ePIC Software & Computing Overview

Torre Wenaus (BNL) for the ePIC Software and Computing Team
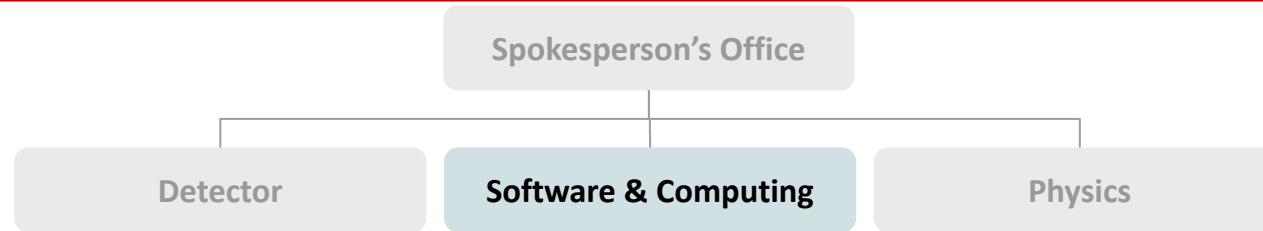ePIC S&C Review
September 26-27 2024
Catholic University of America, Washington DC

ePIC S&C Organization

Spokesperson's Office

The mix of active and *not yet activated* activities remains the same as last year

Since last year Dmitry Kalinkin has joined us as Deputy Coordinator for Development, taking over from Sylvester Joosten

| Detector | **Software & Computing** | Physics |

Guiding Principles:

- *Diversity, Equity, and Inclusion*
- *Statement of Software Principles*
- *Sustainability*

Software and Computing Coordinator
Markus Diefenthaler (Jefferson Lab)

Cross-cutting Working Group:

- *Data and Analysis Preservation*

Deputy Coordinator (Operations)
Wouter Deconinck (U. Manitoba)

Operation Working Groups:

- Production
- User Learning
- Validation

Deputy Coordinator (Development)
Dmitry Kalinkin (U Kentucky)

Development Working Groups:

- Physics and Detector Simulation
- Reconstruction Framework and Algorithms
- *Analysis Tools*

Deputy Coordinator (Infrastructure)
Torre Wenaus (BNL)

Infrastructure Working Groups:

- Streaming Computing Model
- *Multi-Architecture Computing*
- *Distributed Computing*

+ cross-cutting task force efforts (across *all* ePIC Working Groups)

T. Wenaus   ePIC S&

BROOKHAVEN
NATIONAL LABORATORY

ePIC S&C Working Groups - What's changed

Development:

- **Development coordination:** Dmitry Kalinkin (University of Kentucky), an expert ePIC software developer and code reviewer, took over as ePIC Software & Computing Deputy Coordinator for Development.
    - Sylvester Joosten (ANL) left the role when he became detector subsystem leader for the Barrel Imaging Calorimeter (BIC).
- **Simulation:** The current priority of the Physics and Detector Simulation WG is co-design between engineering and simulation efforts, specifically addressing **synchronization of models between engineering (CAD) and detector simulations** (Geant4, DD4hep).
- **Reconstruction:** Priorities are now coordinated with the Analysis Coordinators (ACs) through **weekly joint AC/SCC meetings**, ensuring that reco priorities are consistent with those of the Physics WGs. Key achievements include rapid progress on **studies for recent early science workshop** using ePIC Software and simulation campaigns.

Operations:

- **User learning coordination:** Stephen Kay (University of York) became convener of the User Learning WG, and has been pivotal in continuing and improving the tutorial series.
    - Kolja Kauder (BNL) left the role to focus on simulation including a new EIC simulation R&D project at BNL.

Comments on 2023 review recommendations/findings/comments

- **Recommendation: that ECSJI verify the readiness of simulation and reconstruction for the TDR by May 2024**
    - As Holly described, the TDR has been moved out a year, to end 2025. A pre-TDR still targets end 2024 and simu & reco are meeting the needs of pre-TDR preparations.
    - The key milestone for TDR readiness we discussed a year ago, time frame based track reconstruction, has been met.
    - Dmitry will discuss the TDR software and simulation readiness
- **Recommendation: that ePIC document a first computing needs assessment by the next ECSAC review, in roughly one year.**
    - We provide a first assessment in this review, in Markus' computing model talk.
- **Comment: To ensure sustainability, ePIC should consider contributing to the development and support of the key tools. The work done in DD4hep goes in this direction and we encourage ePIC to take the same approach with other tools.**
    - We agree! Recent developments and (effort-dependent) plans are presented by Dmitry.
- **Recommendation: that the ePIC collaboration start by the time of the next year's review an evolving list of software dependencies that includes the packages, primary supporters, and ePIC contributions**
    - Dmitry will present this.

Comments on 2023 review recommendations/findings/comments - 2

- **Comment on sufficiency of resources: due to the opportunistic nature of the resources, there are uncertainties which could result in a shortfall in time of need.**
    - Opportunistic resources continue to provide most of the processing resources (80% from OSG, as last year). This has continued to be sustainable over the last year. Of course the uncertainty persists. DOE does not yet provide funding for EIC/ePIC computing resources.
    - JLab provides some processing resources
    - For storage, both JLab and BNL have been providing resources, limited but sufficient for the ePIC production program.
        - Storage use has recently consolidated around a new XRootD service at JLab, and a Rucio instance at JLab is being commissioned.
- **Finding: The sharing of the responsibilities between ePIC and ECSJI, particularly on the software side, are still being discussed.**
    - It's been a year of positive developments.
    - We appreciate the partnership with Amber and Alexei and their collaborative spirit and support.
    - The software (NPPS, EPSCI) and computing groups (SDCC, Scientific Computing Operations) at the various labs are strongly involved.
    - We appreciate ECSJI's role in oversight for the purposes of accountability to the host labs and DOE.

