The ePIC Streaming Computing Model
Version 2, Fall 2024

Marco Battaglieri[1], Wouter Deconinck[2], Markus Diefenthaler[3], Jin Huang[4], Sylvester Joosten[5], Dmitry Kalinkin[6], 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.
[6]University of Kentucky, Lexington, KY, USA.

Abstract

This second version of the ePIC Streaming Computing Model Report provides a 2024 view of the computing model, updating the October 2023 report with new material including an early estimate of computing resource requirements; software developments supporting detector and physics studies, the integration of ML, and a robust production activity; the evolving plan for infrastructure, dataflows, and workflows from Echelon 0 to Echelon 1; and a more developed timeline of high-level milestones. This regularly updated report provides a common understanding within the ePIC Collaboration on the streaming computing model, and serves as input to ePIC Software & Computing reviews and to the EIC Resource Review Board. A later version will be submitted for publication to share our work and plans with the community. **New and substantially rewritten material in Version 2 is dark green. The present draft is preliminary and incomplete and is yet to be circulated in ePIC for review.**

Contents

1 Introduction

This report provides an overview of the ePIC computing model, as presently conceived many years before datataking begins. Although ePIC datataking is distant, software and computing is an active and vital part of the ePIC development program today, supporting ongoing detector design optimization, studies of detector and physics performance, and the development of reconstruction and analysis software that will be the foundation for the ePIC physics program. A well-developed software stack and computing capability is also needed early in order to develop an understanding of the complete workflows and their operating scales sufficient to deliver the ePIC physics program, thereby establishing the S&C requirements for the experiment. A capable S&C infrastructure is also needed to support a steady R&D program such that when ePIC datataking begins, it will be with a modern, sustainable software stack that is effective and economical in using the computing facilities of the day; that effectively utilizes the globally distributed computing resources available to the Collaboration; and that strongly leverages the available community software within and beyond nuclear and particle physics. This report addresses these aspects of the computing model and use cases for the present and near term, and the longer term towards datataking.

Section 2 briefly describes the ePIC experiment, the capabilities and characteristics of which inform and drive the computing model, beginning with how the data arrive off the detector. The EIC will deliver e-A collisions (together with backgrounds) in which every collision event is of physics interest to ePIC. The luminosity and cross sections are such that it will be possible, technically and in cost terms, to record every such event. This leads to a defining characteristic of ePIC data acquisition and computing: an integrated streaming readout and processing system will acquire collision events without the more traditional selective physics trigger and stream them down processing workflows that constitute the early stages of the ePIC streaming computing model.

Section 3 describes in some detail the first stage of the streaming computing model, the streaming data acquisition system. The DAQ system receives the data from the experiment's detector subsystems, performs essential data reduction steps such as zero suppression to decrease the data volume without impacting the physics, aggregates the readout data across the subsystems into 'time frames' that record all activity in ePIC's detectors during a time interval, and finally stream the time frame data out to the ePIC computing systems for downstream processing.

When the raw data stream leaves the DAQ system, the immediate actions of the ePIC computing system are to archive it in its entirety and commence prompt processing (reconstruction, validation, diagnostics, calibration). Here we encounter another defining characteristic of the ePIC streaming computing model: mirroring the dual host laboratory organization of the EIC Project, ePIC computing leverages both host laboratories as peers, beginning with the raw data flowing symmetrically to each (the "butterfly model"), such that two geographically distinct, complete copies of the raw data are established in close to real time, with each host lab technically capable of performing whatever downstream processing ePIC and the labs decide upon. Section 4 describes the computing use cases, from these initial steps and on to first-pass reconstruction, reprocessing, simulation, and analysis.

The computing resources available to ePIC will be (are) distributed among collaborating institutions from the host labs to domestic universities and laboratories to the globally distributed facilities of ePIC's international collaborators. Crucially, the available resources also include opportunistic resources which today and in the future account for a significant fraction of the computing (particularly processing; less so storage). The computing model is organized in terms of "Echelons", informed by but distinct from the Tier model of the LHC experiments. Section 5 describes the computing resource requirements of the experiment, in general terms as driven by the streaming data sample and the physics program, and as distributed among the Echelons depending on the use case.

The ePIC streaming computing model is inherently distributed, from the forking of the raw data to two host facilities in the butterfly model, to a sharing of computing responsibilities among institutions and facilities both domestically and internationally. Section 6 describes distributed computing for ePIC: the processing requirements presented by streaming data, and the distributed workflow and data management capabilities that are required.

The ePIC Collaboration gave early priority to establishing a coherent, capable and collaborative software effort and tool set that is able to deliver – first and foremost for the present needs of ePIC, while also establishing a software stack that leverages modern community software and that can evolve sustainably towards ePIC datataking. Section 7 describes the ePIC software, from designing and managing a common stack to (when complete) the particulars of simulation, framework, reconstruction, and other software. The ePIC S&C community puts strong emphasis on user-oriented design to inclusively engage as much of the ePIC community as possible in the software. Accordingly, training, documentation, and other engagement and support tools take high priority and are described.

Section 8 describes ePIC S&C project organization, its collaborations internal to ePIC and participating institutes, and its collaborations with others. Aspects of the organization important to accountability to host laboratories and funding agencies are also described.

Finally, Section 9 describes ePIC's long-term software and computing plan. The timeline and high-level milestones are described, from today's pre-TDR phase through the TDR, supporting test beam needs, optimization and testing during detector construction, conducting computing and analysis challenges that progressively grow in their scale and scope, and on to ePIC commissioning and datataking.

2 The ePIC Experiment

Although the building blocks of the nucleon have been known for decades, a comprehensive theoretical and experimental understanding of how the quarks and gluons form nucleons and nuclei, and how their strong dynamics determines the properties of nucleons and nuclei, has been elusive. Most of the information about the nucleon's inner structure has emerged from the study of deep-inelastic scattering (DIS) process [1–3], an activity which has established QCD as the theory of the strong interaction.

In DIS, a high-energy lepton scatters off a hadron and excites that hadron to a final state with much higher mass. Information on the quark momentum density can be determined by detecting the scattering electron and the additional hadrons produced in the reaction. Correspondingly, information on the gluon density is derived from logarithmic scaling-violations when analyzing DIS data at a range of virtualities [4], or through the photon-gluon fusion process [5]. Information on structure and dynamics beyond a picture of hadrons as collections of fast-moving partons can be obtained by measuring correlations of the struck quark and the further remnants of the hadron. In some cases, the high-energy lepton diffractively scatters ($ep \to epX$), leaving the hadron intact, with no further signature of hadronic products [6, 7]. Such processes offer another context to examine QCD, especially at low x.

Dual advances in perturbative QCD and computation have laid the foundation to imaging quarks and gluons and their dynamics in nucleons and nuclei. The theoretical accuracy of modern perturbative QCD calculations has recently been advanced to next-to-next-to-leading order (NNLO) and beyond, including implementations of heavy-quark mass dependence and thresholds [8–10] in general-mass schemes [11, 12]; these advances enable lepton-hadron scattering as a discovery tool via precision measurements and the observation of new particles, both on its own or in strong synergy with hadron-hadron facilities.

The EIC targets the exploration of QCD to high precision, with a particular focus on unraveling the quark-gluon substructure of the nucleon and of nuclei. It will be designed and constructed in the 2020s, with an extensive science case as detailed in the EIC White Paper [13], the 2015 Nuclear Physics Long Range Plan [14], an assessment by the National Academies of Science [15], and the EIC Yellow Report [16]. The Yellow Report has been important input to the successful DOE CD-1 review and decision. It describes the physics case, the resulting detector requirements, and the evolving detector concepts for the experimental program at the EIC.

In 2021, the host laboratories for the EIC, Brookhaven National Laboratory and Jefferson Lab, invited proposals from detector collaborations to develop the first detector system at the EIC. This detector system, Detector 1, receives its primary funding from the DOE EIC Project and is anticipated to address the scientific objectives described in the EIC White Paper and NAS Report. Three proto-collaborations — ATHENA, CORE, and ECCE — responded by presenting detector concepts. To obtain guidance in selecting the optimal experimental equipment for the EIC, the host laboratories established the EIC Detector Proposal Advisory Panel. By 2022, the panel, composed of renowned and independent scientific-technical experts, concluded that although both ECCE and ATHENA met the criteria for Detector 1, ECCE was the preferable option due to its reduced risk and lower cost. The panel unanimously endorsed ECCE for the first detector system at the EIC. The recommendation also emphasized the importance of the proto-collaboration welcoming more members and expeditiously finalizing its design for a timely transition to CD-2/CD-3A. Immediately following the recommendation, ECCE and ATHENA combined their efforts, culminating in the formation of the ePIC collaboration in 2023.

The ePIC collaboration currently consists of almost 500 members from 171 institutions and is working jointly with the DOE EIC Project to realize the ePIC experiment. Fig. 1 displays a diagram detailing the basic design of the central detector, positioned within a large acceptance solenoid of 1.7 T. The design of the interaction and detector region has been optimized to achieve close to 100% acceptance for all final state particles and ensure their measurement with high precision. The entire integrated detector with the far forward

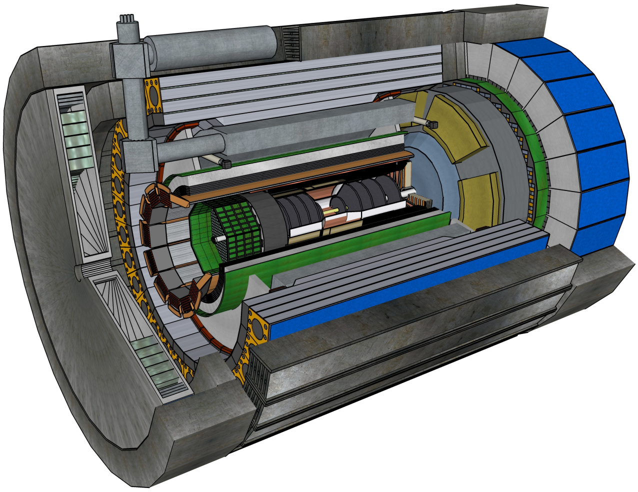


Fig. 1 Drawing showing the ePIC Central Detector.

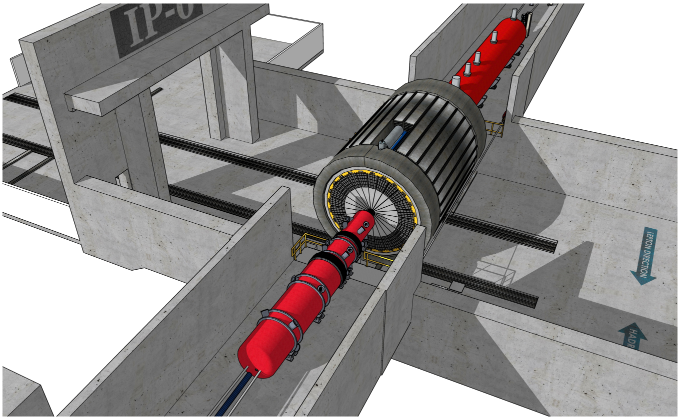


Fig. 2 Drawing of the ePIC Detector, encompassing the far-forward, and far-backward detector regions next to the ePIC Central Detector.

and far backward detector regions spans an approximate length of 90 m, as illustrated in Fig. 2. The primary requirements for the detector include coverage over a broad pseudorapidity range, $-4 < \eta < 4$. Furthermore, maintaining strict control over systematic errors is crucial, necessitating the inclusion of a luminosity monitor and polarimetry for both electron and ion beams.

The EIC is being designed to achieve peak luminosities ranging from $10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ to $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. Considering a luminosity of $10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ combined with strong hadron cooling (where L_{peak} equals L_{avg}) and an operation efficiency of 60% for the collider complex, the resulting integrated luminosity is $1.5\,\mathrm{fb}^{-1}$ every month. The majority of the key measurements can be accomplished with an integrated luminosity of $10\,\mathrm{fb}^{-1}$ [13, 16], which corresponds to a duration of 30 weeks of operations. However, for particular measurements, e.g., the study of the spatial distributions of quarks and gluons within the nucleon using polarized beams, an integrated luminosity of up to $100\,\mathrm{fb}^{-1}$ is necessary. By selecting the beam species and adjusting their spin orientation with care, many measurements can be conducted at the same time.

To guarantee a broad kinematic range and extensive coverage of phase space, the EIC necessitates a variable center-of mass energy $\sqrt{s}$ that falls within approximately 20 GeV to 140 GeV [15]. Some experiments will need variations in $\sqrt{s}$, while others will be conducted at distinct center-of-mass energies.

For the experimental program at the EIC, photoproduction is the dominant physics process. Its cross section is well known and is two orders of magnitude smaller than the cross sections measured at LHC or RHIC experiments. Similarly, particle multiplicities come in at around ten particles in the final state, which is considerably less than those found in pp or pA colliders. The event topologies are known from the DIS measurements from the HERA collider and fixed-target experiments H1, ZEUS, and HERMES. Section 2 offers detailed estimates regarding event rates and data sizes, which include predictions about potential background contributions.

3 The Streaming Data Acquisition System

3.1 Streaming Readout

In its simplest form, streaming readout is the continuous collection of data from the detectors without any selection by a hardware trigger. Each signal over zero-suppression threshold is streamed from the detector with a time-stamp that uniquely identifies its position on the time axes. Along the way to final storage, each stream is independently manipulated applying multiple stages of data. The first stage uses per-channel zero-suppression implemented in detector specific readout electronics. In the early stages, data selection, compression or filtering is performed independently on each channel to the greatest extent possible. This provides the maximum flexibility to change or include new detector components in the readout since channels are not bound to each other. Later stages may use high-level analysis involving sophisticated processes like track reconstruction.

