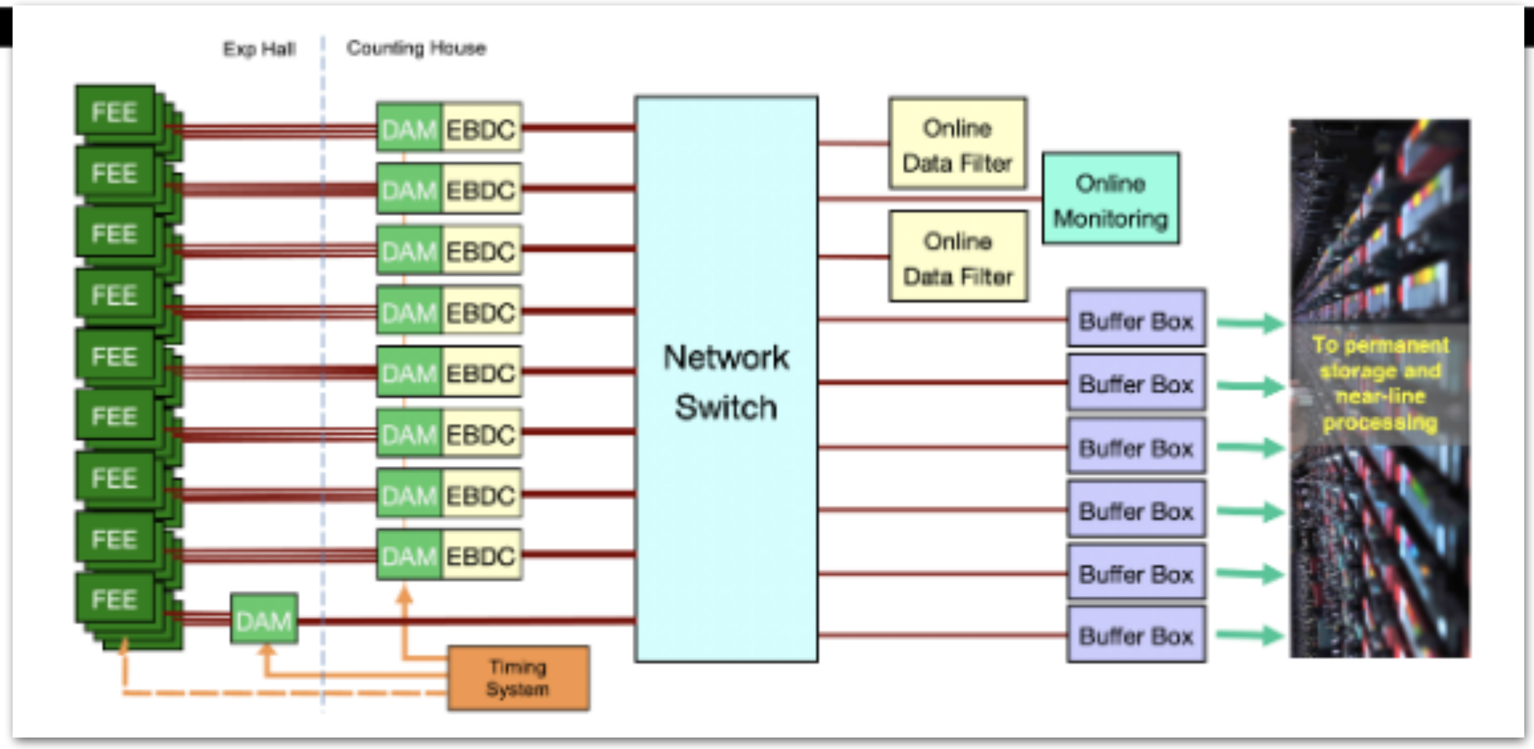
Items discussed

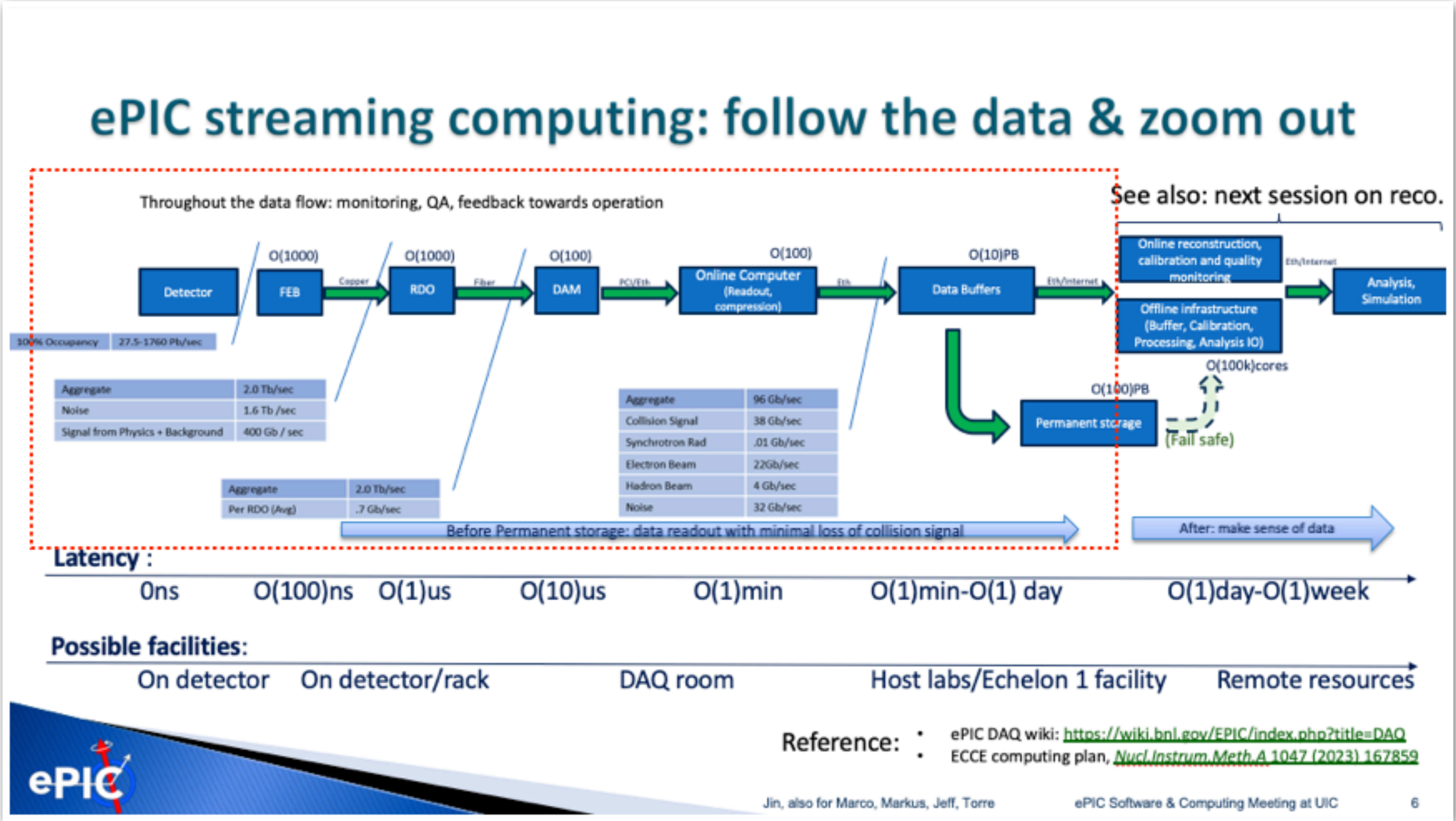
- Streaming RO computing model
- ePIC Data Rates
- The interface between DAQ and computing
- ePIC butterfly computing model (Echelon 0-3)
- Algorithmic workflow data analysis requirements
- (Autonomous) Calibrations
- (raw) data filtering
- Data format, data cooking, …