Selected progress highlights, detailed in subsequent talks

- Simulation
  - A first version of **time frame building** has been incorporated
  - Greater **realism in noise and background**
  - Realistic **modelling of services** in the geometry and materials budget
- Reconstruction
  - **Time frame based reconstruction** implemented
- Tracking based studies
  - The first application of time frame based reco
- <u>Validation image browser</u>
  - Browser based viewer of validation plots gathered in a DB with production metadata
- Framework
  - support for **time frame processing** workflow - frames in, events out
  - onnxruntime integration in EICrecon to support use of **ML inference** in algorithms
- Event display
  - Browser based <u>event display</u> using HSF's <u>Phoenix</u>, for detector/event visualization, debugging of simu geometry and reco

Selected progress highlights, detailed in subsequent talks - 2

- [Production](#) system and operations
  - **OSG based production** continues to be tremendously effective, 1M core-hrs/mo
  - **Monthly production** delivers needed data and helps motivate and drive the software
  - **Expanding internationally** via OSG (Canada, Italy)
  - Benefitting from **distributed computing decisions** (Rucio, XRootD)
- [Website landing page](#) guides users to info and help
  - Onboarding info, ePIC tutorials, HSF training, FAQ
  - Regular tutorials (Jan, Apr, Sep so far this year), typical attendance ~20-45, recorded and available online. April was at CERN, serving different time zones.
  - Also a helpdesk on Mattermost
- Conditions database (CDB)
  - ePIC S&C has followed the development of an HSF reference implementation CDB, now in production for sPHENIX and recently adopted by Belle II (migration to come)
  - **ePIC evaluation has begun** with a test instance successfully established
- Giving back to the community: Dmitry will present our substantial list

Progress on a quantitative computing model and its physical layout

See Markus' computing model talk

- ## Computing resource estimate
  - We provide a **first quantitative computing resource estimates** for processing and storage across Echelon 1 and Echelon 2 sites, broken down by computing use cases
  - The model draws on **WLCG's measurement of ~15%** as the long term average (so far; the trend is down) for yearly constant-cost capacity increases for CPU and storage

- Echelon 0 to Echelon 1 data transfer
  - The EIC's unique two-host-lab organization motivates the 'butterfly' computing model in which **BNL and JLab are symmetric peers** in their capability for post-E0 processing
  - We've begun to address how this can be implemented from IP6 to the BNL and JLab computing facilities
    - Does Echelon 0 (DAQ) computing take place entirely at IP6 or does it leverage BNL's computing facility
    - Raw data stream content, path and workflows from Echelon 0 (DAQ) to the two Echelon 1s

BROOKHAVEN
NATIONAL LABORATORY
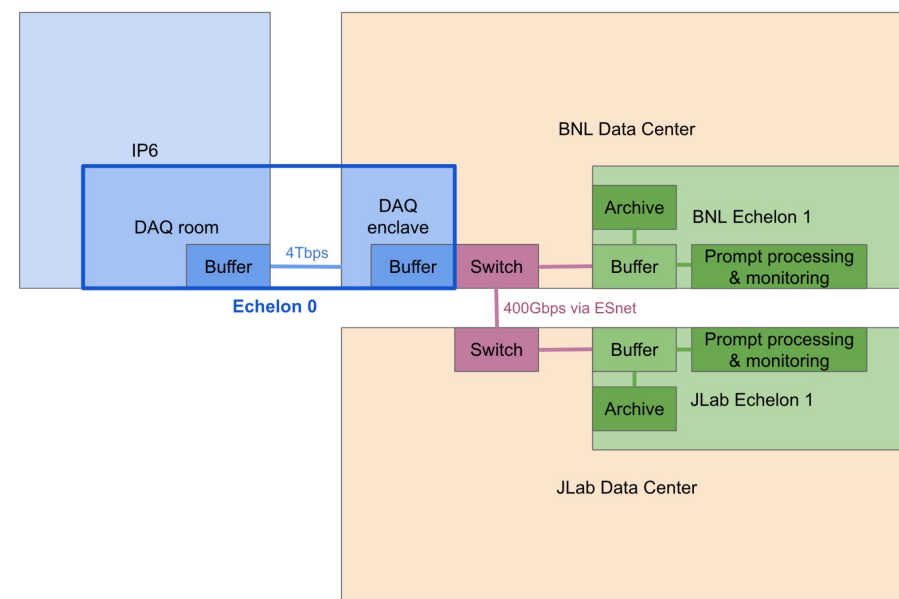
The ePIC Software & Computing effort

- ePIC S&C **remains largely best-effort**, with no dedicated DOE funding
- Vital **university based contributions** from highly motivated experts, including international
- Slow and steady progress on university groups / new institutions joining the effort
- We have **less effort than we'd like on our highest priorities**
- We will comment in this review on how/where we would make use of additional effort to advance our priorities
- Dedicated funds for 2-4 S&C experts embedded among existing experts at labs and/or universities would make a big difference

BROOKHAVEN
NATIONAL LABORATORY