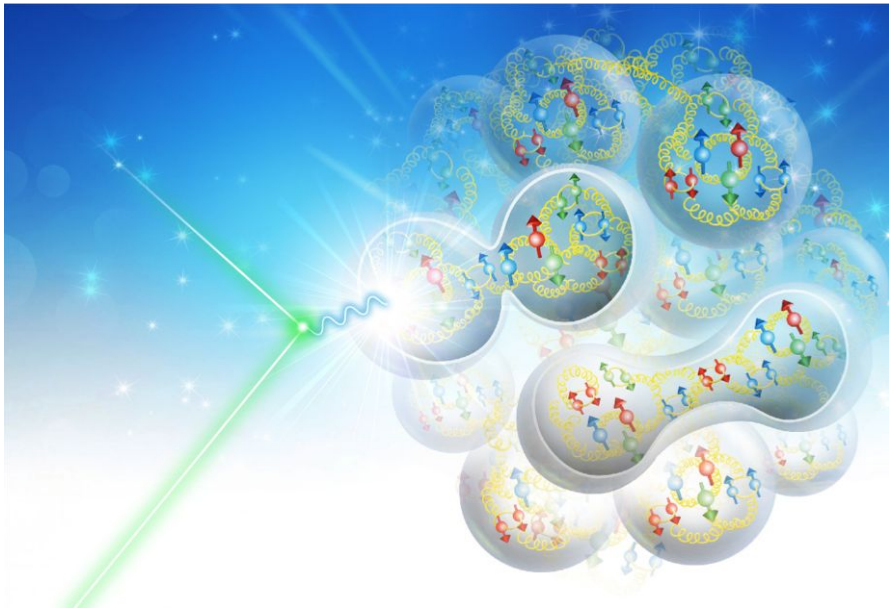
# Efforts at BNL/JLab on the OSG: the EIC and CEBAF

Mark Ito
Jefferson Lab
OSG All-Hands Meeting
September 3, 2020

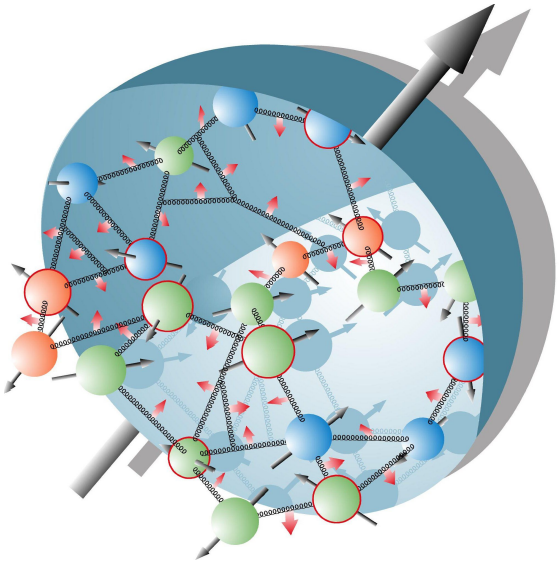
# Outline

- What is the Electron-Ion Collider (EIC)?
- How is computing organized?
- What components of OSG infrastructure are leveraged?
- Current use of the OSG at Jefferson Lab (JLab)
  - GlueX
  - CLAS12



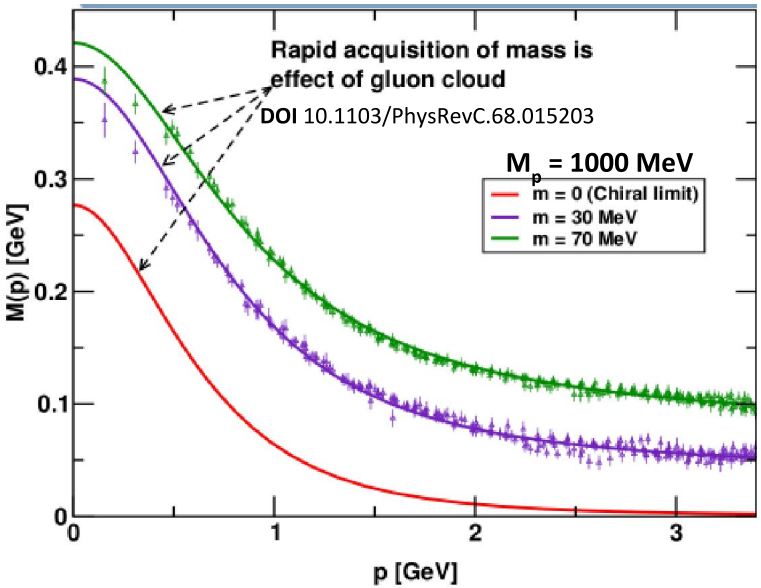
# The dynamical nature of nuclear matter