Hits will be organized into time frames, defined by the timing system and applied to the hits early in the readout process. Time frames from different readout electronics can be aggregated by a 'frame builder'. The frame builder can either use standard CPUs or fast and dedicated hardware such as GPUs or FPGAs. In the streaming readout concept the time frame, the picture of the whole detector taken during a certain time interval, represents the basic and full information collected by the detector with the minimum possible bias.

Each frame is then streamed to Echelon 1 computing farms where a processor analyzes it applying a selection algorithm, a software "trigger" written in a high-level programming language, that using the whole information decides if (at least) an 'event' is present in the time frame and deserves to be further reconstructed. Beside proceeding with real-time data processing, if technically feasible, ePIC is planning to record data frames before applying the software "trigger". This will represent an unbiased raw data set that, if required, could be re-analyzed with improved software triggers.

The reconstruction of ePIC events in Echelon 1 is envisioned to proceed naturally from the arrival of the raw streaming data and be accomplished on a time scale O(3 weeks). In order to produce high quality results on such short timescales the calibration, alignment, and reconstruction must be well organized and rely upon automated systems as discussed in sections 4.4 and 4.3. This will require well-defined procedures for the commissioning period, while gathering initial calibration data, as well as detailed monitoring and iteration of calibration while the detector runs.

Separation of time frame data into distinct events can happen late in the reconstruction process. The streaming system offers the advantage over triggered systems that the event selection process can make use of the full detector information, including analyzed information such as tracking. It also offers the the advantage of being able to study detector noise and backgrounds in complete detail.

Some current generation experiments were designed in the conventional triggering scheme and evolved into streaming readout as technology advanced. These include sPHENIX, and the streaming upgrades for LHCb, ATLAS and the JLAB CODA DAQ system. At LHCb the recently deployed streaming readout has enabled the collaboration to decrease the time-to-publication from months-to-years, down to weeks. The ePIC Collaboration has opted from the very beginning to develop the ePIC DAQ and computing model in streaming mode to maximize efficiency and flexibility.

3.2 The ePIC DAQ System

For the ePIC data acquisition will implement an flexible, scalable, and efficient stream DAQ as outlined by the EIC Yellow Report. This design will provide the advantages of streaming include the replacement of custom L1 trigger electronics with commercial off-the-shelf (COTS) computing, virtually deadtime-free operation, and the opportunity to study event selection in greater detail. These advantages come at the cost of greater sensitivity to noise and background.

The ePIC detector will consist of around 24 detector subsystems using several readout technologies which include Silicon Monolithic Active Pixel Sensors (MAPS), Low Gain Avalanche Detectors (AC-LGAD), High Rate Picosecond Photodetectors (HRPPDs), and Silicon Photomultiplers (SiPMs). A schematic of the overall readout scheme for the ePIC detector is shown in Figure 3.

Readout will be accomplished using front end sensors, adaptors, and detector specific ASICs encapsulated into custom Front End Boards (FEBs). The data from the FEBs will be aggregated into Readout Boards (RDOs) using bidirectional, serial, electrical (copper) interfaces between FEBs and RDOs.

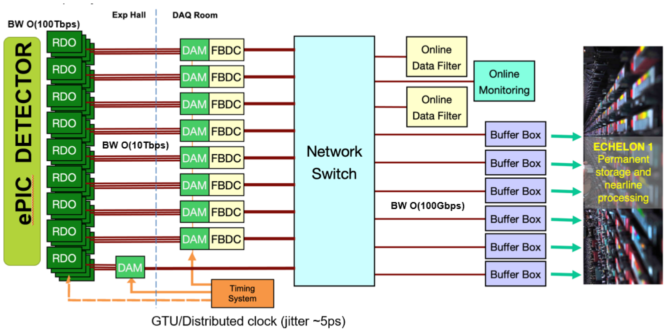


Fig. 3 schematic for the ePIC DAQ

The RDOs will distribute configuration and control information to the FEBs and read hit data as well as monitoring information from the FEBs. These readout components are detailed in figure 4.

The RDOs will also use a bidirectional optical connection to more powerful FPGA-based hardware, the Data Aggregation and Manipulation Board (DAM). The fiber connection between the RDO and DAM will implement a unified, proprietary protocol. This protocol will serve four functions:

- The distribution of configuration information from the DAQ System to configure the RDOs, and to distribute configuration information to the FEBs via the RDOs using their serial links,
- The distribution of real-time control information to the RDO and FEBs,
- The distribution of a high-resolution beam crossing timing signal to the RDO and FEBs,
- The high performance (∼10Gb) transfer of hit data and monitoring information from the FEBs and RDO to the DAM boards.

The Data Aggregation and Manipulation (DAM) boards are envisioned to be the FLX-155 board being developed at BNL for the ATLAS experiment at LHC. These boards will provide the interface between the detector front-end and the “back-end” online computing. These boards are flexible in their function as they can be used as an optional standalone processor (with a 100Gb ethernet output) or as a PCIe interface to a high-performance COTS server (FBDC) as part of the Online Filter.

3.3 High Resolution Clock Distribution

The design of the global timing distribution system (GTU) will be central to the operation of the streaming readout model. The timing system must provide signals to ensure that the data from different detectors can be synchronously aggregated. It must provide a copy of the accelerator bunch crossing clock (running at 98.5Mhz) to all front-end systems. A subset of these systems will require a phase aligned system clock with a jitter on the order of 5ps in order realize required timing resolutions for these detectors (∼20-30ps).

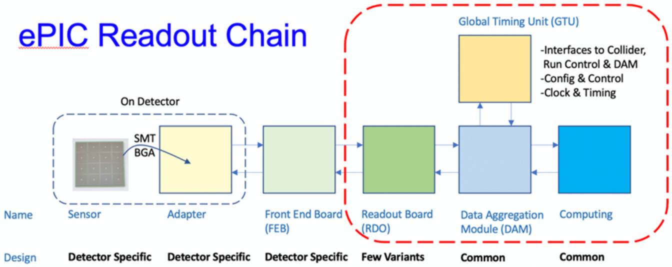


Fig. 4 ePIC full readout chain. Custom, detector specific electronics are required for the readout of each detector. DAQ components common to all detectors are outlined in red.

The GTU is also the only source of real time information provided to the FEB/RDOs, so it must provide information a trigger system would normally provide. These functions include the ability to synchronize data from different detectors, to send flow control signals, to pass bunch information such as spin orientations and bunch structure, the ability to provide user defined signals for signaling special data formatting or calibration needs, and the ability to implement a hardware trigger for debugging or as a fallback option to solve unforeseen readout issues.

The structure of the timing system will include two stages. The first is the GTU electronics which interface to both the collider timing signals and the DAQ control systems. These boards will initially distribute timing signals and information via fiber to the DAM boards. The second stage of the timing system is the communication link between the DAM boards and the RDOs. While there will be several flavors of RDO which each address the specific needs of specific detector electronics, all RDOs will share the common timing, configuration, and data protocol driven by the DAM boards.

We expect the DAM boards to connect to the RDOs using fiber. Each RDO will transmit data to the DAM on a dedicated link. The clock and control connection from the DAM to the RDO can be replicated from a single link at the DAM board. The clock will be reconstructed on the RDO from the transmitted timing system information. This scheme has been demonstrated (CERN TCLink protocol) to be capable of providing a phase resolution of a few picoseconds which is stable even after power cycling, for the Xilinx Ultrascale+ FPGA family and has been tested by the ePIC DAQ group.

For triggered systems it has been traditional to use the bunch crossing signal as the reference clock for digitization. This ensures, once phases are properly adjusted, that the integration windows are oriented on the collisions and that timing windows can be directly applied. In the ePIC detector's streaming readout, the RDOs from all detectors will be required to aggregate zero suppressed data coming from the FEBs tagged by the time. However for streaming, the shaping time and integration time of the signal readout need not be as tightly specified as for a triggered system. Some of the ePIC readout systems use ASICs originally designed for operation at CERN using a 25ns bunch clock. These clocks will be synchronised to the 10ns EIC clock with 2:5

ratio. Additionally, for the case of the ITS-3 based detectors the digitization period is significantly longer than a single bunch crossing. The ePIC DAQ must provide these clocks to the FEBs, account for any phase shift or frequency difference, and provide the information needed to construct the time relative to the EIC bunch crossing signal for all detectors.

Data from the ePIC detectors will be gathered into time frames of approximately 0.6ms for efficient data transfer. This time window has been chosen to balance header efficiency with electronics resources. The time frame window is synchronized for all detectors. Event data from all detectors will be built into a single buffer for each time frame.

3.4 Front End Boards (FEB)

Data Streams being generated on the FEBs need to be driven in a deterministic way, and they must be synchronized to the global clock. Depending on the specific capabilities of the ASICs it may be possible to provide some complementary processing resources at the front-end to support the data framing as well as initial zero-suppression or threshold filtering of the data. These electronics are potentially the most susceptible to radiation effects.

3.5 Readout Boards (RDOs)

The ePIC RDO provides high quality timing information to the FEBs, performs aggregation, and handles initial formatting of hit data. RDOs will use lpGBT or FPGA based processing depending upon the needs of detector FEBs and the radiation environment of the RDO. The RDOs are not expected to have a role in data reduction although simple detector specific algorithms are possible for FPGA based RDOs.

3.6 Data Aggregation and Manipulation Boards (DAM)

For the ePIC DAQ system the DAM boards will be the primary aggregation points for the “raw” detector data streams. Because these are the main aggregation points for the front-end DAQ, there will need to be some well-defined but configurable algorithms for merging streams and managing potential congestion and data loss both for the incoming streams and the outgoing aggregated streams being queued up for back-end processing.

The expected hardware choice for the DAM board the FLX-155 board developed at BNL by the Omega group for use in ATLAS. The FLX-155 Engineering Articles are being tested in the fall of 2024. A render of the FLX-155 is shown in figure 5. Its capabilities are substantial and the updated components ensure a longevity of production, performance and support that are appropriate for the EIC timeline. The board is built around the new Xilinx Versal FPGA/SoC family. This will facilitate using the board both as a PCIe device (supporting both PCIe Gen4 and Gen5 standards) in a server or as a standalone “smart aggregation” switch running a Linux OS. It will support up to 48 serial links to RDOs at the front-end running at speeds up to 25Gbps as

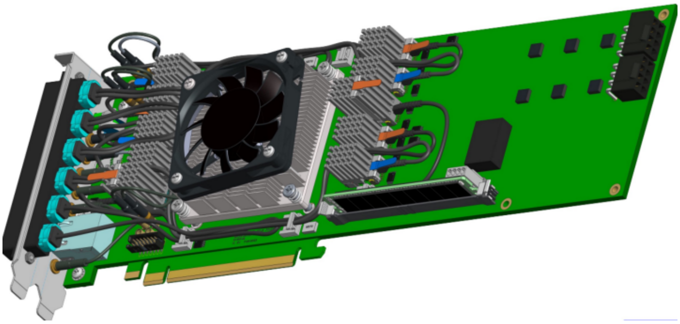
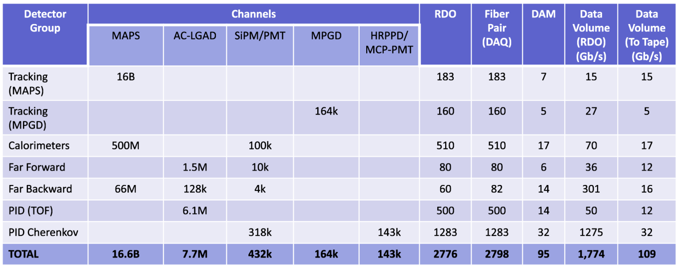


Fig. 5 Image of FLX-155



Detector Group	Channels					RDO	Fiber Pair (DAQ)	DAM	Data Volume (RDO) (Gb/s)	Data Volume (To Tape) (Gb/s)
	MAPS	AC-LGAD	SiPM/PMT	MPGD	HRPPD/ MCP-PMT					
Tracking (MAPS)	16B					183	183	7	15	15
Tracking (MPGD)				164k		160	160	5	27	5
Calorimeters	500M		100k			510	510	17	70	17
Far Forward		1.5M	10k			80	80	6	36	12
Far Backward	66M	128k	4k			60	82	14	301	16
PID (TOF)		6.1M				500	500	14	50	12
PID Cherenkov			318k		143k	1283	1283	32	1275	32
TOTAL	**16.6B**	**7.7M**	**432k**	**164k**	**143k**	**2776**	**2798**	**95**	**1,774**	**109**

