

Scientific Computing at JLAB

Amber Boehnlein

IT Division

Thomas Jefferson National Accelerator
Facility

HSF
UCSD
January, 2017

Jefferson Lab

Jefferson Lab is a nuclear physics user facility in Newport News, Va

— Continuous Beam Electron Accelerator Facility (CEBAF)

- 6 GeV Era: 173 Experiments completed over 17 years
- 1500 users
- ~500 Ph.Ds awarded
- Tight coupling between Theory and Experiment
 - LQCD major driver for the computing program

CEBAF has been upgraded to 12 GeV

- Upgraded/Upgrading the Experimental Halls
- Data collection began last fall for two Halls
- Two major experiments will be constructed
- Over subscribed: 10 year scientific program is already planned

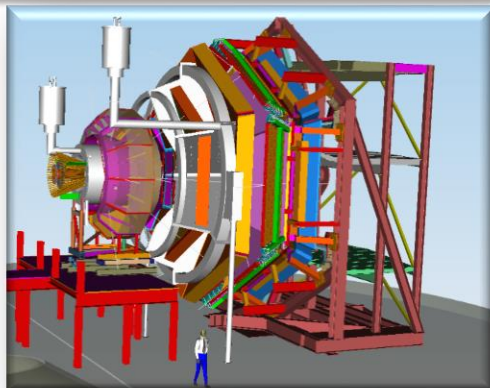
Electron-Ion Collider (EIC) next up after completion of the Facility for Rare Isotope Beam

- National Academy Study in progress
- FRIB construction will complete around 2022

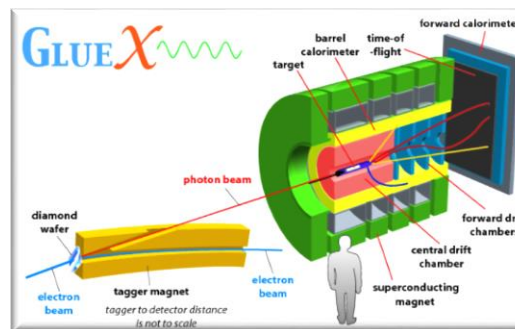
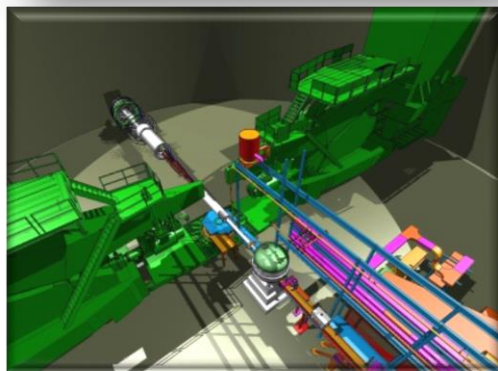
Physics Drivers for the 12 GeV Upgrade

- Explore the physical origins of quark confinement (GlueX)
- New and revolutionary access to the spin and flavor structure of the proton and neutron
- Discovering the quark structure of nuclei
- Probe potential new physics through high precision tests of the Standard Model

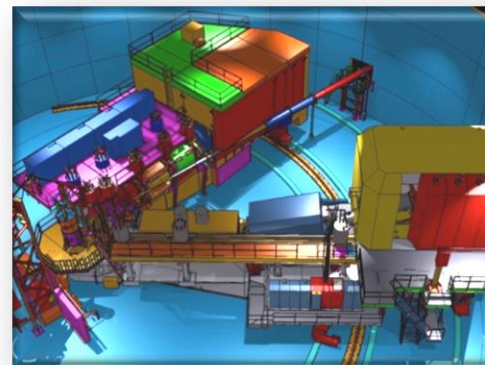
Hall B –
understanding
nucleon structure
via generalized
parton
distributions



Hall A –
form factors,
future new
experiments (e.g.,
SoLID and
MOLLER)