**Nuclear Matter** Interactions and structures are inextricably mixed up



**Ultimate goal** Understand how matter at its most fundamental level is made

**Observed properties** such as mass and spin emerge out of the complex system

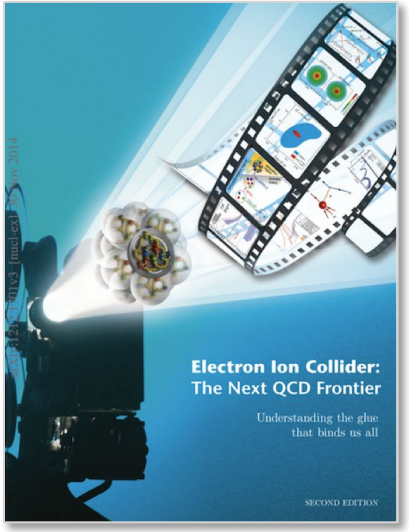


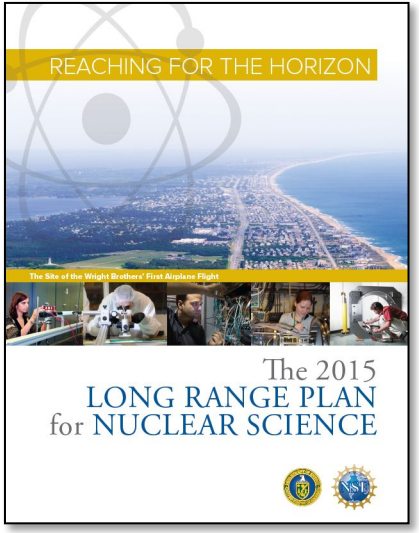

**To reach goal** precisely image quarks and gluons and their interactions

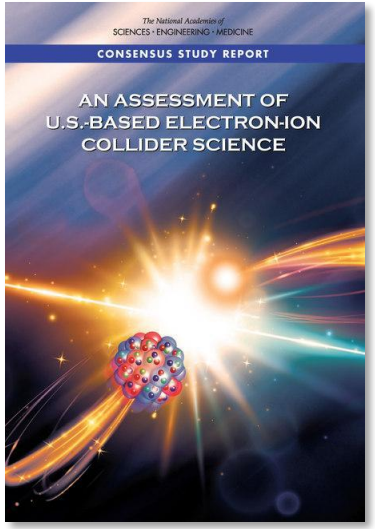


Jefferson Lab

# Why an Electron-Ion Collider?



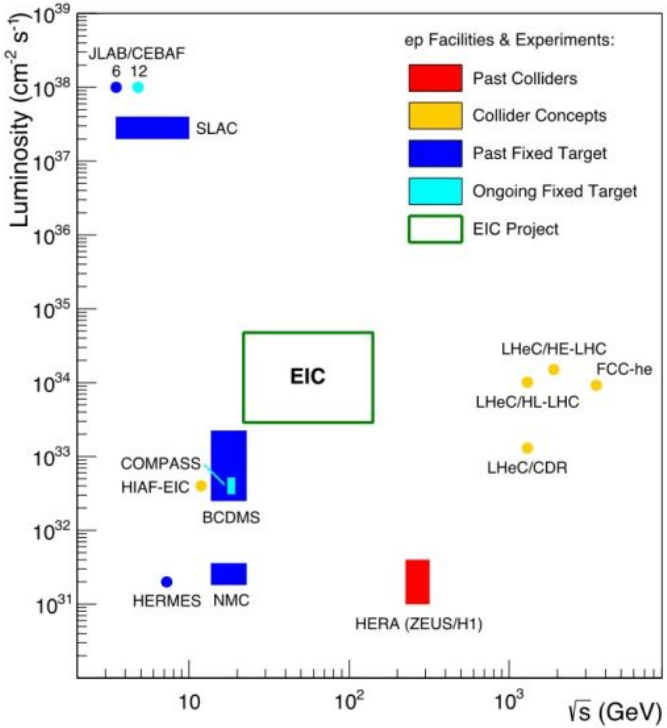




**Right tool**

- to precisely **image quarks and gluons** and their interactions
- to explore the new **QCD frontier of strong color fields in nuclei**
- to understand **how matter at its most fundamental level is made**.