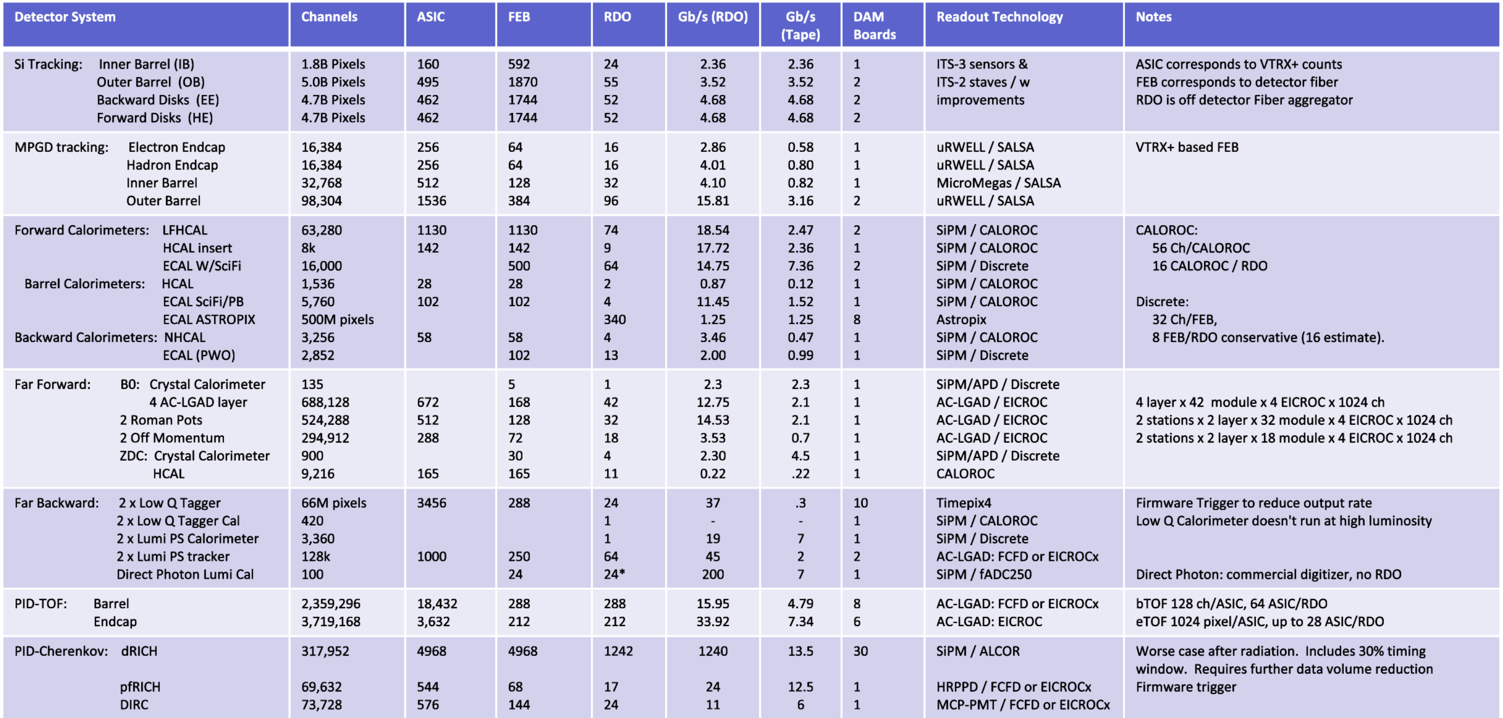
Fig. 6 ePIC DAQ component counts summarized by detector function

well as a 100Gb ethernet link off the board. There is a DDR4 RAM slot available to support buffering and more complex algorithms for data reduction or event identification. The board also supports JTAG and I2C communications.

3.7 Scale of the DAQ System

While the baseline detectors are currently being finalized, our current understanding of the readout technologies, channel counts, RDO, DAM and fiber counts and expected data volumes are summarized in Figure 6 and shown by detector in Figure 7.

The estimated interaction rate for the EIC is up to 500kHz for the highest luminosity of $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. Particle multiplicities in the ePIC detector in comparison to LHC or RHIC are significantly smaller. This means that the majority of bunch crossings will not result in interesting physics. It is important to establish a firm understanding of the sources of background and noise and minimize these rates with respect to the physics signal. For the DAQ system we need to ensure that at the various readout stages there is sufficient bandwidth to comfortably manage expected rates from all detector systems. There are three stages show in Figure 8: digitized data off the detector into the FEB/RDOs at O(100Tb/sec), data into DAM boards and online computing at O(10Tb/sec), and filtered data readout out to disk of O(100Gb/sec). Current data rate estimates are consistent with these values. These estimates



Detector System	Channels	ASIC	FEB	RDO	Gb/s (RDO)	Gb/s (Tape)	DAM Boards	Readout Technology	Notes
Si Tracking: Inner Barrel (IB)	1.8B Pixels	160	592	24	2.36	2.36	1	ITS-3 sensors &	ASIC corresponds to VTRX+ counts
Outer Barrel (OB)	5.0B Pixels	495	1870	55	3.52	3.52	2	ITS-2 staves / w	FEB corresponds to detector fiber
Backward Disks (EE)	4.7B Pixels	462	1744	52	4.68	4.68	2	improvements	RDO is off detector Fiber aggregator
Forward Disks (HE)	4.7B Pixels	462	1744	52	4.68	4.68	2		
MPGD tracking: Electron Endcap	16,384	256	64	16	2.86	0.58	1	uRWELL / SALSA	VTRX+ based FEB
Hadron Endcap	16,384	256	64	16	4.01	0.80	1	uRWELL / SALSA	
Inner Barrel	32,768	512	128	32	4.10	0.82	1	MicroMegas / SALSA	
Outer Barrel	98,304	1536	384	96	15.81	3.16	2	uRWELL / SALSA	
Forward Calorimeters: LFHCAL	63,280	1130	1130	74	18.54	2.47	2	SiPM / CALOROC	CALOROC:
HCAL insert	8k	142	142	9	17.72	2.36	1	SiPM / CALOROC	56 Ch/CALOROC
ECAL W/SciFi	16,000		500	64	14.75	7.36	2	SiPM / Discrete	16 CALOROC / RDO
Barrel Calorimeters: HCAL	1,536	28	28	2	0.87	0.12	1	SiPM / CALOROC	
ECAL SciFi/PB	5,760	102	102	4	11.45	1.52	1	SiPM / CALOROC	Discrete:
ECAL ASTROPIX	500M pixels			340	1.25	1.25	8	Astropix	32 Ch/FEB,
Backward Calorimeters: NHCAL	3,256	58	58	4	3.46	0.47	1	SiPM / CALOROC	8 FEB/RDO conservative (16 estimate).
ECAL (PWO)	2,852		102	13	2.00	0.99	1	SiPM / Discrete	
Far Forward: B0: Crystal Calorimeter	135		5	1	2.3	2.3	1	SiPM/APD / Discrete	
4 AC-LGAD layer	688,128	672	168	42	12.75	2.1	1	AC-LGAD / EICROC	4 layer x 42 module x 4 EICROC x 1024 ch
2 Roman Pots	524,288	512	128	32	14.53	2.1	1	AC-LGAD / EICROC	2 stations x 2 layer x 32 module x 4 EICROC x 1024 ch
2 Off Momentum	294,912	288	72	18	3.53	0.7	1	AC-LGAD / EICROC	2 stations x 2 layer x 18 module x 4 EICROC x 1024 ch
ZDC: Crystal Calorimeter	900		30	4	2.30	4.5	1	SiPM/APD / Discrete	
HCAL	9,216	165	165	11	0.22	.22	1	CALOROC	
Far Backward: 2 x Low Q Tagger	66M pixels	3456	288	24	37	.3	10	Timepix4	Firmware Trigger to reduce output rate
2 x Low Q Tagger Cal	420			1	-	-	1	SiPM / CALOROC	Low Q Calorimeter doesn't run at high luminosity
2 x Lumi PS Calorimeter	3,360			1	19	7	1	SiPM / Discrete	
2 x Lumi PS tracker	128k	1000	250	64	45	2	2	AC-LGAD: FCFD or EICROCx	
Direct Photon Lumi Cal	100		24	24*	200	7	1	SiPM / fADC250	Direct Photon: commercial digitizer, no RDO
PID-TOF: Barrel	2,359,296	18,432	288	288	15.95	4.79	8	AC-LGAD: FCFD or EICROCx	bTOF 128 ch/ASIC, 64 ASIC/RDO
Endcap	3,719,168	3,632	212	212	33.92	7.34	6	AC-LGAD: EICROC	eTOF 1024 pixel/ASIC, up to 28 ASIC/RDO
PID-Cherenkov: dRICH	317,952	4968	4968	1242	1240	13.5	30	SiPM / ALCOR	Worse case after radiation. Includes 30% timing window. Requires further data volume reduction
pfRICH	69,632	544	68	17	24	12.5	1	HRPPD / FCFD or EICROCx	Firmware trigger
DIRC	73,728	576	144	24	11	6	1	MCP-PMT / FCFD or EICROCx	

Fig. 7 ePIC DAQ component counts

have been compiled from detector experts as well as by detailed simulations of collisions, synchrotron radiation, hadron beam gas, and electron beam gas events as applied to the detector configurations at the proposal stage. These results are expected to hold for the current ePIC detector design.

The reduction from O(10Tb/sec) to O(100Gb/sec) performed in the DAM boards or stages of DAQ online computing will arise primarily by reducing the data volume from detectors using SiPM readout at thresholds that need to be sensitive to single photons such as the dRICH and pfRICH. At these thresholds the SiPM readout has a dark current rate of 300 Hz/mm^2 at -40C. These rates will increase to 270 kHz/mm^2 after several years of radiation damage. An efficient online event selection will reduce the effect of the dark current by a factor of 200 at highest running rates. AI techniques are also being investigated to help accomplish this task. The far backwards detectors will be subject to a similar requirement as they will produce up to 150Gb/sec due to high Electron Bremsstrahlung rates. This data must be processed by the DAQ readout system to produce luminosity measurements, but the full data readout to disk will be reduced by software filtering to on the order of 1Gb/sec.

3.8 DAQ Computing Resources

Table 1 outlines the envisioned resources for the streaming DAQ needs. This is based on the elements shown in the DAQ schematic in Figure 3. Several thousand fibers from the RDOs will be aggregated in the DAM boards and the DAM outputs presented to the online farm. Each online farm node represents a multi-core server. The expectation is that they will minimally support 32-64 cores, and selected nodes will support PCIe-based GPUs and/or FPGAs (in addition to the DAM boards in the EBDC nodes). The high performance DAQ network is expected to support 100/400Gbps bandwidth connections. As the majority of the DAQ computing is expected to be COTS hardware, much of it will be acquired as late as is reasonable in the construction phase.

Resource	Totals
DAM/FELIX boards	136
EBDC Servers	92
DAQ Compute Nodes	108
File Servers (Buffer Box)	6

Table 1 DAQ Computing Resources

In the ePIC streaming model, there will be many independent streams of data coming off the detector electronics (FEB). These streams will be aggregated initially at some level by RDOs and further aggregated/processed by the DAM boards. The DAM output streams will be made available to the “back-end” processing farm for the streaming DAQ. The expectation is that all the stream processing will be done on COTS based networks, servers, and other high performance computing hardware (GPUs, FPGA boards etc.). The scale of this infrastructure is dependent on both the aggregate bandwidth of the streams and the level of processing required to reduce the aggregate data set to a level allowing for permanent storage.

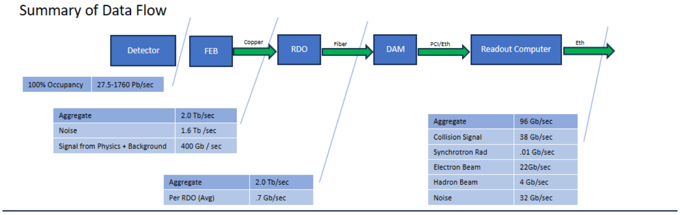
The primary function of the DAQ computing farm is to read the data from the DAM boards, package it in data files, buffer it, and send it downstream for further processing. It will need to apply low-level data processing and reduction to accomplish this. It must also provide sufficient resources for monitoring to ensure the proper operation of both the DAQ and the detectors. All these tasks will involve correlating data between different detectors. A critical part of the monitoring system must, in fact, ensure that the correlation between detectors is robust. The DAQ system will also need to construct and display information in real time, including beam and background scalers. It will need to provide databases (DBs) to track configuration history and to track data produced. It will need to provide real time monitoring and logging.

Two options are being studied for the physical placement of DAQ computing resources, as shown in Figure **??**. Examination of the technical implications of the two scenarios has recently begun, with cost and risk analyses still to come.

In the first option all DAQ computing resources are located in the DAQ room at IP6. The data stream that leaves IP6 is the raw data stream destined for the two Echelon 1 sites (host lab computing facilities). This has the advantage of all resources being consolidated in one location and fully under the control of the ePIC DAQ team. It requires an extensive computing infrastructure build in the DAQ room to support the full system. Tight space constraints will limit the headroom and expandability of the system.

In the second option BNL's recently completed computing facility in B.725, where SDCC and the BNL Echelon 1 are located, is used as the host of a 'DAQ enclave', an extension of the DAQ system into a well-provisioned computing facility where computing infrastructure and space are amply available. In this scenario DAQ's online farm resides in the enclave, requiring the farm's inputs to traverse fiber connections from IP6 to the enclave totalling about 4 Tbps. In this scenario the integrity of the integrated, but now distributed, DAQ system as a 'DAQ owned' system must be preserved. The enclave must be supported as an extension of the DAQ room in terms of physical access for DAQ experts, cybersecurity architecture, and full sysadmin level control of computing equipment. The IP6-enclave network (fiber and switches) would similarly have to be 'DAQ owned', with full control and no traffic sharing. These requirement could be ensured through an SLA, and a draft SLA would be a concrete and concise means of framing the requirements. [will we write one?]

For both options, the data stream leaving the DAQ system, with egress either from the DAQ room or from the enclave, travels to a switch in the computing facility where it forks to the two Echelon 1 facilities at BNL and JLab.



100% Occupancy	27.5-1760 Pb/sec

Aggregate	2.0 Tb/sec
Noise	1.6 Tb /sec
Signal from Physics + Background	400 Gb / sec

Aggregate	2.0 Tb/sec
Per RDO (Avg)	.7 Gb/sec

Aggregate	96 Gb/sec
Collision Signal	38 Gb/sec
Synchrotron Rad	.01 Gb/sec
Electron Beam	22Gb/sec
Hadron Beam	4 Gb/sec
Noise	32 Gb/sec

Fig. 8 Expected worst-case data rate contributions for the ePIC detector

3.9 Integration of Slow Controls

There will be myriad slow controls information associated with both the EIC collider and the ePIC detector. These include various systems on the beam line, magnets, detector biases, gas flows, temperatures, pressures, etc. While the design and implementation of these slow control systems will be driven by the relevant subsystems they are associated with, it is the defined responsibility of the DAQ to provide software tools to facilitate the integration of all this information with the streaming physics data. This will include synchronizing the times associated with readout of slow control systems and the bunch-crossing clock that will be driving the DAQ system. Online slow control databases to support calibration and reconstruction processing will also be developed. Finally, a general network infrastructure in the experimental hall and counting house, independent of the high performance DAQ network, will be provided to support all slow control systems.