Hall D –
exploring origin
of confinement
by studying
exotic mesons

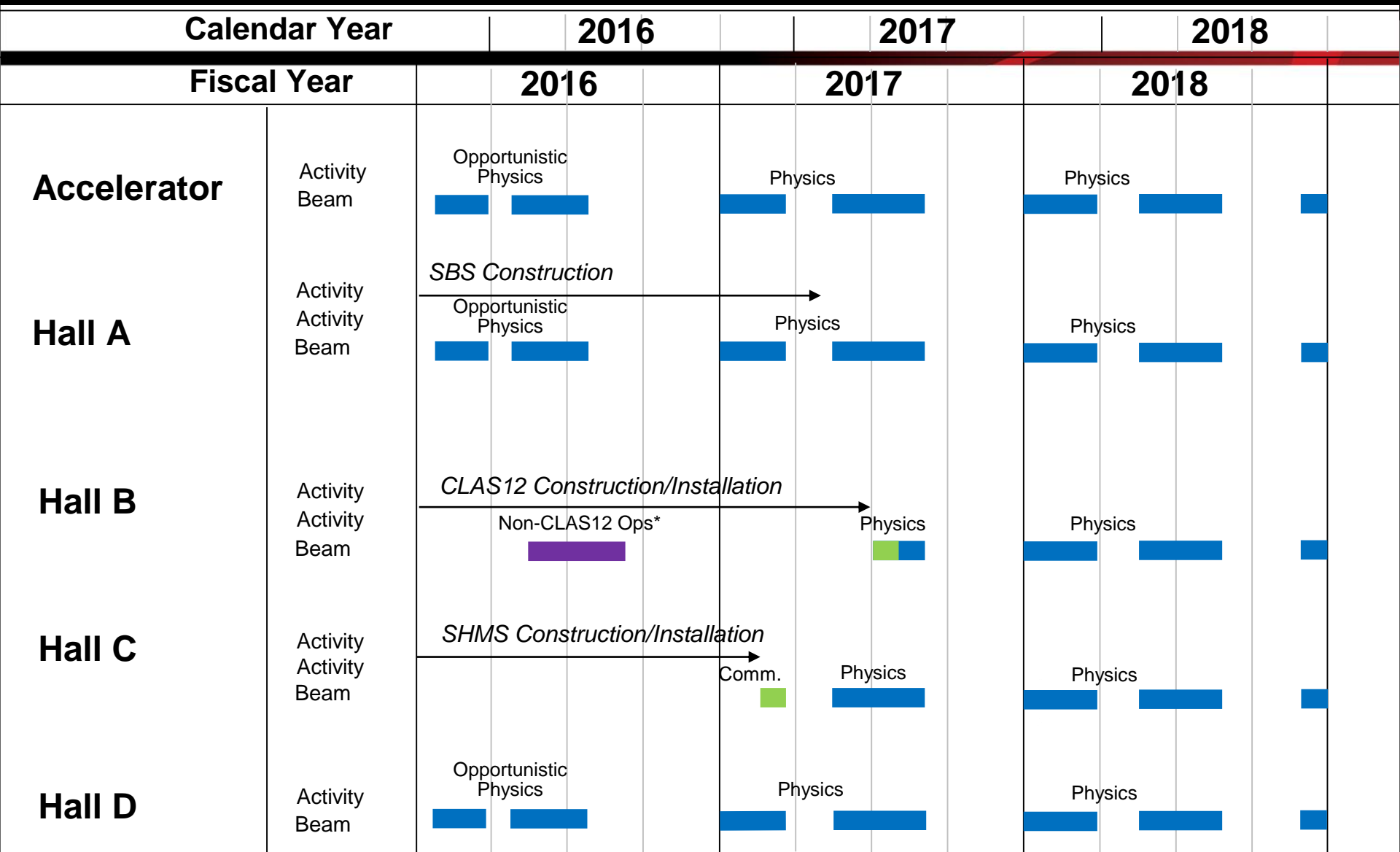


Hall C –
precision
determination
of valence
quark
properties in
nucleons
/nuclei

Scientific Computing At JLAB

- Physics computation -Large Scale Parallel Computing
 - Lattice QCD: JLAB known for Science; software; hardware
 - Accelerator Simulation
- Experimental computing is coordinated by Physics Division
 - Physics Division Staff often play lead roles
 - Relatively small efforts: in the few FTEs
- Scientific Computing group in IT supports
 - Hardware: Disk; Tape; Compute clusters;
 - Runs key services: the batch system
 - Provides in-house tools: eg monitoring and workflow
 - Provides technical support and expertise, especially for parallel computing
- IT provides cyber security, networking and desktops

