



Report from the EIC Computing Coordination Group

Jérôme Lauret for the ECCG

Outline

- ECCG charges
- EIC activities at JLab
- EIC activities at BNL
- BNL/Jlab common activity
- EIC Joint Institute for Computing and Software
- Conclusions



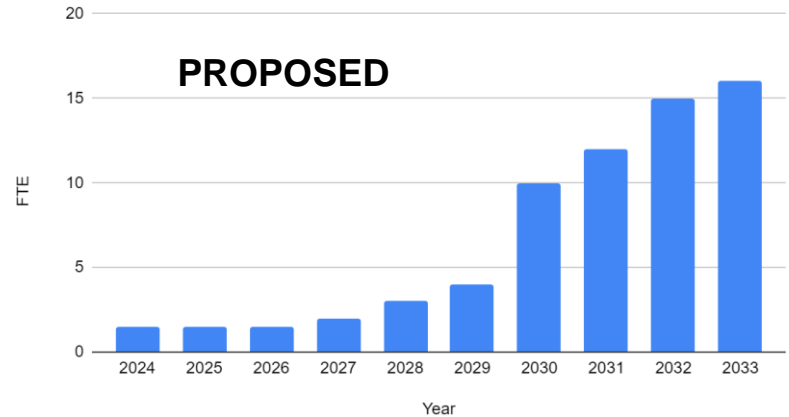
ECCG charges

Coordination of EIC Computing efforts. *G. Heyes, J. Lauret, A. Bressan, W. Deconinck, S.J. Joosten, C. Fanelli*

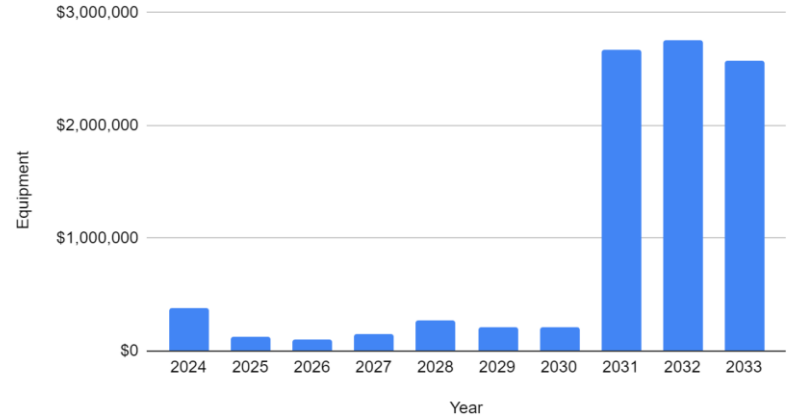
- Coordination of resources among EIC computing efforts & Assure computing infrastructure follows DOE and host lab mandates and guidelines
- Keep records of required, available resources and resource usage
- Providing an access point for external resources
- Infrastructure assessment to fold in international computing resources
- Developing in-common data flow and management for ePIC and detector II

The ECCG provided the needed resources based on known requirements and served as a vehicle to communicate needs (lab). Recommend to constantly keep requirements up to date. Budget discussion January 2023, estimates for 2024-2033 at EACH lab (came with similar numbers) - FTE is cumulative, equipment is per year (~ 9.5 M\$ x 2 needed).