3.10 Event Rates and Data Sizes

The effort to estimate the expected data volume from the ePIC detector is in progress. Collision and beam gas backgrounds from both the electron and hadron beams have been studied, but there are continued efforts to ensure that all detectors are included using proper energy thresholds and digitization schemes. The effects and mitigation of synchrotron radiation also continues to be studied. The method for converting hits to data volume is to assume a constant detector-specific bit size based on current assumptions of the digitization for each detector. The distribution of hits within to specific regions of detectors has also been studied and is being incorporated into the design of the readout systems.

The hit rate for collision signal is taken from simulated hits for DIS events generated by the ePIC physics and detector simulation. The simulated data set was taken for 18x275 GeV collisions with $Q^2 > 0$ with luminosity $1.54x10^{33}cm^2s^{-1}$. The collision rate was 83kHz, but the hit rates were scaled to the maximum rate of the EIC collider of 500kHz. Hadron and electron beam gas events were generated using the simulated vacuum profile after 100Ah of pumping. Noise calculations are currently based on the ePIC detector group expert estimates.

The general strategy of the ePIC DAQ is to apply as few data reduction strategies as is required to successfully store the data. However, the data rates from some detectors will require DAQ processing. Figure 8 shows the expected contributions from signal, background, and noise at each stage in the ePIC data flow. The maximum contributions are summarized by detector in Figure 6 and Figure 7. There are several notable features of the expected data rates that will require data processing.

- The SiPM dark current rates are included in these calculations as noise. These increase with radiation damage, so the quoted numbers are after several years of expected operations. After the damage reaches these levels an annealing process is planned to partially mitigate these rates.
- The SiPM dark currents are expected to be particularly problematic for the dRICH detector because it must be run with thresholds sensitive to single photons. The electronics have sufficient bandwidth to read all of the data to the level of the DAM board but in this case we expect an online event selection to be necessary to reduce the data volume by a factor up to about 30 to fit into the ePIC data budget.
- The far backward detectors are expected to see up to 18 tracks per bunch crossing due to very high bremsstralung rates. These hits will be summarized into bunch-by-bunch luminosity calculations at the DAM board level, but we also expect it to be necessary to apply an online event selection for the full data.

3.11 Transferring Data from DAQ to Offline

Many varieties of data and metadata will be transferred from DAQ to offline. For each subdetector, the data sent can include – as well as "regular" detector data – samples of data not processed by DAQ's data reduction algorithms for monitoring and data integrity checks, summary data (luminosity measures, scalers), and detector metadata (bad channels, threshold information, run information).

The details of the raw data model and the format of the data being transferred from DAQ to offline are under development. After the data reduction algorithms of the DAQ have been applied, the DAQ system will aggregate hits from across all detector subsystems into 'time slices' that record all activity in the detector over the duration of the slice (0.6 ms). One of the primary objectives of streaming computing is a holistic reconstruction using all the information from each detector system. Understanding biases that might arise from low-level data processing and reduction in the DAQ and prior to archiving the data stream is of fundamental importance, and it is essential to circumvent these biases when feasible. DAQ's minimal data reduction and assembly of time slices is designed to deliver the raw data downstream with the minimum possible bias.

The raw data stream containing event data in time slice format is flows into a disk buffer after the stream's creation in the online farm. DAQ inserts file markers in the stream, with a filesize (to be determined) that optimizes low-latency realtime processing and storage system efficiency. The buffer ensures latency tolerance to avoid deadtime, assuring smooth streaming operation and robustness against data flow interruptions. The buffer depth is foreseen to be about 1 week. The stream simultaneously flows out of DAQ for distribution to the two Echelon 1 facilities.

Upon arrival at an Echelon 1 site the stream again flows into a disk buffer. This again supports smooth streaming operation and protection against processing flow interruptions, and the buffer serves as an input pool for Echelon 1 processing. The foremost Echelon 1 responsibilities are to archive the raw stream (delivering the full stream to each Echelon 1 would be a natural way to assure two complete geographically separated copies) and, simultaneously with the buffer as input, to perform prompt processing.

For the first few years of ePIC datataking the event data stream arriving from Echelon 0 will be archived in full, untouched. Once the detector, its readout and its event data characteristics are well understood, ePIC may introduce data reduction processing at Echelon 1 prior to archiving. The algorithms that would perform this reduction are likely to be operating in passive mode from the beginning. The delivered system will support archiving 100% of the raw stream, while also accommodating pre-archive data reduction at the discretion of ePIC.

The per-event data volume in the raw stream will be much higher than nominal in the first years of ePIC datataking, blown up by extra detector data for developing detector understanding and debugging. This will be counterbalanced by the expected low luminosity early on, and less noise (before dRICH radiation effects build up). A safe assumption is that the bandwidth and storage volume that the system makes available will be fully used at any given time.

The system is designed to support the Echelon 1 sites as symmetric peers for post-DAQ computing. It will be up to the ePIC Collaboration together with the sites to determine as a matter of policy the specific computing roles of the Echelon 1 sites. The responsibility of ePIC DAQ and S&C is to provide the capability needed by the Collaboration to flexibly establish policy and evolve it over time.

4 Computing Use Cases

In this section, we outline the computing use cases for the Streaming Computing model. In Sec. 5, the use cases are associated with the four tiers of the ePIC Streaming Computing Model computing fabric, Echelons 0 through 3. Echelon 0 refers to the ePIC experiment. Echelon 1 pertains to the host labs. Echelon 2 encompasses global processing and data facilities. Echelon 3 concerns home institute computing.

4.1 Interface between DAQ and Computing

Where the interface lies between “online” and “offline” in the ePIC streaming data and processing flow is still a matter of discussion. The working definition for the purposes of this document is the point at which data flows to archival storage. In aspects both technical and sociological, this is the point at which substantial differences exist on the two sides. All processing prior to delivering the archival stream is at risk of permanently losing data in case of error or reduced live time. Post archival, the requirements and latencies are less stringent, the environment is more open.

This Section describes the computing use cases on the offline side of this definition, beginning with the stored data stream and its associated monitoring.

4.2 Stored Data Streaming and Monitoring

The first and foremost responsibility of the data stream processing as it receives archive-ready raw data from DAQ is to archive it. ePIC’s butterfly model provides for geographically separated replicas of raw data as it is archived, by symmetrically receiving the raw data stream at both BNL and JLab facilities, and archiving to tape at both sites. The data is also retained on disk at both sites for near real time workflows such as calibration and prompt processing, discussed below. Monitoring of the raw data stream and other data and metadata received from DAQ provides for examination, validation and alarming of the data stream, both by automated means and via UI. Monitoring can also consume the reconstructed objects produced by prompt reconstruction. Background analysis and subtraction can take place to ready the data stream for subsequent processing.

4.3 Alignment and Calibration

ePIC aims for rapid turnaround from datataking to full calibrated reconstruction, making a prompt alignment and calibration loop vital. It will operate off the same buffered raw data stream (and prompt reconstruction data set) that is available at each site, and will be as automated and autonomous as possible in its operation. Workflows may ingest raw data or (by definition incompletely calibrated) reconstructed data as input. Alignment and calibration data products as used in the reconstruction and other downstream workflows are delivered to a conditions database available globally, and refined until final, ready for final reconstruction. Initial prompt reconstruction based alignment and calibration is restricted to Echelon 1 (like prompt reconstruction itself). Refinements towards a final calibration can proceed elsewhere as well.

4.4 Prompt Reconstruction

A defining characteristic of ePIC's streaming data model is the events are reconstructed in near real time from the streaming data, modulo time varying calibrations that will require later reprocessing for a final fully calibrated reconstruction. The prompt availability of reconstructed data, and concurrent calibration cycle consuming it, is a crucial element of ePIC's objective to have a rapid, near real time turnaround of the raw data to production, as expressed in the software principles[17]. The stringent low latency and high availability requirements of prompt reconstruction, together with the locality of its inputs at the Echelon 1 sites, makes this a processing activity limited to Echelon 1. Prompt reconstruction uses streaming based processing described in Section 6.1 below, taking time frames as produced by the DAQ as input and producing event (single interaction) based data as output, for processing by analysis software.

4.5 First Pass Reconstruction

It is expected that the Echelon 1 facilities will have insufficient compute resources to perform the complete first pass reconstruction for incoming data. The prompt reconstruction workflow at Echelon 1 will process, at a minimum, the sample necessary for monitoring, diagnostics, quick-turnaround calibration and so on. The remaining first pass reconstruction processing will be shared with Echelon 2 facilities. The maximum acceptable completion time is about 2-3 weeks. This timescale is driven by calibrations. Given the expectation of relatively low data rates during commissioning and early running, and the need to commission, validate and stabilize the use of Echelon 2s for first pass reconstruction, it is likely that Echelon 2s will be integrated after the first pass reconstruction workflow at Echelon 1 is operating smoothly and Echelon 2s are validated as ready.

4.6 Reprocessing

The reprocessing use case can take several specific forms: full reprocessing from time frames (expected to be infrequent, after commissioning), re-reconstruction of event-factorized data with updated reconstruction and calibration (as soon as calibrations are available, plus a few more times per year), and regeneration of analysis object data as selections against the full data sample evolve (frequent). The analysis object data will be compact enough to "take home". All reprocessing workflows are amenable to batch style processing and can utilize Echelon 1-2 and opportunistic resources.

4.7 Simulation

Monte Carlo simulation in ePIC will encompass physics simulation (event and background modeling) and (with physics simulation as input) detector simulation, both fully detailed (Geant4) and fast (parameterized, ML based). At least one order of magnitude more simulated events than data will be needed for ePIC's various run configurations in order to estimate systematic uncertainties, ensuring simulation will remain a substantial production workload and resource consumer after datataking is underway. The output of simulation and subsequent digitization will have the frame-based streaming structure matching that of real data, such that the reconstruction operates on simulated data exactly as it does on real. (This is not yet implemented.) However in its production, simulation data has more in common with conventional batch processing than streaming. That said, we aim to set up the simulation workflows to mimic streaming data production workflows in an active attempt to gain experience with these workflows prior to datataking.

From a workflow and resource utilization perspective, reconstructing the simulated data within the same workflow is preferable, e.g. avoiding a storage-consuming output stage after the simulation, and avoiding the complication of distinct MC simulation/production workflows. Technical and sociological considerations may however separate these workflows at certain times, for example if the lifetime of simulated data (slow release cycle, determined mainly by experimental setup changes and major software releases) is substantially longer than for reconstructed data (fast release cycle determined by rapidly evolving reconstruction algorithms). Both workflow configurations should be foreseen. Simulation workflows can utilize Echelon 1-2 and opportunistic resources.

4.8 Analysis

The EIC has a broad science program. The analysis effort in ePIC categorizes its studies into inclusive, semi-inclusive, and exclusive measurements, the investigation of jet and heavy-flavor physics, and the exploration for physics that goes beyond the standard model. Each category encompasses numerous observables under examination. The feasibility of analysis prototyping and some types of analysis aligns with the capacities of Echelon 3. Nonetheless, many studies, such as imaging the quark-gluon structure of the nucleon, necessitate the computing resources of Echelon 2 or 1. The traditional approach for these analyses is rooted around immediate data reduction of large amounts of detected particles into multi-dimensional histograms. Corrections for experimental effects, such as background effects, limited detector acceptance and resolution, and detector inefficiencies can then be deconvoluted from the observable of interest through simple arithmetic and matrix transformations. This procedure of deconvoluting experimental effects from histogrammed observables is referred to as unfolding. In contrast, there are emerging analysis techniques at the event level. The event-level approach requires a reversal of the traditional procedure of correcting and unfolding measured histograms: here, idealized events from theory have to be folded with the relevant experimental effects. After folding, the theoretical calculations can then be directly compared with the experimental events at the detector level. The accuracy and precision of these methods depend on intricate simulations in the unfolding scenario and detailed modeling of experimental effects in the folding scenario.

4.9 Modeling and Digital Twin

The streaming data will be used as input for modeling the background for detailed studies of the background under various conditions of the EIC and ePIC detector. Furthermore, ePIC plans to use the complete information from the experiment to create a digital twin of the experiment. This digital twin will complement the detailed detector simulations. It will provide a model of the experiment to be used as input for experimental control in situations where immediate feedback from the model is necessary. The digital twin also offers a model that can be easily shared, facilitating the reproduction of results without the necessity of running computationally intensive detector simulations. The digital twin also allows for the exploration of different scenarios, providing complementary information to gain deeper understanding and optimization of experimental conditions. This, along with the data analysis and detector simulations, will offer valuable insights into improving run plans and potential upgrades for the experiment. Modeling workflows can utilize Echelon 1 and Echelon 2.

5 Computing Resources

5.1 The Computing Model's Resource Requirements

Figure 9 shows the Butterfly Computing Model that will be used for ePIC. In this model the detector together with the DAQ system constitute Echelon 0. Two Echelon 1 sites are located in the computing centers of the host labs BNL and JLab, with the capability to serve as functional peers for host-site ePIC computing, both providing storage and processing resources. Echelon 2 sites contribute processing resources and may also provide storage resources convenient for access by distributed collaborators (see Section 5.4 for details).