CEBAF Three-Year Schedule



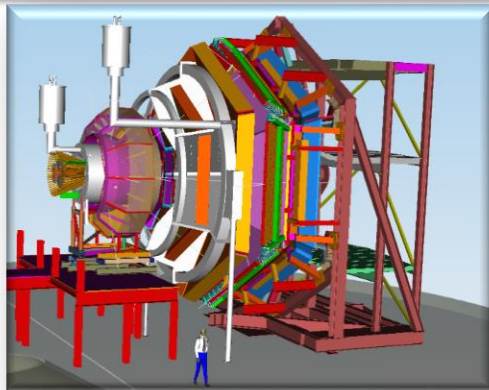
* PRad Summer 2016 run

■ Beam for Commissioning
 ■ Beam for Physics
 ■ Non-CLAS12 Ops

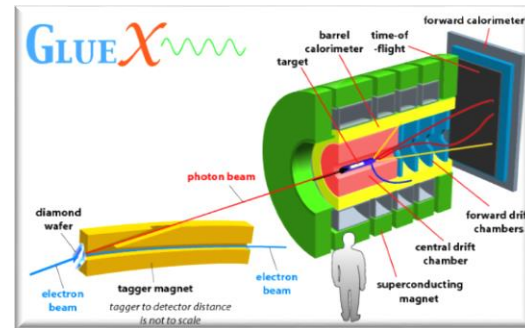
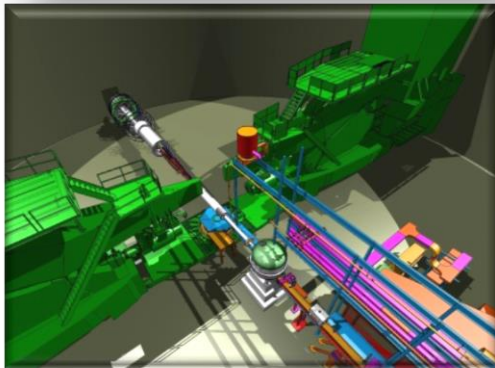
Experimental Computing

- Each experiment is responsible for their own software with the Hall.
- Hall B and Hall D have standing collaborations; Hall A and C collaborate with each other
- Hall D/GlueX is new—strong focus on getting results to publication
- 12 GeV era introduced MC data challenges, nightly builds.
- Working to make connections outside lab—best practices and maybe tools.

Hall B –
JAVA CLARA
Framework
GEMC (Geant4)