FTE vs. Year



Equipment vs. Year

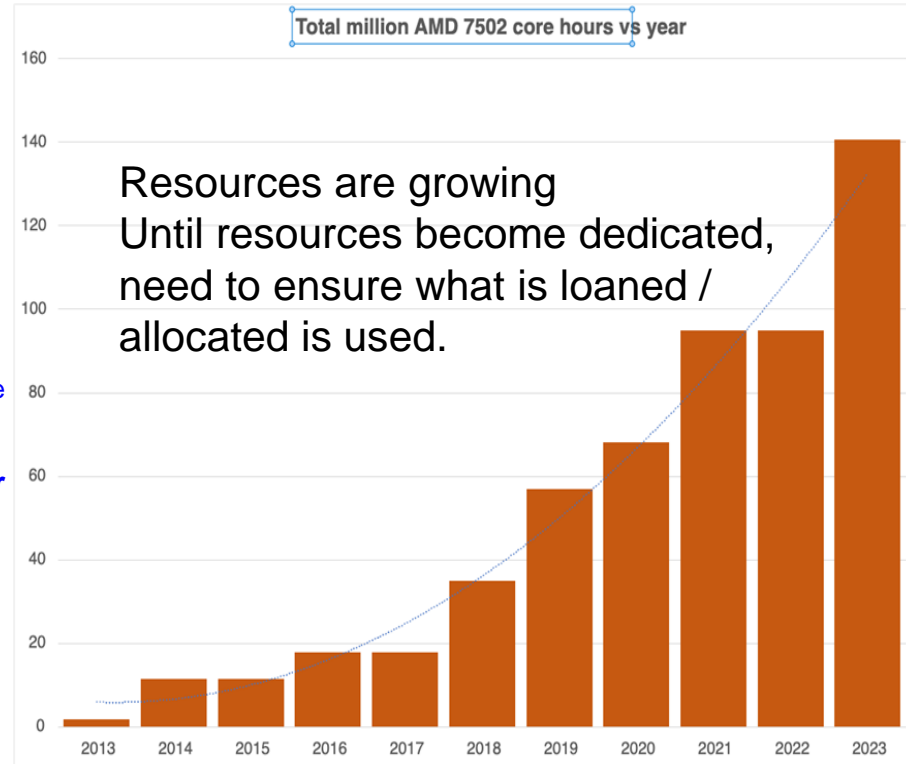


PRESENT



EIC resources at Jefferson Lab

- 3072 cores of EPYC 7763 AMD “Milan” added in FY23
 - FY19 – FY21 Were EPYC 7502 “Rome”
 - Rome vs Milan vs ... see [here](#).
- Current capacity is 140 Million Core Hours (Using Rome to normalize) - **Capacity is now 1.5 x FY22**
- Fair share allocations are used to balance load
 - Bursts beyond share occur when cycles are free
 - The cluster is routinely over 80% busy but...
 - **The EIC fare share is 14% and now have 1.5x more available via that share than last year**
- **EIC actually only used 0.6% of the allocation this year**
 - Of this 0.6% only **11.1% are OSG jobs** all others are **submitted locally** via Slurm. There have been some issues getting OSG jobs to migrate to JLab (being investigated).
- JLab is moving to a model where each Project has its own server (ex: OSG submit host)
 - Enables pilot job tuning and data storage strategies that match the collaboration's needs.
 - Use of OSG is encouraged where practical

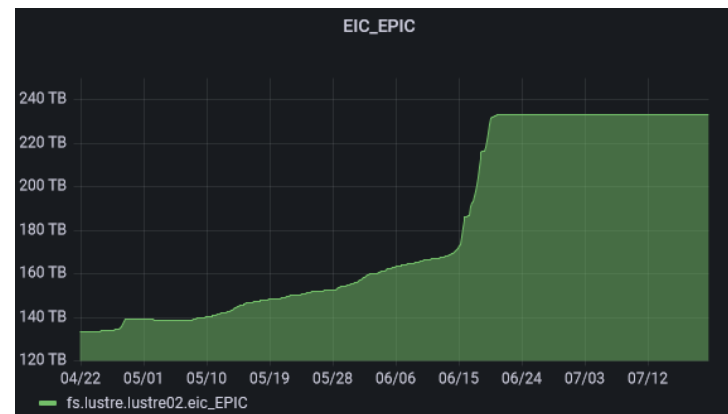
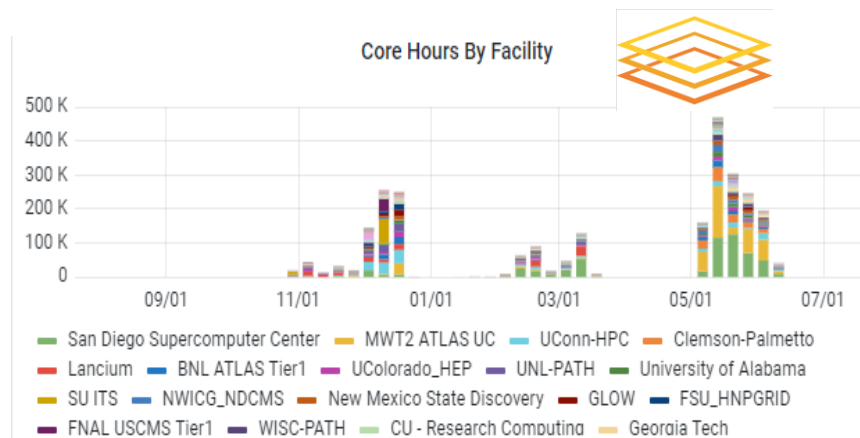