The computing resources needed for all phases of data processing after the raw stream leaves Echelon 0 will be distributed across multiple facilities. The overall resource requirements are therefore cumulative among them, together with the networking required for robust distributed operation.

Overall, Echelon 0 will need to send raw data at 200Gbps and each Echelon 1 site will need to be able to receive data at 100Gbps. Additional bandwidth will be needed at the Echelon 1 sites to send data to Echelon 2 sites for processing and to receive the results. Storage bandwidth and volumes are driven by these rates and are detailed in the following sections.

The resource requirement estimates are based on the EIC/ePIC timeline and expected luminosity ramp as of Fall 2024. The expected maximum luminosity of the EIC for ePIC is $\approx 10^{34} cm^{-2} s{-}1$[16]. It will take some years beyond the start of physics operations to attain this level. Present planning puts the luminosity target for initial 'Phase 1' physics operations at $\approx 10^{33} cm^{-2} s{-}1$. The resource estimations below are based on this Phase 1 scenario, unless specified otherwise, taking 2034 as a nominal Phase 1 datataking year. The computing model is designed to ultimately serve the full luminosity

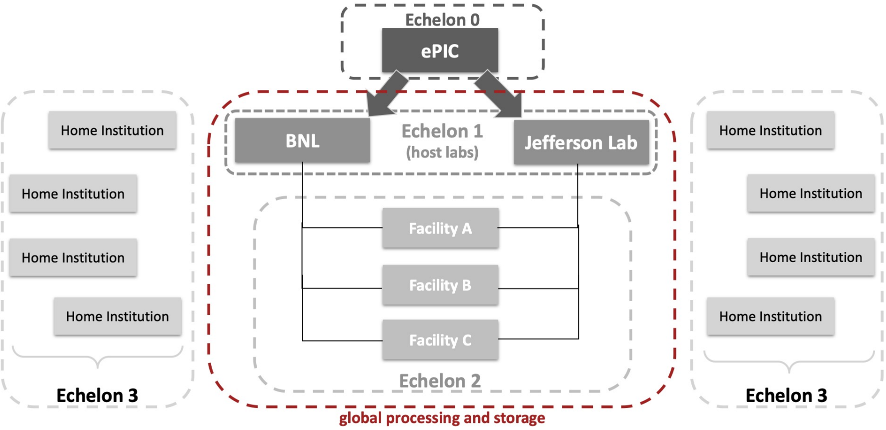


Fig. 9 Butterfly Computing Model *see text for details).*

delivered by the EIC, while tailoring the deployed scale to current needs at any given time, thereby gaining cost savings from computing cost reductions over time.

5.2 Echelon 0: The Stored Data Stream

The maximum EIC luminosity of $10^{34}cm^{-2}s{-}1$ is expected to correspond to a nominal data egress rate from DAQ of ≈100Gbps (see sec. 3.10). During early Phase 1 running at lower luminosity the rate is expected to be circa 1/3 of that, the lower luminosity counterbalanced by a larger volume of detector data used for debugging and developing a detailed understanding of detector behavior. Given the uncertainties, for the purposes of resource estimation we use the full 100Gbps rate.

This is an instantaneous rate that will be reduced to the average rate via a data buffer in Echelon 0 just prior to the exit. While the average rate may be around 50% of the maximum, the system will be designed to accommodate the full 100Gbps bandwidth between Echelon 0 and each of the Echelon 1 sites. This will allow for closer to real-time processing of the data offsite. Both of the host labs will therefore receive the full storage-level data stream in real time.

Thus, Echelon 0 will use 200Gbps of outgoing bandwidth. A small amount of additional outgoing bandwidth will be needed for monitoring streams, slow controls data, and misc. metadata artifacts. These are expected to contribute $\leq$ 1% to the total. The utilized bandwith on the network should not exceed 50% of the maximum throughput, putting the network bandwidth requirement out of Echelon 0 at 400Gbps.

A summary of the Echelon 0 rates can be seen in Table 2.

The incoming bandwidth to Echelon 0 is expected to be small by comparison to the total outgoing bandwidth. This will include incoming monitoring data from higher Echelons and relevant calibration values (see section 3.2.5 of [18]).

The Echelon 0 storage will be primarily short term disk in the form of the output Data Buffers. The buffers will serve to smooth out fluctuations in the DAQ rate as well as provide a means to store data for a short period of time in the event of a temporary loss of communication to the Echelon 1s. The buffer will be sized to hold at least 24 hours of raw data produced at the full 100Gbps rate, corresponding to on the order of 1PB.

Resource	**Type**	**Amount**
Outgoing bandwidth	raw data	200Gbps
	Monitoring, slow controls, misc. meta data	$\leq$ 1Gbps
	Capacity headroom	$\approx$ 200Gbps
	TOTAL	**400Gbps**
Incoming bandwidth	monitoring, calibration	$\leq$ 1Gbps
Storage	Disk (outgoing data buffer w/ 24hr)	1PB

Table 2 Echelon 0 networking and storage requirements.

5.3 Echelon 1: ePIC Computing at the Host Labs

The host labs at Echelon 1 will each receive a full copy of the data. Current planning calls for the bandwidth, storage and compute requirements to be the same for both Echelon 1 sites. An Echelon 1 site will need to be capable of receiving 100Gbps and permanently storing both the raw data and the reconstructed data for the full data set.

The following Subsections describe the computing resource requirements for networking, storage and processing at Echelon 1, necessarily describing also the proportion of computing not done at Echelon 1. Processing not performed at Echelon 1 will primarily be done at Echelon 2s, described in the next Section. The ePIC Collaboration will also draw on opportunistic resources for computing needs that can be met away from the Echelon 1 centers; opportunistic and special resources are discussed in a later Section.

5.3.1 Echelon 1 Networking

The Echelon 1 sites will require sufficient incoming bandwidth to receive 100 Gbps of raw data and outgoing bandwidth to serve the Echelon 2 sites they connect to.

Preliminary plans have the near real-time computing for reconstruction of the raw data stream ('prompt processing') being split equally between each of the Echelon 1 sites. Over time, prompt processing is likely to be shared also with Echelon 2 sites. At any rate, Echelon 2 sites will also collectively host some number of raw data copies for reprocessing and calibration. This means each Echelon 1 site will need additional outgoing bandwidth at a level of at least 1/6 of the total raw data stream or $\approx$ 35Gbps with capacity headroom for steady state running.

The calibration, monitoring, and slow controls data will be needed by each Echelon 2 site. While the bandwidth for all of these combined is small relative to the full raw data stream, the Echelon 1 sites will need to supply multiple Echelon 2 sites with copies of those values.

Resource	Type	Amount
Outgoing bandwidth	Raw data -*immediate* ($\frac{1}{6}$ of total)	17Gbps
	Raw data -*replay* (contingency)	50Gbps
	monitoring, slow controls, misc. meta data	1Gbps
	TOTAL	**68Gbps**
Incoming bandwidth	monitoring, calibration, slow controls *(from E0, E1, and Echelon 2)*	1Gbps

Table 3 Echelon 1 networking requirements. Values shown are for a single E1 site.

The bandwidth requirements for each Echelon 1 site to the Echelon 2 sites it serves is shown in Table 3.

5.3.2 Echelon 1 and 2 Storage

Each Echelon 1 site will require enough archival and disk storage to hold the entire raw data set as well as the Echelon 1's share of downstream data processing and simulation. The estimated raw data size for one year of running during Phase 1 is ≈70PB (see *Table 4* of [18]). The total Echelon 1 storage required across all workflows for a year of data is ≈200PB. This assumes that the Echelon 1s between them hold a complete simulated data sample (≈100PB), a complete first full reconstruction dataset and one re-reconstruction dataset (each ≈10PB).

Fast disk access will be needed to store raw data while calibrations are done and data is processed at either an Echelon 1 or 2 site. Raw data files will not be deleted from disk until their corresponding reconstruction artifacts are stored in both Echelon 1 tape archives. This process is currently estimated to take up to 3 weeks allowing for an extended calibration period. The prompt processing sample, currently estimated (conservatively) as 50% of the full sample, also will have a steady-state disk presence. These disk resident samples are expected to total about $5 - 10$PB in Phase 1.

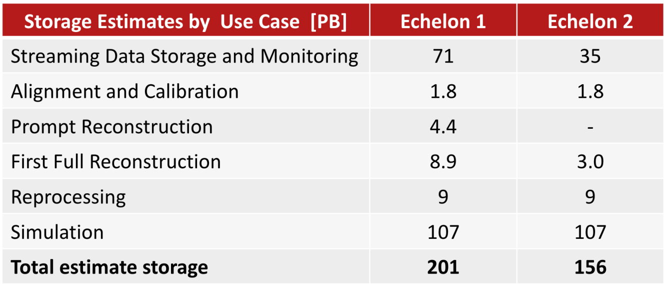
The raw data as well as simulation and reconstruction data may be primarily tape resident in order to economize on disk, with a 'data carousel' workflow used to dynamically orchestrate the placement of a sliding window of tape-resident data on disk for processing. This approach has been used successfully for many years at RHIC and ATLAS, for example.

Additional disk space will be required for individual user analyses. Some of this will be distributed throughout the Echelon 2 sites, but it is anticipated that the Echelon 1 sites will also be used for this purpose. To estimate this, we assume these analyses will require an additional 10% of the reconstructed data volume (1% of the raw data volume) and that it will be distributed amongst the Echelon 1 and 2 sites in the same proportions. This is considered negligible for the purposes of this estimate, smaller than the uncertainties.

Echelon 2 sites will be important contributors to all workflows apart from raw data streaming and (at least in early running) prompt reconstruction. We use the following estimates of their participation share for the purposes of this resource estimation.

Fraction of alignment/calibration done outside Echelon 1	50%
Fraction of first full reco done outside Echelon 1	40%
Fraction of reprocessing reco done outside Echelon 1	60%
Fraction of simulation done outside Echelon 1	75%

Table 4 Fractions of production workflows performed outside of Echelon 1 sites (i.e. at Echelon 2 and opportunistic sites)



Storage Estimates by Use Case [PB]	Echelon 1	Echelon 2
Streaming Data Storage and Monitoring	71	35
Alignment and Calibration	1.8	1.8
Prompt Reconstruction	4.4	-
First Full Reconstruction	8.9	3.0
Reprocessing	9	9
Simulation	107	107
Total estimate storage	**201**	**156**

Fig. 10 Storage resource estimates aggregated over Echelon 1 and Echelon 2 sites, for a nominal Phase 1 datataking year.

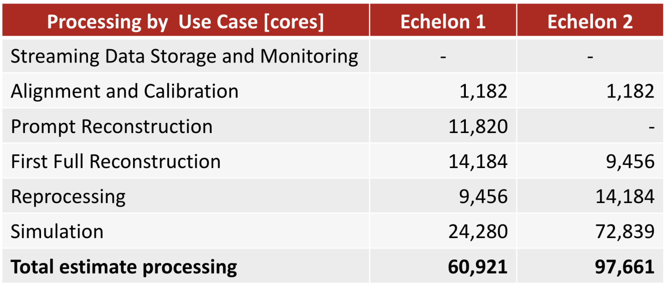
Storage resource requirement estimates are summarized in Figure 10.

5.3.3 Echelon 1 and 2 Compute

Determining the scale of the ePIC Streaming Computing and planning for computing resource needs during the commissioning and operation of the experiment are essential. For a dependable estimate, a prototype for the holistic reconstruction of physics events from time slices is required. This reconstruction needs to include jet reconstruction and the identification of leptons and hadrons using all PID systems in the ePIC Detector. It is important to have reliable estimates of the fraction of background events in the data stream and their impact on the reconstruction performance in the time slices, and to understand how quickly these background events can be discarded without the need for full reconstruction. Defining the alignment and calibration methods for each subsystem and having detailed discussions about fast alignment and calibration techniques are crucial to estimate the computing resources required for alignment and calibration.

Substantial progress has been made in the year since the first release of this report, and while much work remains to be done and the uncertainties remain large (and will remain so for many years), we provide here a first estimate of the processing resources required by ePIC. Important steps during the year making this estimate possible were the implementation of time slice based track reconstruction; a careful examination of the physics and background rate environment; a similar examination of the data and processing flow of the raw stream from DAQ; and a detailed survey of alignment and calibration methodologies and timelines (during datataking) across ePIC's detector subsystems. As for the storage estimates above, the benchmark is a nominal Phase 1 year, $10^{33} cm^2 s^{-1}$ luminosity, circa 2034.

Processing resource requirement estimates are summarized in Figure 11. Raw data streaming to storage and real-time monitoring requires a relatively small amount of processing, smaller than the uncertainties. Alignment and calibration processing is estimated based on a 50/50 participation of E1 and E2; this is very approximate, the main point is that calibration and alignment processing will engage both Echelons. Prompt reconstruction will take place at Echelon 1 only in the early years, potentially extending to Echelon 2s later. The capacity estimate includes a factor of two for headroom sufficient to catch up if prompt processing is interrupted. The Echelon proportions for first full reconstruction, reprocessing and simulation are as for the storage case above. Note that the total processing numbers are an overestimate given that not all workflows are active at any given time, and all available computing resources will be applied to active workflows.



Processing by Use Case [cores]	Echelon 1	Echelon 2
Streaming Data Storage and Monitoring	-	-
Alignment and Calibration	1,182	1,182
Prompt Reconstruction	11,820	-
First Full Reconstruction	14,184	9,456
Reprocessing	9,456	14,184
Simulation	24,280	72,839
Total estimate processing	**60,921**	**97,661**

Fig. 11 Processing resource estimates aggregated over Echelon 1 and Echelon 2 sites, for a nominal Phase 1 datataking year.