**Understanding of nuclear matter** is **transformational**,
perhaps in an even more dramatic way than how the understanding of the atomic and molecular structure of matter led to new frontiers, new sciences and new technologies.

# The Electron-Ion Collider: Frontier accelerator facility in the U.S.



Department of Energy

## U.S. Department of Energy Selects Brookhaven National Laboratory to Host Major New Nuclear Physics Facility

JANUARY 9, 2020

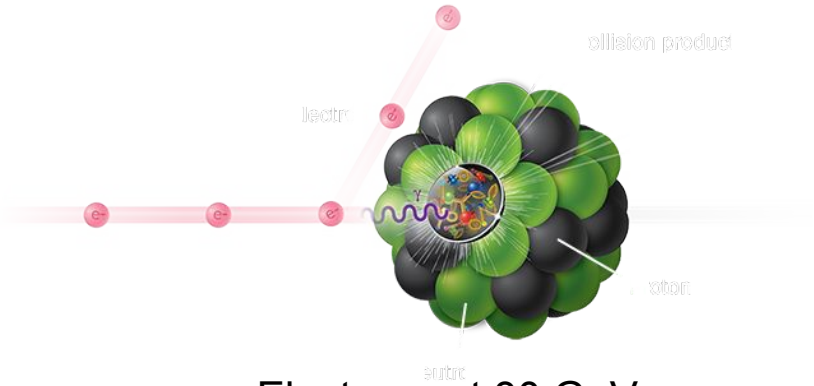


Thomas Jefferson National Accelerator Facility in Newport News, VA will be a major partner in realizing the EIC, and several other DOE laboratories are expected to contribute to EIC construction and to the groundbreaking nuclear physics research program that will be accomplished there.

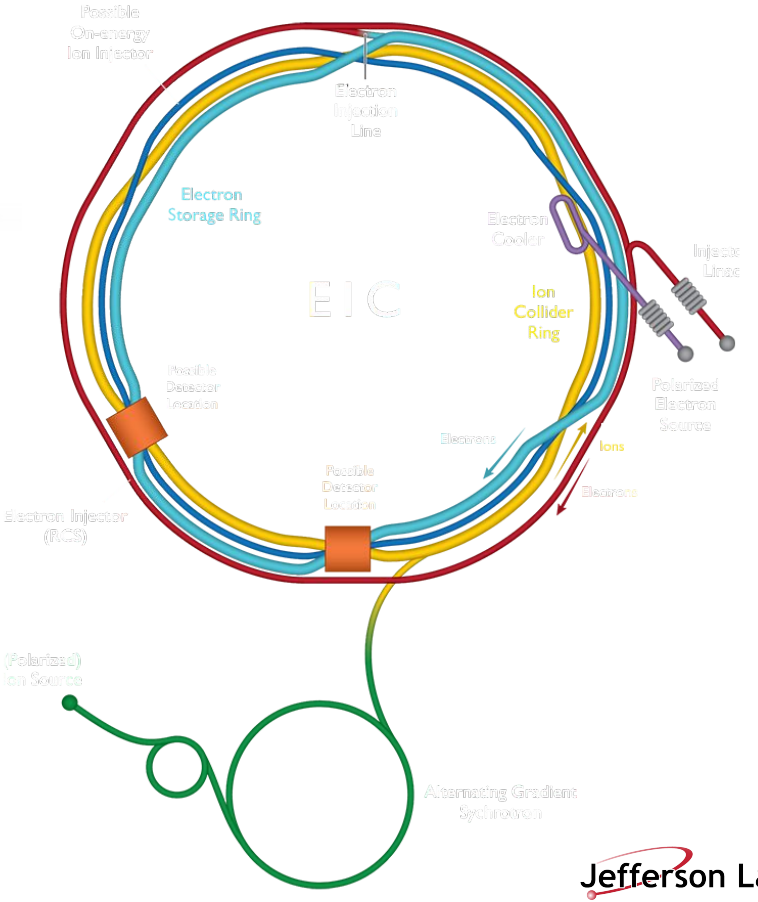


Jefferson Lab

# Schematic Views of the EIC



- Electrons at 30 GeV
- Protons at 250 GeV
- or Ions at 100 GeV/nucleon
- Polarized