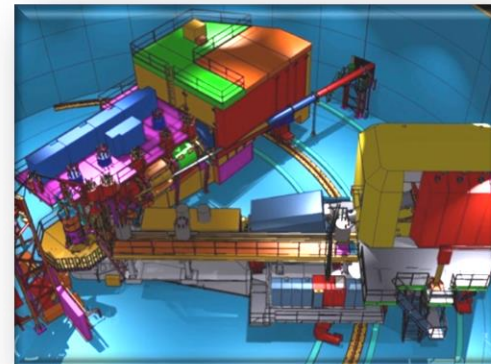


Hall A –
Individual
experiments
In house built
ROOT framework
~10-40 TB/week

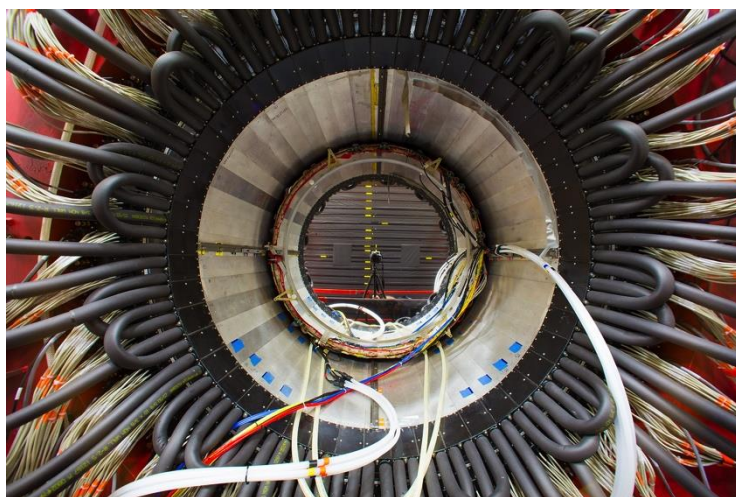


Hall D –
C++ JANA
framework
OSG
Geant4
ROOT

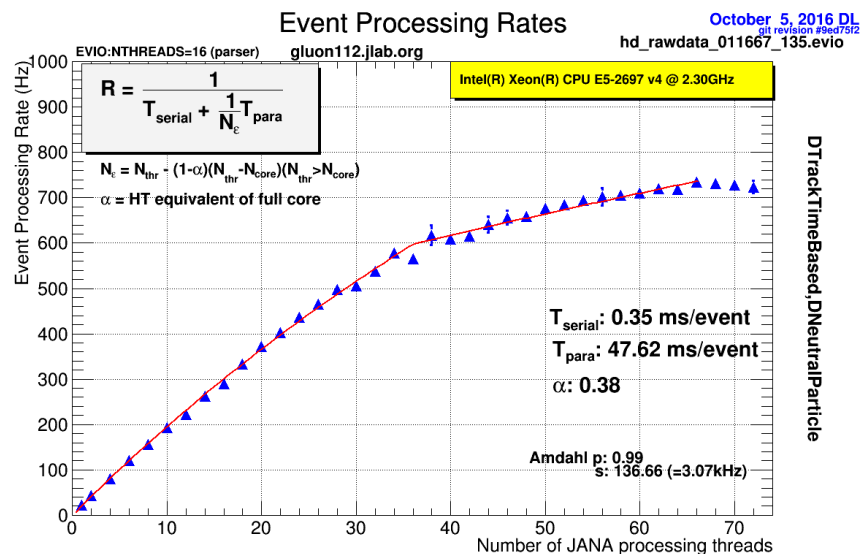
Hall C – In
house built
ROOT
Framework
Collaborates
with Hall A
1.5TB/week



Parallel Event Reconstruction in GlueX



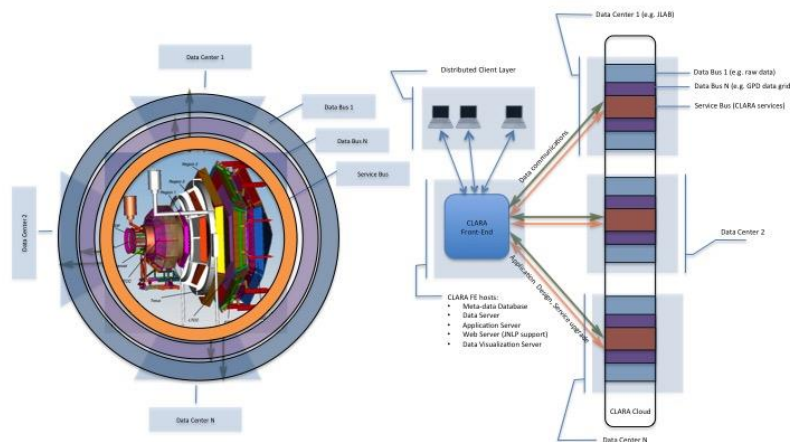
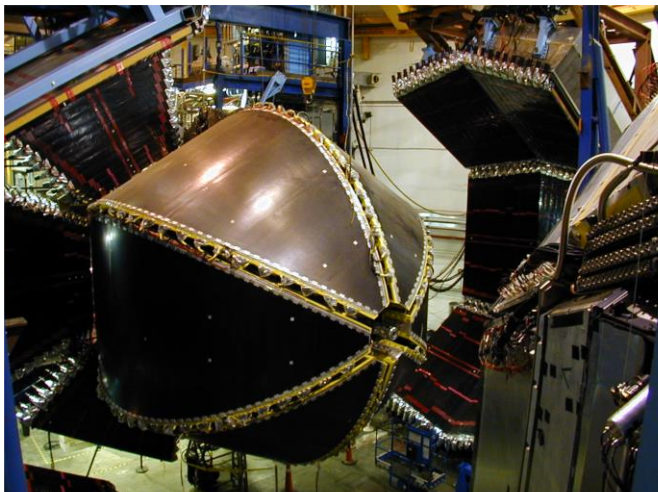
The GlueX experiment in Jefferson Lab's Hall D was designed to look for exotic hybrid mesons, a previously unobserved form of matter. The experiment will produce as much as 10 PB of data a year.