5.4 Echelon 2: Global ePIC Computing

The ePIC Collaboration is international and its computing will be as well. This is expressed in the computing model as soon as it extends beyond the Host Labs to become global, at Echelon 2. An essential component of ePIC computing, relied upon to achieve the computational scale necessary to meet the experiment's scientific goals, will be the resources contributed formally by ePIC's collaborating institutions around the world, which represent the Echelon 2 component. The computing model must be designed to effectively integrate these resources and manage their productive use, wherever they may be located, dependent of course on factors such as network connectivity.

The dual Echelon 1 structure of the ePIC computing model, the "butterfly model", already places distributed computing requirements on the model. Effectively integrating and leveraging globally distributed resources at Echelon 2 extends this requirement. The experience of the LHC experiments, well represented within the ePIC Collaboration, is relevant and applicable to developing an effective model for ePIC. Because Echelon 2 resources will be formally relied upon to meet computing requirements, they must come with appropriate MOUs specifying service requirements and assuring technical implementations compatible with the ePIC computing model. The ePIC Collaboration for its part commits to a joint effort on facility integration, and the provisioning of sufficient testing/validation protocols, monitoring and diagnostics to convey to the Echelon 2 facility, in sufficient detail to guide remediation, the faults and performance lapses that occur.

Connectivity of the Echelon 2 sites to Echelon 1 will be the same to both Echelon 1 sites (Host Labs). The connectivity will ultimately be to the ESnet network backbone to which the Host Labs are both connected. Echelon 2 sites will not have connectivity just to one or the other Echelon 1. Similarly, the Echelon 2 sites themselves will be interconnected as determined by their network environment, and these connections will be exploited by the computing model, e.g. for data replication among sites. A clear lesson from the LHC, which evolved from a hierarchical model to an interconnected mesh as experience was gained, is that the latter is far more effective.

5.5 Echelon 3: Home Institute Computing

The Echelon 3 component of the computing model is where the ePIC collaborator doing analysis or developing software directly encounters the computing system. People will access ePIC computing from their institutional cluster, their work desktop, their personal laptop, and so on. Serving these use cases is the role of Echelon 3, and the Echelon 1 and 2 based services that support the Echelon 3s (like access to data). Like Echelon 2, Echelon 3 resources are global, as well as local to the user. These resources are numerous, diverse, often volatile and opportunistic, restricted in their use, and not suited to be managed as Collaboration resources. Rather the Collaboration will provide the tools, interfaces, connection points, data access mechanisms and support mechanisms to make such resources effective portals and analysis processing resources for ePIC analysis.

5.6 Opportunistic and Special Resources

Among the software and computing principles[17] guiding ePIC are those expressing the importance of leveraging as many computing resources for the collaboration as is possible and practical. ePIC software should be able to run on the architectures and platforms available, effort permitting, while leveraging system characteristics such as the presence of accelerators (GPUs, TPUs, etc.), again effort permitting. ePIC S&C should support distributed workflows on the computing resources available to the worldwide EIC community, leveraging not only conventional cluster “high throughput computing” (HTC) but also high performance (HPC) systems with good usability and thereby a rewarding cost/benefit calculation.

The most productive computing resource currently used by ePIC is the Open Science Grid (OSG)[19], where a concurrent core count of 5-10k is stably attainable. As ePIC builds up its own computing resources we expect opportunistic resources like the OSG to continue to play a role, in particular for simulation production (detector and physics simulation). Simulation is a relatively simple workflow that puts moderate demands on resources (storage needs, I/O intensity, memory), steady state processing, and a relatively relaxed time to complete requirement. While ePIC’s essential simulation requirements should be accommodated by planned and assuredly available resources, anticipating that ePIC science will be compute limited the exploitation of opportunistic resources should be foreseen. OSG has its origins as the US component of the Worldwide LHC Computing Grid (WLCG). The WLCG is evolving to also support non-LHC experiments (e.g. DUNE, SKA) and we can anticipate that opportunistic resources will be available to ePIC internationally as well.

Commercial clouds are being actively used by science communities (Rubin Observatory and ATLAS are examples) with their capabilities and cost models under study. Opportunistic (preemptible) usage modes together with workflows that elastically spike into the resource to support fast-turnaround use cases such as analysis are the most promising in terms of cost effectiveness. In ePIC we will monitor such developments and participate as we are able, and will decide at a later date whether such resources will have a role in our computing model.

Special resources include non-x86 processor architectures such as ARM, accelerators such as GPUs and TPUs, and no doubt others yet to emerge over the next decade. A requirement on ePIC S&C infrastructure is to have the flexibility and extensibility in the software and CI to add support for architectures of interest as they appear. The ARM architecture is already supported, and we anticipate it will have an important role in coming years given its cost effectiveness per dollar and per watt, and the relative ease of the port. FPGAs are used in the Streaming DAQ for low-level data processing and reduction. GPUs are highly likely to play a role online; whether the same is true offline is unclear. Nonetheless support for high concurrency in the software will be needed, with requirements such as multithreading support, and advantages such as efficient memory utilization. The rise of AI/ML and accompanying proliferation of specialized accelerators such as TPUs makes it probable we will exploit them, perhaps largely transparently behind software APIs. We will track the technologies as we pursue our own AI/ML R&D and applications.

Large supercomputers such as the leadership class facilities (LCFs) developed by the DOE and NSF are most often constituted by what we've called special resources. Whether such machines are effective for ePIC use will be a case by case evaluation. Today's GPU based machines offer limited potential given the dearth of GPU-capable workloads in ePIC (a common situation in NP and HEP), though we are doing R&D in GPU-amenable areas such as Cherenkov detector simulation. The US will have its first leadership class ARM machine in 2026, at the NSF's TACC facility[20], with Japan and Europe hosting others; such machines we would already be able to use effectively. LCFs are increasingly being designed as AI/ML factories; such machines we will assuredly be able to use for at least training and optimization. We have since Sep 2023 an R&D project underway to leverage large scale resources for the processing-intensive AI application of EIC detector design optimization.

5.7 Authorization and Access

Authorization and access mechanisms are evolving both in their technical aspects and the institutional policies that govern their use, thereby impacting the accessibility for users. The foremost priority of the ePIC Collaboration is to ensure that every collaborator has access to the resources of the collaboration, including data, websites, collaborative tools, information systems, document repositories and so on, today and reliably in the future. This consideration can be a leading or determining factor in the tools and services we use, and where they are hosted. It has been a factor in choosing GitHub as code repository and a cloud-based Mattermost instance, for example. We will continue to make this a requirement.

6 Distributed Computing

The ePIC collaboration consists of a globally distributed community of scientists engaged in the experiment's data and compute intensive scientific program. Section 4 described the use cases and workflows that the ePIC computing infrastructure must support. Section 5 described the computing resources of ePIC from the detector to the host labs and on to the globally distributed data and processing centers providing the collaboration with resources, and finally to the local resources used by analysts at their institution or from their laptop. This Section describes the distributed computing software and services that will be needed in order to knit these resources into a coherent computing fabric for ePIC that serves the full spectrum of use cases.

The ePIC experiment follows a lineage of "big science" collaborations using computing resources on a global scale, the most prominent example to date being CERN's LHC experiments, which in their development towards the High Luminosity LHC (HL-LHC) are also preparing for a rich and data intensive physics program in the 2030s. The LHC's ALICE and LHCb experiments have further commonality with ePIC in having introduced streaming computing models for the LHC's present Run-3. The LHC experiments and their collaborators in the WLCG community have built and continue to develop expertise, tools and global infrastructure that the proliferating big science community can draw on. The ePIC approach to distributed computing described here is built on leveraging and collaborating with this community, bootstrapping our distributed computing infrastructure from existing components and approaches where possible so our own efforts can focus on the extensions and tailoring needed to support the unique aspects of ePIC's streaming computing model and global collaboration.

6.1 Processing Requirements for ePIC Streaming Data

The processing of ePIC streaming data has characteristics that are markedly different from the workflows commonly found in NP and HEP experiments to date. Current convention is that data is acquired in online workflows that deliver the data to hierarchical storage as large files, and then processed by offline workflows with a typically substantial latency period after acquisition (apart from promptly processed subsets for monitoring, data quality and possibly calibration purposes). In this scenario the offline processing maps readily onto the batch queue based resource provisioning mechanisms of computing centers. Offline processing payloads are sent to batch queues and consume input files distributed appropriately for resource locality. Keys to the applicability of this straightforward approach are the discrete, coarse grained processing units in the form of files and collections of files (datasets), and the decoupling of processing with respect to real time data acquisition. The case of ePIC streaming data processing, however, has neither of these characteristics.

In ePIC streaming data processing, a quasi-continuous flow of fine-grained data must be processed promptly with the dynamic flexibility to match in near real time the inflow of acquired data to processing resources that stand ready to consume it. Prompt processing is necessary to ensure data quality and detector integrity during datataking, and while processing of a subset could achieve those aims, processing the full dataset quickly is necessary to minimize the time required for calibrating the detector and delivering analysis-ready reconstructed data promptly, a primary goal of ePIC. For ePIC data processing, with the two host labs symmetrically serving as Echelon 1 processing centers, the processing resources used at any given time must be transparent to the workflow engine, effectively a requirement that a distributed processing capability be an integral part of the system. The data sources are distributed as well; in a streaming computing model that dissolves much of the distinction between online and offline, the system must be flexible towards decisions as to the parallelism of data delivery received from the DAQ, i.e. where the event builder function occurs. The system must support processing parallel streams of data from subdetector, accelerator, beamline and other sources, augmented by sufficient metadata to make their association and merging fault-proof. The minimized latency and high system complexity require that a high level of automation and resilience to changing conditions be built into the streaming processing system, necessary also to keep the operations effort at a manageable level.

Summarizing the driving characteristics of ePIC streaming data processing, it is time critical, proceeding in near real time; it is data driven, consuming a fine-grained and quasi-continuous data flow across parallel streams; it is adaptive and highly automated, in being flexible and robust against dynamic changes in datataking patterns, resource availability and faults; and it is inherently distributed in its data sources and its processing resources. This model presents challenges for an infrastructure based on batch jobs and coarse grained files. However, the safe assumption for the infrastructure of the 2030s is that batch-style processing and coarse grained files – particularly as they map onto archival storage – will remain. A robust approach to building the ePIC streaming computing model and system will be to accommodate, but effectively hide, those underlying characteristics of the infrastructure. We may ultimately not need to accommodate them, for example Kubernetes or similar mechanisms of dynamic processing resource provisioning may displace the batch model. We should accommodate both and be resilient against technology evolution.

6.2 Workflow Management

As described, the requirements of ePIC's streaming based prompt reconstruction are distinct from the typical workflow management practices of contemporary experiments. Streaming is however a fertile and rapidly evolving field, in our community and well beyond. Many streaming data processing frameworks and tools exist and evolution is rapid. ePIC should be ready both to take judicious advantage, and avoid technology lock-in. The tools generally share a fundamentally similar distributed parallel model, and have common features that do not risk lock-in such as the use of standard workflow descriptions (e.g. DAG, CWL). Some systems directly manage the processing resources, such as Apache Storm and Spark, others can overlay on conventional batch or dynamic resources (such as Kubernetes); HEP/NP's own PanDA is such a system. The underlying facilities must support high availability and service quality, though a distributed system mitigates against very stringent requirements on a single facility. The facility and the streaming workflow management system in tandem must support data flow optimization in real time.

Resources should be flexible across use cases and workflows, readily usable for other purposes when datataking is not active. For example, applications should be able to scale elastically and exploit heterogeneous hardware such as an AI/ML application spiking into an accelerated resource for low-latency turnaround. Some workflows such as simulation and reprocessing are served well by conventional batch processing, lending advantage to all ePIC's major resources supporting batch.

The international nature of the ePIC Collaboration and its computing makes it essential that workflow management tools support the use of computing resources around the world, for essentially all managed workflows apart from prompt reconstruction, and for physics analysis.

6.3 Data Management and Access

Prompt processing of data streaming from the detector will yield file based data suitable for consumption by hierarchical storage and by file-based data management tools. Raw data copies will be written to archival storage at the two host labs, with the expectation that retrieval is rare. (Under normal operation, no production workflow involves archival data retrieval.) Data management tools must support the distribution and use of data around the world, serving ePIC's global processing resources and community of analysts. Disk resident replicas at Echelon 1 and 2 sites will be managed by the data management system. Client tools for accessing and storing data at managed data stores will be usable at all Echelons including local/personal computers with appropriate authentication. Authentication and authorization (AA) mechanisms must support access for all ePIC collaborators globally.

The broad acceptance of the Rucio[21] data management system as a standard, within HEP and increasingly within NP, makes it the likely system for ePIC datataking use, in its evolved 2030s form. Rucio is being integrated and tested in ePIC now, and ePIC will engage with and contribute to the (very open) Rucio community. Rucio and the distributed computing community is migrating to SciToken based AA mechanisms which enable a federated ecosystem for uniform authorization across distributed scientific computing infrastructures, and should be capable of meeting the collaborator access requirement.

Data movement tools are in a state of flux. The long-used third party copy tool gridftp was recently retired, with http chosen as the basis for replacing it. XRootD is a powerful community-standard tool with data movement functionality tuned to the needs of HEP/NP (e.g. efficient handling of ROOT based data, in terms of both movement and caching). FTS is the data mover underpinning Rucio as used by the LHC experiments. Object store based data storage and movement (supporting the S3 API) are increasingly common. Some DOE computing facilities require the use of Globus data mover tools. Fortunately Rucio can hide much of this fragmentation (Rucio is not in itself a data mover, it interfaces with them). ePIC will leverage this encapsulation and avoid lock-in.

7 Software

7.1 Designing and Managing a Common Software Stack

Giving importance to common community software is one of the guiding principles of ePIC, discussed in Section 9.3.