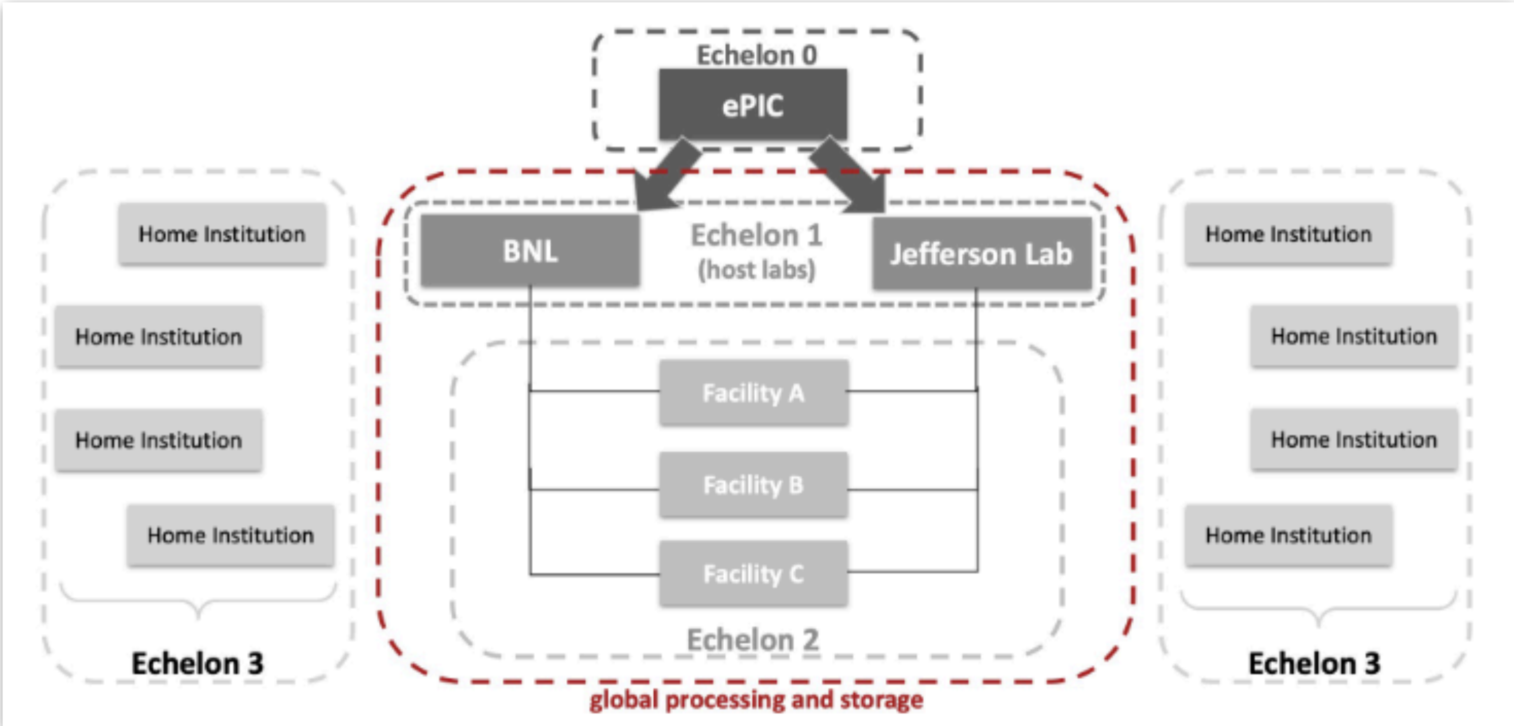
Achievements

- ePIC computing review (Oct 19-20 23)
- *The ePIC Streaming Computing Model* paper

# ePIC Streaming Computing Model Working Group




## Streaming RO for ePICS

- Full consensus for SRO within the EIC community (Yellow Paper, ECCE, ATHENA, …)
- ePICS rates are not comparable to LHC HI-LUMI but the advantages of SRO remain
- Holy Grail: to manage (storage) an unbiased (un-triggered) data set for further analysis

### ePICS rates

- Established a strong link with DAQ WG
- Define an envelope for the workflow (<100/400Gb/s)
- First estimate of ePICS data rate/volume to be updated based on new information
- Provide feedback to sub-systems groups to reduce the data stream to a compatible level

### Interfaces

Within the 'control room' - Project driven (boundary based on permanent data storage)

- Each stage in data flow will require IO specs (based on CPU, GPU, FPGA reduction)

Outside the control room - Collaboration driven

- Networking? CPU/GPU farm? Local/remote resources? on/off-line analysis?