Other activities at Jefferson Lab

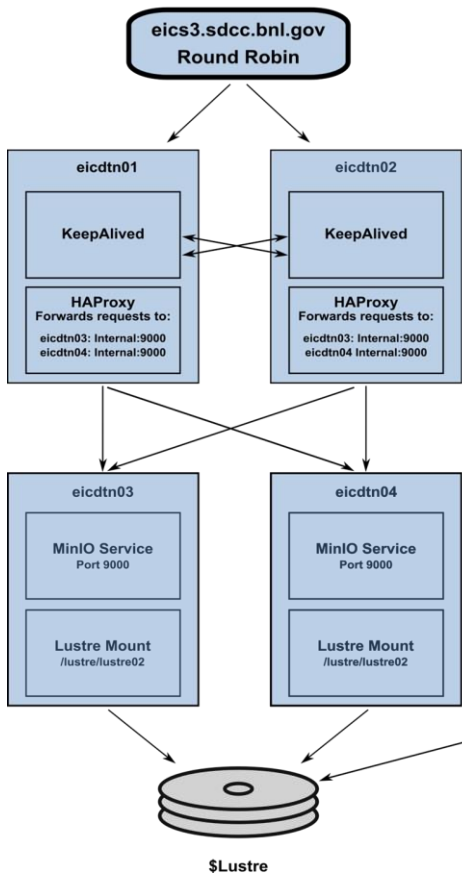
- The Lab's Internet is being upgraded from 2x10 Gbit/sec to 2x100 Gbit/sec by ESNNet
 - Installation delays due of outside cable work and long lead time on optical switch gear.
 - Current forecast is for 100 Gbit operations *before the end of September*.
- The Current storage (Lustre System for /cache and /volatile) is reaching EOL. Replacement on order on the way:
 - Double I/O sustained throughput to >40GB/sec
 - Double capacity to 10 PB raw, 8 PB useable
- The cluster is presently running CentOS 7.9, which is approaching EOL.
 - The next OS for the Farm is AlmaLinux9
 - This is not without its complications
- JLAB is gradually pruning legacy dependencies on local filesystems that have cybersecurity or operational gaps. We will do this in small, announced steps
 - For example, removing: /site, /u/scratch. Deprecating (but still supporting): /apps

Activities @ Brookhaven[®] National Laboratory

- Resources remain the same @ 2,000 slots for the EIC (125k HTC cores available)
- Dedicated OSG submit hosts available
 - osgsub01.sdcc.bnl.gov, osgsub02.sdcc.bnl.gov
 - Little ends at BNL, some local processing (Condor based). Most done via OSG connect.
- BNL provided 3 PBytes of storage to the EIC
 - Supported via *Program Development* funds
 - S3 interface allowed a [seamless access to the storage in read and write \(from anywhere\)](#)
 - Currently ~ 1.14 PBytes / 3 PBytes used, 233 TB ePIC
 - ATHENA and ECCE space still needs to be clean-up
- Recent storage consolidations
 - Redundant service access to S3/monIO (next slide)
 - Monitoring of downtime (inexplicable, server crash reboot - in contact with Dell but no finding so-far)