The design decisions for the ePIC Software stack are based on lessons learned from the global NP and HEP community. Developers of the ePIC Software have been closely following the "Software & Computing Round Table" [22], which is jointly organized by the host labs and the HEP Software Foundation. This monthly round table forum aims for knowledge transfer and to encourage common projects within our scientific community. Notably, members of the ePIC Software & Computing Coordination also play roles in organizing the round table.

For the EIC community, the round table has proven essential. It enables developers to stay informed about software and computing advancements in the NP and HEP and to create a network of significant contacts for collaboration and cooperation.

In addition, the organizers of the "Software & Computing Round Table" also host the "Future Trends in NP Computing" workshop series [23]. These workshops delve into the next generation of data processing and analysis workflows, aiming to optimize scientific output. The workshop topics address questions how to strengthen common efforts in the NP and HEP communities and to outline a roadmap for software and computing in Nuclear Physics for the upcoming decade. Other topics discussed in these workshops include machine learning for enhancing scientific productivity, reusability and common infrastructure components, scaling up and down computing, and how to make analysis easier by addressing issues around metadata handling or the estimate and treatment of systematic uncertainties. resource management, the relationship between I/O, the role of machine learning in amplifying scientific productivity, software portability, reusability, shared infrastructure components, and the challenges of scaling computing capacities. They also focus on simplifying data analysis processes.

Furthermore, the organizers of the "Software & Computing Round Table" also host the "Future Trends in NP Computing" workshop series. These workshops explore the next generation of data processing and analysis workflows, with the goal of optimizing scientific output. The workshop topics addresses questions how to strengthen common efforts within NP and with HEP and to draft a roadmap for software and computing in NP for the next decade. They also cover subjects like machine learning for enhancing scientific productivity, reusability and shared infrastructure components, scaling computing resources, and improving analysis by addressing challenges related to metadata handling and the estimation and treatment of systematic uncertainties.

As ePIC Software & Computing develops, the S&C Round Table and the Future Trends in NP Computing Workshop will continue to be important mechanisms to ensure that ePIC software development continues to have close communication and collaboration channels to the global HEP and NP software community, such that opportunities for common software projects are brought to light and developed.

ePIC is planning and developing a software stack that is common within the collaboration as well as having commonalities outside through common software projects. Within ePIC, one of the software principles[17] is to have tight compute-detector integration, including a common software stack for online and offline software that encompasses the processing of streamed data, aiming for rapid, near real time turnaround of the raw data to online and offline productions. The principle recognizes the convergence between online and offline software in modern NP/HEP experiments with sophisticated high level software triggers, and even more so in a streaming computing model like that of ePIC. The full ePIC prompt reconstruction using "offline" reconstruction software occurs in the critical workflow delivering data from the detector to near real time downstream processing. Developing and using that algorithmic software and the infrastructure around it will be a collaborative effort between online and offline.

This online/offline commonality and shared development requires recognizing the different requirements and environments of online and offline, which are not dissolved by commonalities in software. The real time and near real time online environment has more stringent requirements in software stability, robustness, latency, security and other aspects than the more forgiving and open offline environment. ePIC's software and infrastructure systems must accommodate differing release schedules, stability requirements, testing protocols and so on within a shared software base.

7.2 JANA2 Software Reconstruction Framework

The JANA2 framework is a software framework for applying complex configurations of independent algorithms to experimental nuclear physics data[24]. It is a second-generation framework following the original JANA framework[25–27] which was written initially for the GlueX experiment[28]. JANA2 was developed to modernize the framework and allow implementation of features in newer C++ standards. The ePIC collaboration adopted JANA2 and has been instrumental in introducing additional new features such as *JOminfactory* (see section 7.2.1) and support for *PODIO* (see section 7.2.2).

JANA2 is designed to process independent quanta of data such as events or time frames in a streaming readout configuration. A thread pool is used to parallelize the execution of algorithms efficiently by implementing an arrow-queue topology tailored to the specifics of the experiment. Complete sets of algorithms can be assigned to either a sequential or a parallel arrow (see Figure 12).

A sequential arrow will never have more than one execution thread assigned to it while parallel arrows may have many. For example, time frames may be placed into the first queue by a sequential arrow that is reading them from a file or network socket. A parallel arrow may then execute many threads, each pulling a single time slice from the first queue and identifying events to place in the second queue. Another arrow may then read from the events queue to apply event-level algorithms.

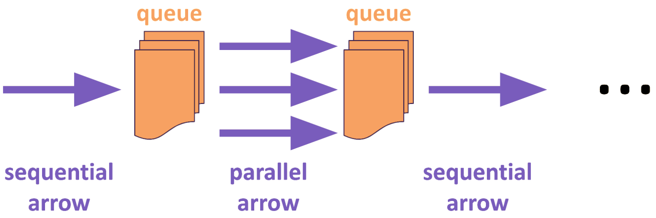


Fig. 12 Schematic example of the arrow queue feature of JANA2. Algorithms (factories) are run by arrows which connect to queues on one or both ends. Parallel arrows may run more than one instance of the algorithm simultaneously while sequential arrows will only one a single instance at any given time.

7.2.1 JANA2 Algorithm Organization

Multiple design patterns are now supported for how experiments implement their algorithms in JANA2. Three of these can be described be the type of factory class used to implement the algorithms:

- JFactory - On-demand algorithm execution
- JMultiFactory - Prescriptive algorithm execution
- JOmniFactory - Prescriptive with generic algorithm support

The *JFactory* class supports a decentralized algorithm configuration with only algorithm outputs declared while the *JMultiFactory* and *JOmniFactory* patterns supports prescriptive configurations with algorithms able to declare multiple inputs and outputs. Figure 13 illustrates a *JMultifactory* configuration where the algorithm calling sequence is completely defined prior to data processing.

The ePIC software (EICrecon) utilizes the *JOmniFactory* which extends the *JMultiFactory* class, adding a small layer to allow the use of generic algorithms specified at run time as opposed to compile time. Generic algorithms here are ones that are not detector specific, but are also framework agnostic. JANA2 can utilize these algorithms through the *JOmniFactory* class even though the algorithms themselves know nothing about JANA2.

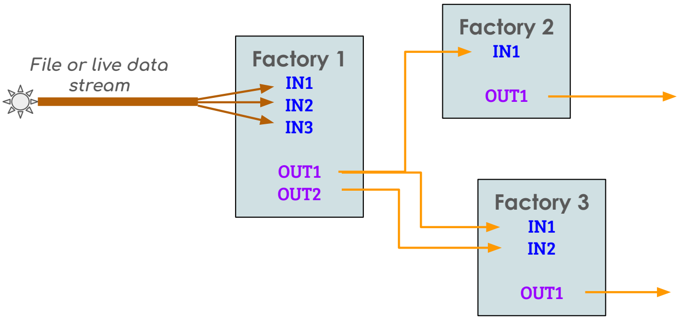


Fig. 13 *JMultiFactory* allows multiple inputs and outputs to be specified directly in the class definition. The JANA2 framework then uses this to execute the algorithms in dependency order so that the inputs are pre-populated before an algorithm is executed.

7.2.2 JANA2 PODIO Support

The ePIC collaboration selected the *PODIO* package for implementing the data model. In response, JANA2 has integrated PODIO support directly in the framework making it the first such framework to do so. Among the features present in PODIO is object association. This allows both one-to-one and one-to-many associations that can be written to and read from persistent storage. When working with PODIO objects, JANA2 utilizes the *collection* mechanism in PODIO.

7.2.3 JANA2 Streaming Readout Support

ePIC plans to implement Streaming Readout Data Acquisition. This means the data will be readout in the form of time frames. Each time frame may contain multiple events and algorithms are needed to identify them. Such algorithms will need to be run at the time frame level, while others will be run on event objects. To support this, JANA2 has implemented an "unfolding" feature that allows quanta of data at one level (e.g. time frame) to generate quanta of data at the next level (e.g. event). Sub-event level quanta are also supported.

8 Serving Users

The EIC community formed largely out of the coalescence of separate user communities: the traditionally separate US-based hadronic physics communities of BNL and JLab, and the heavy ion physics communities of BNL and LHC. Through the detector proposal process, this user community oriented itself around two major initiatives: the ATHENA and ECCE detector proposals. The origin of the communities is relevant here to the extent it informs user attitudes towards software and computing. There is often an affinity for the familiar and a resistance to change, in particular when the change is to a product from a (previously) competing group. So too for the software and computing practices of the ePIC collaboration, which formed out of the merging of the ATHENA and ECCE detector proposal proto-collaborations. To ensure alignment of the ePIC collaboration with a single software and computing strategy, we have used direct user engagement and involvement in decision making using aspects of the *user-centered design* methodology.

In contrast to other communities with other funding mechanisms, the EIC community has only a modest allocation of professional developers allocated to supporting the development of the EIC. The majority of the development is therefore accomplished by a combination of dedicated users who have chosen research software as a focus area, and by users who are driven to contribute where it furthers their analyses but have no training in modern software practices. This highlights the need for *user learning* as a key approach to enhancing maintainability and increasing the quality and number of user contributions.

8.1 User-Centered Design

Anticipating the need for alignment even before the formation of the ePIC collaboration, we developed in concert with the broader EIC user community the EIC Statement of Software Principles[17], which encapsulates our goals for sustainable and future-oriented software. The eight tenets are summarized as follows:

- We aim to develop a diverse workforce, while also cultivating an environment of equity and inclusivity as well as a culture of belonging.
- We will have an unprecedented compute-detector integration.
- We will leverage heterogeneous computing.
- We will aim for user-centered design.
- Our data formats are open, simple and self-descriptive.
- We will have reproducible software.
- We will embrace our community.
- We will provide a production-ready software stack throughout the development.

After the formation of the ePIC collaboration, these Software Principles were the guiding principles for several software decision stages on the tools to center our geometry definition, simulation, and reconstruction workflows around. For each decision, the software and computing coordinators selected an expert to develop a set of requirements, solicited user feedback on these requirement, invited various proposed solutions to the need, and organized an open discussion where all options were presented and discussed. Sustainability, maintainability, but also familiarity and external project control were significant drivers of decisions. Ultimately, decisions were reached (typically after several lively rounds of discussion) to structure our software stack around the following products:

- code repository: hybrid of GitHub as primary repository, and self-hosted Gitlab for continuous integration/delivery,
- geometry definition and detector interface: DD4hep,
- event data model: podio to manage the data model, and EDM4hep as initial data model,
- reconstruction framework: JANA2.

After each decision, and again after the completion of the process, the ePIC collaboration was invited to endorse these decisions.

Since the software stack key components were decided upon, we have aimed to keep the stack *highly modular*. Whenever possible we avoid deep enmeshing of technologies and prefer to have clear separations between components in case a future change is necessary. For example, reconstruction algorithms in the JANA2 framework do not themselves know about JANA2 (nor does their user/contributor/developer need to know about JANA2), since they interface primarily with event data model structures.

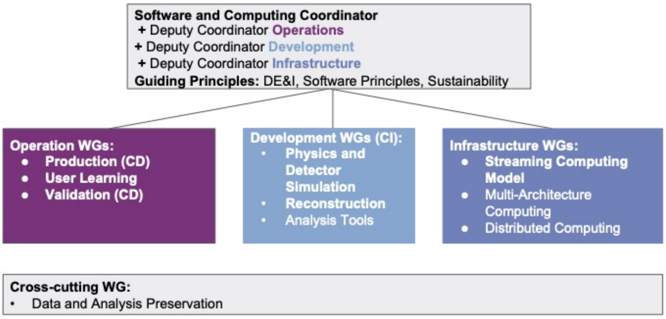


Fig. 14 Organizational chart of the Software and Computing Effort in ePIC. The Streaming Computing Working Group is joint with Electronics and DAQ Working Group in the Technical Effort to ensure that DAQ and Computing are developed together.

8.2 User Learning

Since the growth of the software and computing workforce in the ePIC collaboration comes primarily from users with physics training and without exposure to modern software engineering practices, we have developed a tutorial program that introduces new users to the components of the software stack. New users with an interest in contributing code are directed to Carpentries tutorials that cover underlying software engineering tools (git, bash, python, CI/CD on GitHub and Gitlab, singularity containers).

Although the scalability of virtual tutorials and in-person tutorials at ePIC collaboration meetings has allowed for reaching hundreds of users, two important secondary goals of our user learning strategy are to place developers in contact with the users at a time of learning and exploration, and to provide a pathway for developers to gain the mentorship and leadership experience that is often a precursor to working group convener positions.

9 Project Organization and Collaboration

9.1 Organization of DAQ and Computing in ePIC

The scientific management of the ePIC collaboration is organized in three efforts that report to the spokesperson and deputy spokesperson: an analysis effort with currently two Analysis Coordinators, a technical effort with a Technical Coordinator, and a software and computing effort with a Software & Computing coordinator (SCC). The SCC oversees all aspects of software and computing in ePIC and has three deputies sharing the responsibilities for development, operations, and infrastructure.

Development currently has two active working groups: Physics and Detector Simulations as well as Reconstruction. Another working group on Analysis Tools is being planned. Operations comprises three active working groups: Production, User Learning, and Validation. Among the Infrastructure working groups, which consist of Streaming Computing Model, Multi-Architecture

Computing, and Distributed Computing, only the Streaming Computing Model group is active at present, the others not being an immediate priority. Moreover, there is a planned cross-cutting working group on data and analysis preservation. The activation of the working groups will depend on the number of people actively participating in software and computing.

Two of the three conveners of the Streaming Computing Model WG are also conveners of the Electronics and DAQ WG that is part of the technical effort. Both working groups have regular meetings, and a significant fraction of the attendees of these meetings are the same. This ensures that the DAQ and Computing are developed together with well-defined and well-understood interfaces, and ePIC builds a group of experts familiar with data processing from the DAQ to the analysis.