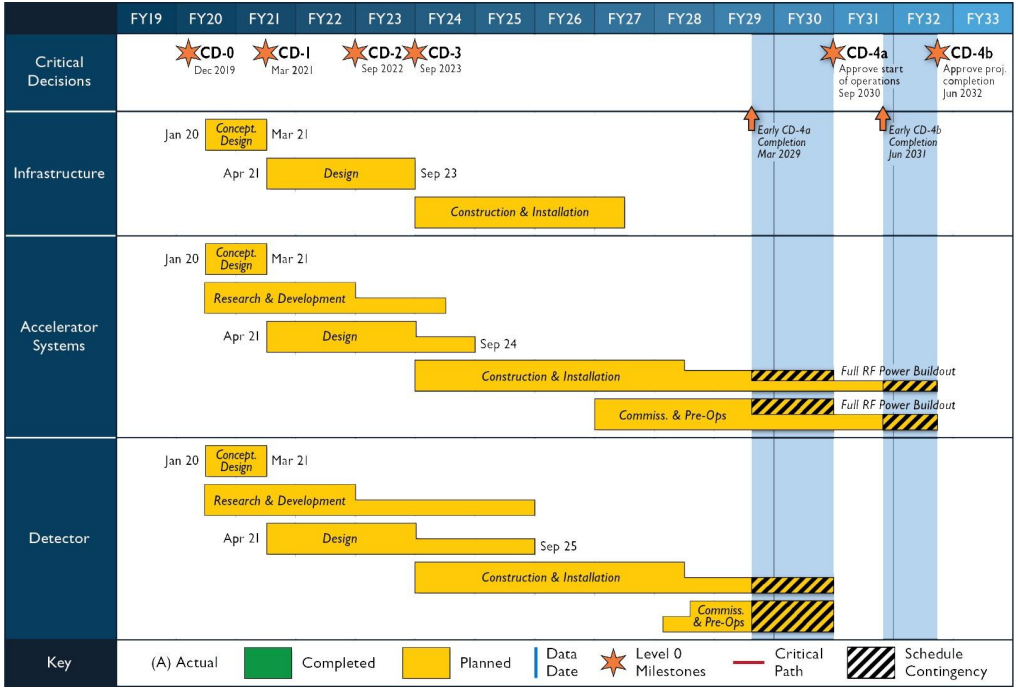



Physics Topics:

- 3D imaging of the proton
- Solving the proton spin puzzle
- Quark and gluon confinement

# Project Status

- After decade-long competition with designs for BNL and JLab, site selection complete: eRHIC at BNL
- Effort is co-managed
  - Computing coordinators:
    - BNL: Jerome Lauret
    - JLab: Graham Heyes
  - Software coordinators:
    - BNL: Torre Wenaus
    - INFN Trieste: Andrea Bressan
    - JLab: Markus Diefenthaler



# Federated Computing @ EIC

- "Federated": add resources at any end-point, available to all (if site accepts jobs)
  - Flexible submit host setup: Users at both JLAB and BNL have the option to run in both proprietary EIC resources and OSG opportunistic resources.
  - Usage of owned EIC resources not counted against the priority on the OSG opportunistic pool.
  - Sites have maximum flexibility: they can either support OSG, EIC or both.
  - Software provisioning: CVMFS – Stratum 0 at BNL fully operational
  - Pending: Federated storage (Stashcache / XCache), CE, Stratum 1 at JLab
- BNL/JLab would form the "core" of an East-Coast (Federated) Super-Facility
  - There is a need for a "fat" pipeline between the 2 labs!
  - Processing may happen anywhere in the Federation.