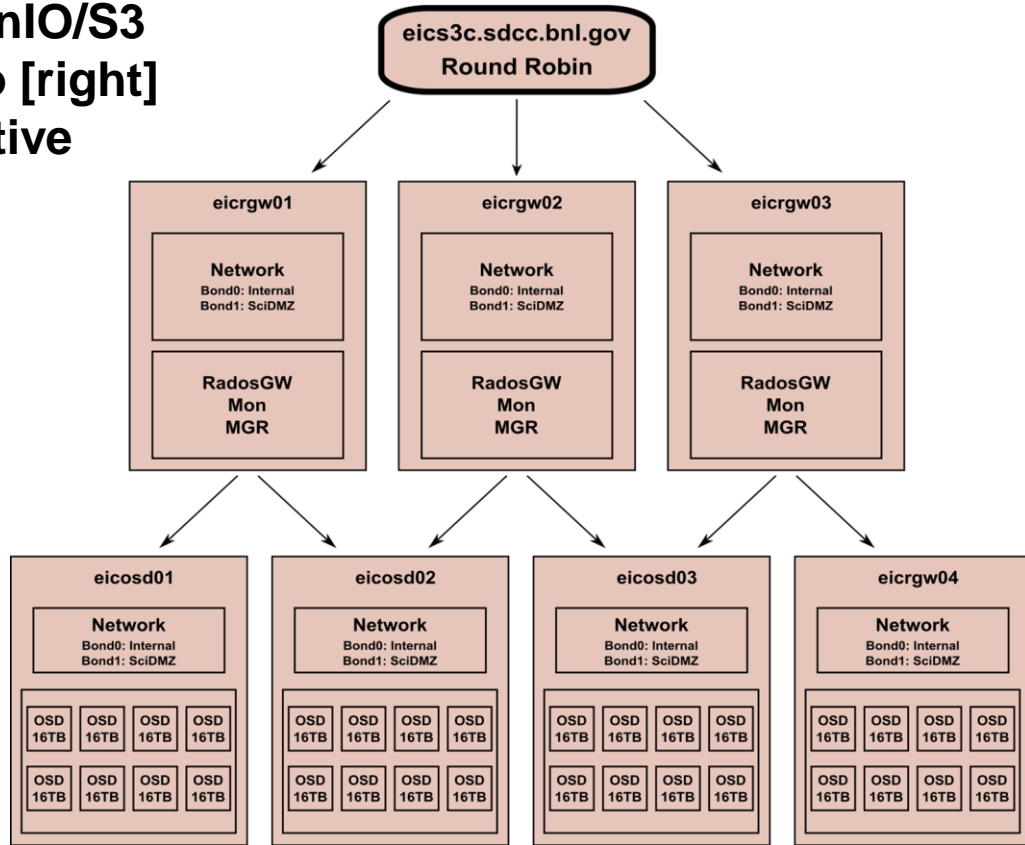


EIC MinIO GW S3



Architecture evolution from [left] minIO/S3 (HA Proxy) to [right] CEPH/S3 (native Rados GW)

EIC Ceph S3



Year 1, single host was kept for comparison



- Program Development funding => part of it was an extendable storage system able to support external contributions (providing an infrastructure with international storage resources / contrib).
- CEPH Object Storage supports native S3 interface and allows more features
 - Long-distance replication under the hood is possible with no need for a 3rd party data management tool
 - BEST feature: Geolocation (caching of part of the whole data) + would offer a way for international collaborators to contribute to a global federated storage pool (add storage from anywhere) and chose what to get or replicate. Single namespace!
 - Supports Federated ID access to storage (anyone from any accepted Federation could read/write without having an account at BNL or Jlab). Possible to support OSG SciToken considered.
 - CEPH/S3 is gaining momentum in the community (see CHEP 2023)
 - Object storage (believed to *enhance IO in HEP/NP*) - **ROOT RNTuple was tailored for Object Storage** [collateral benefit]
- Our current status
 - Only 450 TB of storage for now
 - Performance of a modest setup is ~5.5 GB/s (in exact alignment with predictions for raw IO performance and erasure coding). It is x2.6 better that the MinIO setup over Lustre.
 - Doubling the capacity would x2 the performance. Doubling the servers another x2.
 - Single Federation access to storage tested and seems to be working ([python SDK](#) / [boto3](#) tested)