9.2 ePIC, the ECSJI and the RRB

The ePIC collaboration welcomes the establishment of the ECSJI with its associated bodies including the EIC International Computing Organization (EICO) to provide the organizational structure overseeing and coordinating the complex computing fabric of ePIC and the EIC, extending from the crucial and innovative Echelon 1 partnership between the host labs, to global contributions represented at Echelon 2, to the full support of the analysis community at Echelon 3 and beyond. As well as the host labs, the partnerships represented in ECSJI include partnering with ePIC and future experiments who bring their computing requirements and interests, and with the international community of collaborating countries and Echelon 2 facilities.

It is the computing aspects where ePIC sees a crucial role for the ECSJI. Regarding software, as stated in the formative charge for the ECSJI, the experiments have responsibility for designing and developing their computing models and software, consistent with the computing fabric developed under the oversight of ECSJI. Similarly, ePIC computing operations is an activity developed and executed within the ePIC Collaboration, in close consultation and collaboration with ECSJI, computing resource providers and others. Both the ECSJI and the software and computing efforts of the experiments are subject to oversight and review, the October 2023 review being the first instance.

The ePIC Software and Computing Coordinator serves as ePIC Point of Contact to the ECSJI.

The EIC Resource Review Board (RRB) oversees the resources for the EIC, including those for software and computing. It is the essential mediating and decision making body to reconcile the computing needs of the EIC detector collaborations with the resources available. ePIC has the responsibility to report its computing and software status, its multi-year resource requirements and their justification to the RRB.

9.3 Collaboration with Others

ePIC adheres to the EIC Statement of Software Principles[17] (ePIC members having played leading roles in developing them) and as stated there, we embrace the wider software community, both within our field and the open software community in general. Common software tools from NP and HEP already play a substantial role in ePIC software. The ePIC and EIC community has developed collaborative projects in areas that are both important and ripe for collaborating with and leveraging the wider community. These include AI4EIC[29], a workshop series on developing and AI/ML techniques and tools to EIC science; and MC4EIC[30], a workshop series on Monte Carlo physics generators for EIC which draw heavily on the wider NP/HEP community, including of course theorists.

The EIC Detector 2 software community now beginning to take shape have been our colleagues in developing the statement of principles. Software collaboration between ePIC and Detector 2 should be expected, and early indications are that this will begin to happen soon. ePIC's early start and tight timeline mean that while ePIC software is a natural starting point for Detector 2, ePIC does not have the available effort to develop common software products for two experiments, and common components will need to be established as common development efforts soon, with agreed understandings on development responsibilities and processes.

In drawing on software from the wider NP/HEP community, such as JANA2[31], Acts[32], and Key4HEP[33] and its components, ePIC's role is both user and contributor. ePIC chooses and uses packages like these because behind them are responsive, reliable, collaborative open software communities that ePIC engages with and contributes to. These decisions have been made in an open, well defined and documented process [34], which continues in ePIC for areas yet to be defined.

10 Long Term Software and Computing Plan

10.1 Timeline and High Level Milestones

ePIC Software & Computing has developed closely with the collaboration and with input from the EIC project a timeline of high level milestones, including the long-term. The milestones are grouped by the anticipated date for CD-2 approval, the detector construction phase, and the commissioning and operations phase. Priority is always given to meeting near-term needs, with the longer range timeline progressively exercising the streaming computing model to deliver for the needs of the CD process, for specific applications (e.g. test beams), for scaling and capability challenges, and ultimately for the phases of datataking. The series of milestones ensures that the agile development process is continuously confronted with real world exercising of the software and the developing realization of the computing model.

10.1.1 Preparations for CD-2 and the Technical Design Report

ePIC is providing support to the EIC Project in the preparations for CD-2 decision process and the Technical Design Report for the detector, both of which are due in November 2024:

- Software and simulation readiness for TDR preparation (and subsequent phases of the CD process).
- Provide for each use case Sec. 4 detailed estimates on the compute resources; update the networking and storage estimates according to format of streaming data format that is currently being defined.

10.1.2 Detector Construction Phase

The subsystems of the ePIC Detector must be constructed by 2030. During the construction, the detector designs will be further optimized, considering changes in costing and the availability of materials. Additionally, prototypes of the various detector systems will undergo testing. This will facilitate software testing and provide an opportunity for validation of the simulations based on the results from the test-beam measurements. The relatively extended duration of the construction phase presents an ideal timescale for the further development and implementation of the ePIC Streaming Computing Model:

- Provisioning DAQ and software sufficient for test beams, which can serve as small scale real-world testbeds for the developing DAQ and software.
- Streaming challenges exercising the streaming workflows from DAQ through offline reconstruction, and the Echelon 0 and Echelon 1 computing and connectivity.
- Data challenges exercising scaling and capability tests as distributed ePIC computing resources at substantial scale reach the floor, including exercising the functional roles of the Echelon tiers, particularly Echelon 2, the globally distributed resources essential to meeting ePIC's computing requirements.
- Analysis challenges exercising autonomous alignment and calibrations.
- Analysis challenges exercising end-to-end workflows from (simulated) raw data to exercising the analysis model.

10.1.3 Detector Commissioning and Early Datataking Phase

In approximately a decade, when the ePIC Detector is completed, the experiment's commissioning phase will commence. This phase will come with distinct expectations and requirements in comparison to steady-state operation. For instance, it will involve the utilization of semi-triggered data-taking modes, initial calibrations, introduction of zero suppression, and the gradual extension of near real-time processing from Echelon 1 to Echelon 2, among other tasks. Successful software and computing efforts during the commissioning phase will necessitate careful planning, drawing from the experience gained in the data and analysis challenges during the detector construction phase.

During the initial data-taking phase, immediately following the commissioning phase, simpler and more conservative approaches will be adopted as the ePIC Streaming Computing Model is gradually being deployed and validated.

10.2 Data and Analysis Preservation

A guiding principle[17] of ePIC S&C is that data and analysis preservation (DAP) will be an integral part of EIC software and workflows, aiming for analyses that are fully reproducible, re-usable, and re-interpretable, based on reusable software and amenable to adjustments and new interpretations.

The ePIC Collaboration is planning to incorporate DAP into its software and computing from an early stage. A cross-cutting working group is foreseen in the org chart and will be activated during the next year. It will address DAP requirements and a timeline for DAP developments, prioritizing those with value for ePIC computing and analysis in the near as well as the long term, such as a robust and user friendly infrastructure for containerization in analysis, which is already well advanced in ePIC.

The S&C infrastructure that ePIC is establishing now will facilitate DAP, including containerization of the ePIC software stack, automation of well defined workflows using workflow definition languages (currently used in GitLab based CI), centralized workflow and metadata management (supporting distributed production on OSG), a curated and sustainable code repository and web presence (GitHub and its website publishing tools), and data management supporting the full data life cycle and provenance (Rucio[21] integration is in progress). A prominent missing component at present is document management, being addressed at the Collaboration level.

11 Acknowledgements

Ongoing development of the ePIC streaming computing model and this document benefits from discussions in the ePIC Streaming Computing Model Working Group.

References

[1] Breidenbach, M., Friedman, J.I., Kendall, H.W., Bloom, E.D., Coward, D.H., DeStaebler, H.C., Drees, J., Mo, L.W., Taylor, R.E.: Observed behavior of highly inelastic electron-proton scattering. Phys. Rev. Lett. **23**, 935–939 (1969). https://doi.org/10.1103/PhysRevLett.23.935

[2] Bloom, E.D., *et al.*: High-Energy Inelastic e p Scattering at 6-Degrees and 10-Degrees. Phys. Rev. Lett. **23**, 930–934 (1969). https://doi.org/10.1103/PhysRevLett.23.930

[3] Miller, G., *et al.*: Inelastic electron-Proton Scattering at Large Momentum Transfers. Phys. Rev. D **5**, 528 (1972). https://doi.org/10.1103/PhysRevD.5.528

[4] Gluck, M., Hoffmann, E., Reya, E.: Scaling Violations and the Gluon Distribution of the Nucleon. Z. Phys. C **13**, 119 (1982). https://doi.org/10.1007/BF01547675

[5] Ellis, R.K., Nason, P.: QCD Radiative Corrections to the Photoproduction of Heavy Quarks. Nucl. Phys. B **312**, 551–570 (1989). https://doi.org/10.1016/0550-3213(89)90571-3

[6] Collins, J.C.: Proof of factorization for diffractive hard scattering. Phys. Rev. D **57**, 3051–3056 (1998) arXiv:hep-ph/9709499. https://doi.org/10.1103/PhysRevD.61.019902. [Erratum: Phys.Rev.D 61, 019902 (2000)]

[7] Aktas, A., *et al.*: Measurement and QCD analysis of the diffractive deep-inelastic scattering cross-section at HERA. Eur. Phys. J. C **48**, 715–748 (2006) arXiv:hep-ex/0606004. https://doi.org/10.1140/epjc/s10052-006-0035-3

[8] Witten, E.: Heavy Quark Contributions to Deep Inelastic Scattering. Nucl. Phys. B **104**, 445–476 (1976). https://doi.org/10.1016/0550-3213(76)90111-5

[9] Aivazis, M.A.G., Collins, J.C., Olness, F.I., Tung, W.-K.: Leptoproduction of heavy quarks. II. A unified QCD formulation of charged and neutral current processes from fixed-target to collider energies. Phys. Rev. **D50**, 3102–3118 (1994) arXiv:hep-ph/9312319. https://doi.org/10.1103/PhysRevD.50.3102

[10] Collins, J.C.: Hard scattering factorization with heavy quarks: A General treatment. Phys. Rev. D **58**, 094002 (1998) arXiv:hep-ph/9806259. https://doi.org/10.1103/PhysRevD.58.094002

[11] Guzzi, M., Nadolsky, P.M., Lai, H.-L., Yuan, C.-P.: General-Mass Treatment for Deep Inelastic Scattering at Two-Loop Accuracy. Phys. Rev. D **86**, 053005 (2012) arXiv:1108.5112 [hep-ph]. https://doi.org/10.1103/PhysRevD.86.053005

[12] Gao, J., Hobbs, T.J., Nadolsky, P.M., Sun, C., Yuan, C.-P.: General heavy-flavor mass scheme for charged-current DIS at NNLO and beyond. Phys. Rev. D **105**(1), 011503 (2022) arXiv:2107.00460 [hep-ph]. https://doi.org/10.1103/PhysRevD.105.L011503

[13] Accardi, A., *et al.*: Electron Ion Collider: The Next QCD Frontier: Understanding the glue that binds us all. Eur. Phys. J. A **52**(9), 268 (2016)

arXiv:1212.1701 [nucl-ex]. https://doi.org/10.1140/epja/i2016-16268-9

[14] Aprahamian, A., et al.: Reaching for the horizon: The 2015 long range plan for nuclear science (2015)

[15] National Academies of Sciences, Engineering, and Medicine: An Assessment of U.S.-Based Electron-Ion Collider Science (2018). https://doi.org/10.17226/25171

[16] Abdul Khalek, R., et al.: Science Requirements and Detector Concepts for the Electron-Ion Collider: EIC Yellow Report (2021) arXiv:2103.05419 [physics.ins-det]

[17] EIC Software Statement of Principles. https://eic.github.io/activities/principles.html

[18] Bernauer, J.C., *et al.*: Scientific computing plan for the ECCE detector at the electron ion collider. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **1047**, 167859 (2023). https://doi.org/10.1016/j.nima.2022.167859

[19] Open Science Grid. https://osg-htc.org/

[20] TACC Leadership Class Computing Facility. https://lccf.tacc.utexas.edu/

[21] Barisits, M., et al.: Rucio: Scientific data management. Computing and Software for Big Science **3**(1) (2019). https://doi.org/10.1007/s41781-019-0026-3

[22] Software & Computing Round Table. https://www.jlab.org/software-and-computing-round-table

[23] Future Trends in NP Workshpp Series. https://www.jlab.org/FTNPC

[24] Lawrence, David, Boehnlein, Amber, Brei, Nathan: JANA2 framework for event based and triggerless data processing. EPJ Web Conf. **245**, 01022 (2020). https://doi.org/10.1051/epjconf/202024501022

[25] Lawrence, D.: Multi-threaded event processing with JANA. PoS **ACAT08**, 062 (2009). https://doi.org/10.22323/1.070.0062

[26] Lawrence, D.: Multi-threaded event reconstruction with JANA. Journal of Physics: Conference Series **119**(4), 042018 (2008). https://doi.org/10.1088/1742-6596/119/4/042018

[27] Lawrence, D.: The JANA calibrations and conditions database API. Journal of Physics: Conference Series **219**(4), 042011 (2010). https://doi.org/10.1088/1742-6596/219/4/042011

[28] et al., S.A.: The GlueX beamline and detector. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **987**, 164807 (2021). https://doi.org/10.1016/j.nima.2020.164807

[29] AI4EIC - Artifical Intelligence for the EIC workshop (series). https://indico.bnl.gov/event/19560/

[30] MC4EIC - Monte Carlo event simulation for the EIC workshop (series). https://indico.bnl.gov/event/17608/

[31] Lawrence, D., Boehnlein, A., Brei, N.: JANA2 Framework for Event Based and Triggerless Data Processing. EPJ Web Conf. **245**, 01022 (2020). https://doi.org/10.1051/epjconf/202024501022

[32] Acts - A Common Tracking Software. https://acts.readthedocs.io/en/latest/

[33] Key4HEP software stack. https://key4hep.github.io/key4hep-doc/

[34] EIC Software Decision Process. https://wiki.bnl.gov/EPIC/index.php?title=EIC_Single_Software_Stack_2022