# OSG Developments at JLab

- After rejuvenating the JLAB VO and giving access to several experiments (CLAS, JAM, SOLID) there are three main thrusts
  - CVMFS (Oasis)
  - XRootD (Stash Cache)
  - Running OSG jobs at JLab

# CVMFS

- GlueX has been successfully using CVMFS (Oasis) for the last few years
  - Holds the software stack

- Set up the JLab VO to use Oasis and provided a seamless means for users to distributed their software stack for use on the OSG

# XrootD

- JLab had been successfully running an XrootD server used by the GlueX and CLAS12

- Other groups interested in using XrootD
  - and/or Stash Cache

- Rolling out a Stash Cache Origin to allow users to seamlessly manage their data

# Running at JLab

- With some core infrastructure work done JLab is planning to allow some OSG jobs to run at JLab
- It will involve a phased deployment in 2 steps:
  - 1) Run only jobs from the GlueX and JLab VO (who are credentialed at JLab)
  - 2) expand to allow EIC members to use the OSG to access JLab computational resources (both JLab and outside contributors)

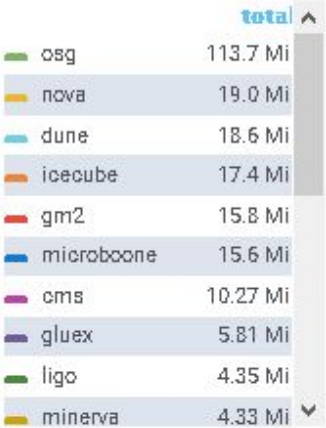
Open Science Grid

# MCwrapper

- MCwrapper manages Monte Carlo production for **GlueX**
- Widely used in the collaboration and works at a variety of institutions
- MCwrapper's autonomous part takes requests via a web form and automatically produces and monitors a workflow to fulfill the request.
    - Primarily runs on the Open Science Grid (OSG)
- To date MCwrapper has autonomously produced more than 13 Billion simulated events with more than 2.1 Million jobs and over 4 million core hours. Totaling over 52TB of simulations

## VO OSG Core hours Top 10

## (Last ~6 months)




## Projects

Show 10 entries

Search:

| Progress % | ID | Email | Status | RunNumLow | RunNumHigh | NumEvents | Generator | BKG | RCDBQuery | VersionSet | ANA_VersionSet |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 589 | | | 30381 | 30381 | 100000 | genr8 | None | | version_recon-2017_01-ver03_8.xml | None |
| 98.2 | 590 | | | 30803 | 30803 | 10000000 | gen_omega_3pi | Random:recon-2017_01-ver03 | | version_recon-2017_01-ver03_8.xml | |
| 93.4 | 594 | | | 30496 | 30496 | 10000000 | genr8 | Random:recon-2017_01-ver03 | | version_4.5.0.xml | None |

Showing 1 to 3 of 3 entries

Previous 1 Next

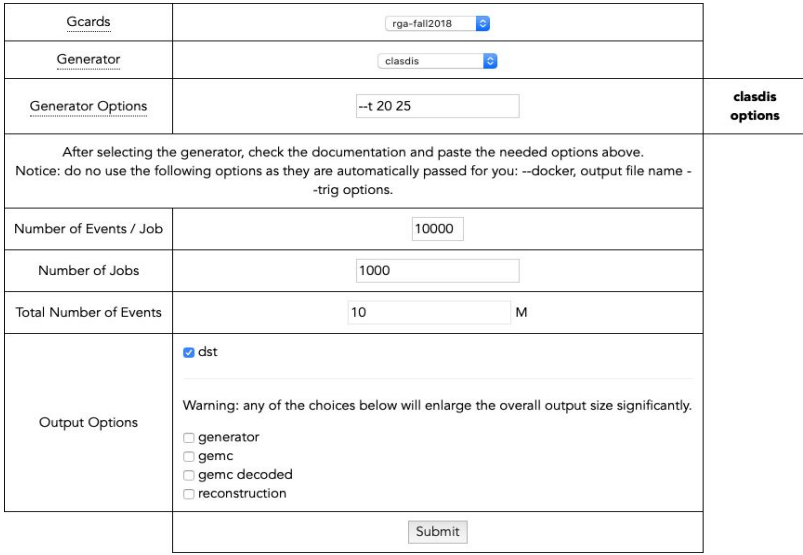
# OSG CLAS12 Monte-Carlo

CLAS12 jobs submitted to OSG through a web portal:

## gemc.jlab.org/web_interface

- different experiment configurations available through dropdown menu
- use of container generators or generated files
- xrootd access to background files @JLAB for background merging
- cvmfs access to CLAS12 software and data such as magnetic fields

| name | executable name | output ok |
|---|---|---|
| clasdis | ✅ | ✅ |
| claspyth | ✅ | ✅ |
| dvcsgen | ✅ | ✅ |
| genKYandOnePion | ✅ | ✅ |
| inclusive-dis-rad | ✅ | ✅ |
| JPsiGen | ✅ | ✅ |
| TCSGen | ✅ | ✅ |