### Streaming Data processing

- Time-critical, data-driven, autonomous, distributed
- Four tiers of the ePIC Streaming Computing Model
  - Echelon 0: ePIC experiment and SRO
  - Echelon 1: shared between BNL and JLab
  - Echelon 2: global contributions
  - Echelon 3: analysis community

# ePIC Streaming Computing Model Working Group

## The ePIC Streaming Computing Model

Marco Battaglieri[1], Wouter Deconinck[2], Markus Diefenthaler[3], Jin Huang[4], Sylvester Joosten[5], Jeffery Landgraf[4], David Lawrence[3] and Torre Wenaus[4]
for the ePIC Collaboration

[1]Istituto Nazionale di Fisica Nucleare - Sezione di Genova, Genova, Liguria, Italy.
[2]University of Manitoba, Winnipeg, Manitoba, Canada.
[3]Jefferson Lab, Newport News, VA, USA.
[4]Brookhaven National Laboratory, Upton, NY, USA.
[5]Argonne National Laboratory, Lemont, IL, USA.

**Abstract**

This document provides a current view of the ePIC Streaming Computing Model. With datataking a decade in the future, the majority of the content should be seen largely as a proposed plan. The primary drivers for the document at this time are to establish a common understanding within the ePIC Collaboration on the streaming computing model, to provide input to the October 2023 ePIC Software & Computing review, and to the December 2023 EIC Resource Review Board meeting. The material should be regarded as a snapshot of an evolving document.

### Contents



## ePIC Software &Computing review (October 19-20, 2023)

### Review charge

1. At this stage, approximately ten years before data collection, is there a comprehensive and cost-effective long-term plan for the software and the computing of the experiment?

2. Are the plans for integrating international partners' contributions adequate at this stage of the project?

3. Are the plans for software and computing integrated with the HEP/NP community developments, especially given data taking in ten years?

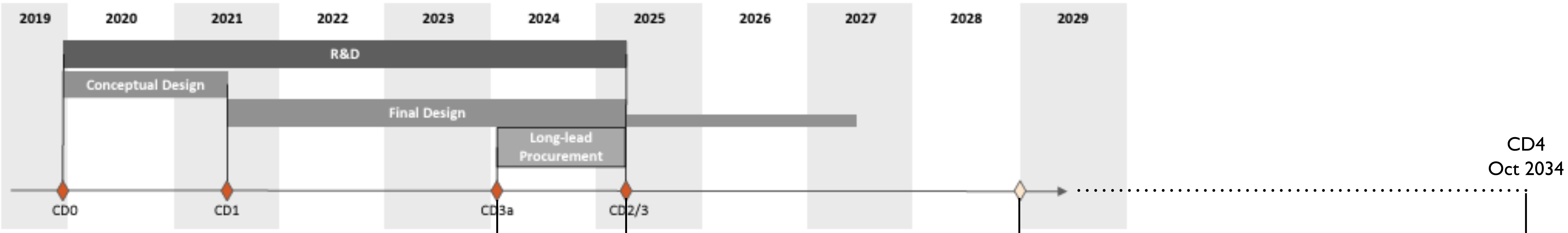
4. Are the resources for software and computing sufficient to deliver the detector conceptional and technical design reports?

5. Are the ECSJI plans to integrate into the software and computing plans of the experiment sufficient?

# Summary and outlook

- ePIC DAQ will be streaming
- SRO is expected to facilitate (and extend!) the ePIC science and shorten the time between data taking and physics output
- SRO provides new opportunities but poses challenges (e.g. new detector-computing relationship)
- The SCM WG is working to define a suitable computing model (based on ePIC Coll requirements) for SRO
- Strong link with DAQ, SW, and Physic WGs to design a sound framework (+ trigger/analysis algorithms)
- Interaction with ePICS sub-detector groups will be the next step for an optimized SRO DAQ design
- On-field validation needs to be pursued in parallel
- It is a long way to go but we are progressing!

## ePIC Milestones




**TDR**
- SW and sim ready for TDR
- Computing resources estimate

**Detector construction**
- SRO DAQ test prototype
- Streaming tests
- Data and analysis challenges

**Detector Installation and commissioning**
- Triggered and SRO in parallel
- calibrations
- data reconstruction (Ech1, Ech2)

**ePIC is running!**
- SCM validation