BNL / JLab activity



- ePIC data is syncing between BNL S3 and JLab XRootD using a Data Transfer Node (DTN) at JLab
- A project, driven by ePIC, is aiming at integrating Rucio / FTS to ease data management
 - Set up rucio and its database running on http
 - Set up rucio and its database running on https using Jlab wildcard host certificate (cert)
 - Set up rucio and its database running on https using InCommon host cert
 - Set up fts and its database using using Jlab wildcard host cert
 - Set up fts and its database using using InCommon host cert
 - Set up xrootd without tpc (third party copy) support
 - Set up xrootd with tpc support using Jlab wildcard host cert
 - Set up xrootd with tpc support using InCommon host cert
 - Set up xrootd with http tpc support
- Jlab Internet-Facing Data Transfer Node (DTN) expected Milestones (in progress)
 - Deploy to Data Transfer Node, Test XRootd to BNL
 - Test FTS to BNL
 - Configure Rucio with BNL



EIC Joint Institute for Computing and Software

- BNL & JLAB, as co-host Laboratories, are establishing an EIC Computing and Software Joint Institute (ECSJI)
 - Operational October 1, 2023
 - Co-directors: Amber Boehnlein (JLAB) & Eric Lancon (BNL)
- The institute provides the EIC eco-system with a single interface for long-term support in matters of Computing and Software for the EIC project and the collaboration(s)
- The scope of the Institute includes
 - Execution of Host Lab responsibilities in a coordinated manner between BNL & JLAB
 - A coordinating body for interacting with international partners providing computing resources as in-kind contributions.
 - Maintaining Service Level Agreements and statements of work outlining the host labs' and international partner contributions
- Operational advantages:
 - Increased reliability & availability of resources
 - Leverage complementary expertise at the two Labs
 - Simplicity for the collaboration(s) and the users
- The proposed governance model will ensure communication, transparency, and assessment of the delivered services
 - In discussion with all stakeholders, including the Host Labs and the ePIC collaboration (POC: Markus Diefenthaler)

Host Labs responsibilities

Responsibilities of host Labs include and are not limited to the following:

- Assurance oversight functions for software and computing designs and execution.
- Supporting the ePIC developed Computing Model,
- Provisioning and operating standard infrastructure solutions consistent with supported Lab infrastructures,
- Support for the future EIC International Computing Organization,
- Support and computing resources for EIC activities that extend beyond ePIC, such as developing and designing other detector concepts and supporting the EIC accelerator modeling,
- Interface for local resources and policies at the respective Labs,
- On-going computing operations in support of the accelerator and detectors design and construction,
- Operational Support Functions for data processing and analysis, collaboration support, and simulations.

Conclusions

- JLab and BNL, via the ECCG, have coordinated, provisioned and put in operation a kernel of computing resources (CPU and storage) that have been (and remains) an integral part of the EIC detector design and simulation phase
 - Labs (for now) rely on sparse funding for supporting the EIC efforts (not always EIC earmarked). It is crucial to continue to focus on the EIC collaboration(s) requirements (iterate, revise, ...)
 - The resources “donated” to the EIC effort are (a) dedicated in some cases but (b) very much underused a short term work plan would go a long way to ensuring the resources remain allocated at the current level.
 - OSG Connect used preferentially (those jobs do not necessarily land on National Labs resources) - we need to ensure utilisation is maximized.
- Evolution
 - While the computing models crystalize, it is essential to think of the “*access from anywhere*” / “*contribute from anywhere*” as well as fostering a collaborative environment where all (international partners, ...) can contribute resources. CPU is easy, storage we may have “a” way.
 - Federated ID is a leading solution to address widely distributed / international collaborations. More services could be provided: Federated ID Mattermost? Federated ID content management system (Web site)? Federated ID Indico [done]? InvenioRDM (document repository)? ...
- Proposed ECSJI would be in place October 1st - single interface for long-term support in matters of Computing and Software for the EIC project and the collaboration(s)