JANA is a highly optimized event reconstruction framework designed at Jefferson Lab for leveraging the parallel computing power of modern day many-core processors. This plot shows the excellent scaling of the system, even through the hyperthreading region.

CLARA Data-stream Processing Framework

Microservices architecture for data analytics



The CLAS12 detector from the Thomas Jefferson National Accelerator Facility's (JLAB) large-Acceptance Spectrometer (CLAS) meets the basic requirements for the study of the structure of nucleons and nuclei within the 12-GeV upgrade program. The major focus of CLAS12 is to study the generalized parton distributions (GPD). CLAS12 is design to carry out the core program for the study of the internal dynamics and 3D imaging of the nucleon and quark hadronization processes.

JLAB's CLARA (Clas12 analysis and reconstruction) framework makes data provisioning, processing and preservation agile and efficient. CLARA enables custom data processing tools to be presented as services and be integrated into the data processing workflow in a fraction of the time previously required. Inherent modularity and loose coupling of the system simplifies software long term maintenance. Elasticity of CLARA makes efficient use of multi-core architecture systems as well as commercial cloud resources.

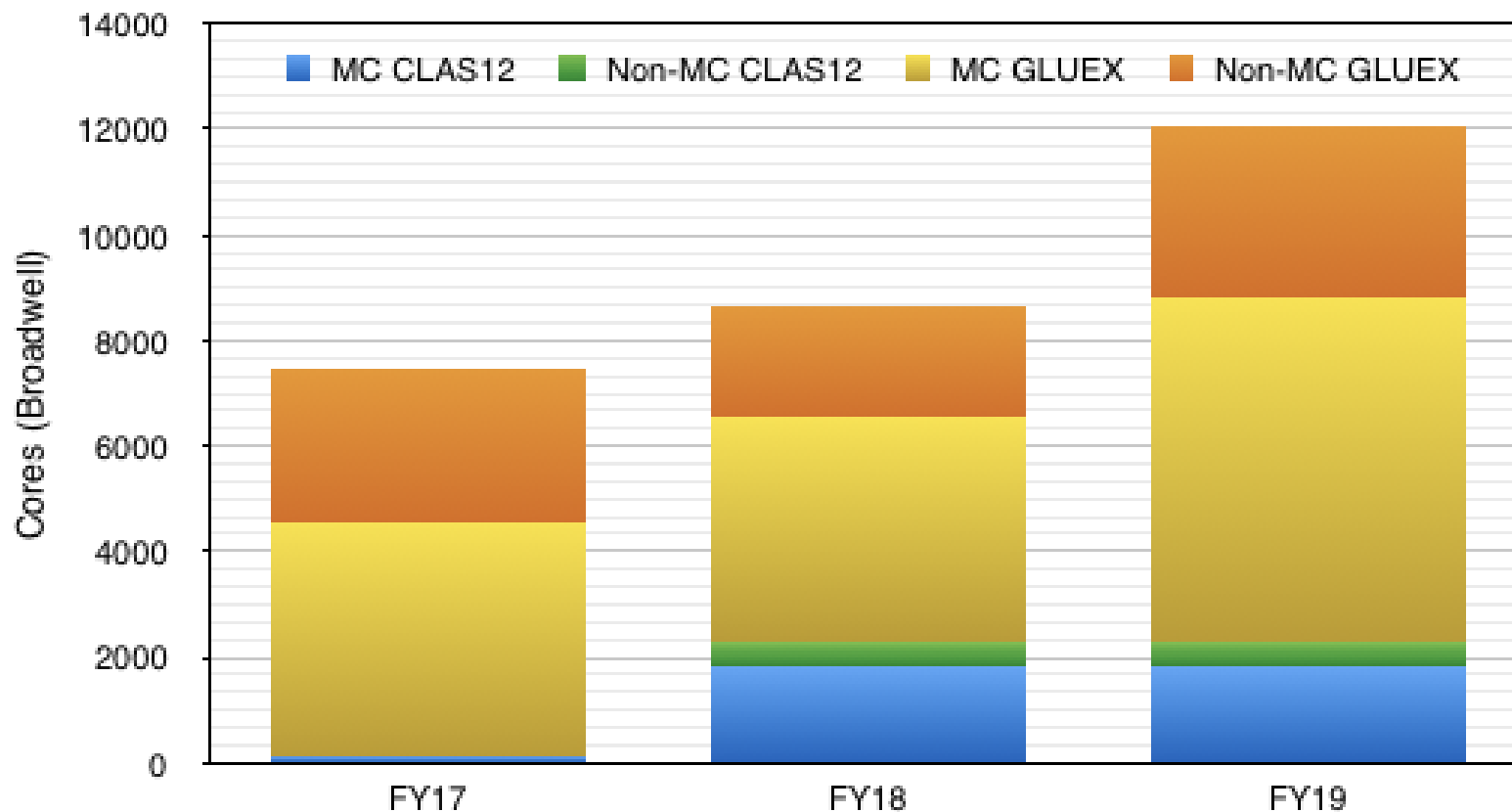
Local Hardware Resources

- 830 servers run by IT in SciComp
 - IT provides interface to the batch system (SWIF)
- Serves the LQCD program and Experimental Physics
- SciPhy Cluster: 397 on top 500
 - 17,152 Xeon Phi KNL cores (LQCD; i.e. 200 KNL nodes)
- 10,000 conventional Intel cores (6,000 LQCD, 5,000 Physics)
- 328 Nvidia GPUs
- 12 Petabyte tape library (4 LQCD, 8 Physics)
- 2.2 Petabyte disk (1.2 LQCD, 1.0 Physics)
 - LUSTRE
- Network: 10GB to ESNET
 - Collaboration with local universities
 - Most external traffic currently is LQCD



Three Year Compute Estimates

Chart showing change in compute requirements



Nucleon Tomography

Major Scientific goal of 12 GeV program

Construct 3D maps of the nucleon by performing global fits of world wide experimental data

- General Parton Distributions (GPDs)
- Transverse Momentum dependent functions (TMDs)

Emerging field: interplay required to further Fundamental Theory, Experiments and Phenomenology and Computational/Computer Science

Multi-disciplinary for the WIN!!