**Generators**:

- Collected in the container
- Have to satisfy certain requirements to be included

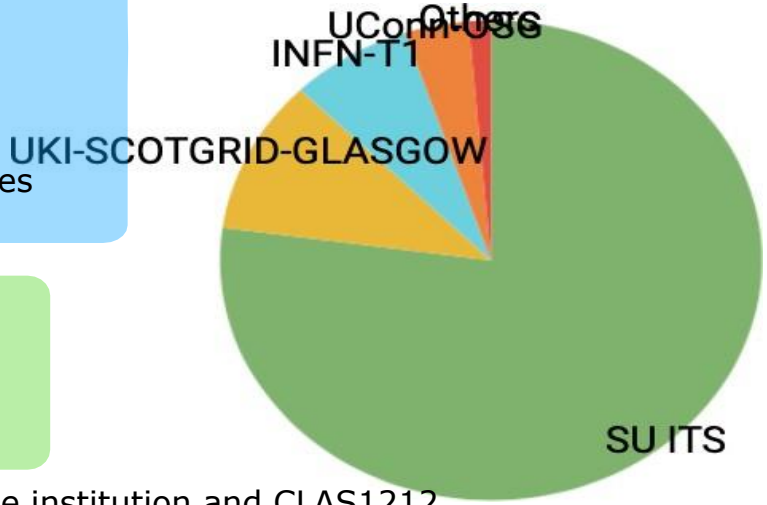
# OSG CLAS12 Monte-Carlo

Dedicated, guaranteed:

- INFN-T1 ~500 nodes
- SCOTGRID ~400-700 nodes
- GRIF (tested, upcoming) ~400 nodes

Semi-dedicated:

- UCONN 0-1000 nodes
- Syracuse 0-10000 node
- Opportunistic: 0-5000 nodes

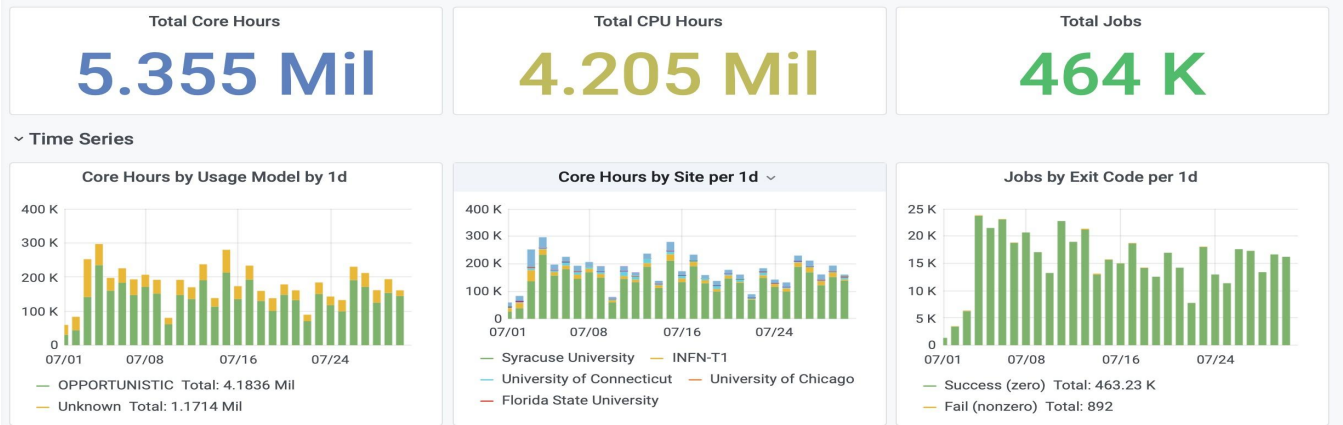



Priority system in place:

- User do not have to wait for queue to complete for theirs to start
- High priority queue

More upcoming: advantageous to both the institution and CLAS1212.

July Usage



# Summary

- Electron Ion Collider will open an exciting new scientific frontier
  - Design and construction to proceed over the next decade
- EIC management a partnership between BNL and JLab
  - Includes federated computing space
  - Joint management of computing and software design/development
- JLab moving ahead with OSG-based capabilities