- Large Scale Data Base to gather and represent the disparate data sources
 - Different kinematic ranges; different precision
- New data will constantly be added from experiments; data management and access; data preparation and 'preservation'
- Easy to use workflows-capturing the parameters
- Development of appropriately parameterized theoretical models
- Development of mathematical methods for numerical integration; for appropriately treating the errors; for weighting data sets
- Development of visualization and functional representations

EIC Computing

Golden opportunity to completely rethink computing model.

- Focus on novel uses of computing and detectors-Petascale at the beam line
- Advances software: algorithms; methodology
- *Focus on usability*

Current EIC priorities are Monte Carlo Detector Simulations for design studies

- Many different groups/many different solutions
- Settling on interoperability of data format
- Highest priority is an easy to use fast simulation.

Upcoming Workshops

**FUTURE TRENDS IN
NUCLEAR PHYSICS
COMPUTING**

Jefferson Lab, Newport News, VA
MAY 3-5 SYMPOSIUM TO BE HELD MAY 2

The workshop "Future Trends in Nuclear Physics Computing" is held from May 3 - 5 at Jefferson Lab. The goal of the workshop is to discuss trends in scientific computing and to collect ideas on how to improve analysis workflows at existing nuclear physics programs as well as to work towards analysis techniques and tools for future projects like the Electron-Ion Collider. We will not only review best practices in scientific computing and the future development of existing analysis techniques and tool, but also to examine how future computing trends like Big Data or machine learning could improve the productivity at existing or future experiments.

ORGANIZING COMMITTEE
Wes Bethel (JLab)
Amber Boehnlein (JLab)
Kyle Cranmer (NYU)
Markus Diefenthaler (JLab)
Graham Hayes (JLab)
Alexander Kiselev (BNL)
Naomi Makins (JUIC)
Katherine Riley (ANL)
Thomas Rockwell (FRIB/NSCL)

WWW.JLAB.ORG/CONFERENCES/TRENDS2017

Jefferson Lab

Summary

- **12 GeV-era Research Program underway:**
 - Computing for GlueX (Hall D) has proven successful
 - CLAS-12 (Hall B) running readiness tests
 - Sharing hardware between LQCD and Experimental physics beneficial in adapting to modern architectures
 - Depend on standard tools: Geant4; ROOT
 - Are adding distributed resources to the mix
- **Realizing the 12 GeV Physics program will require combinations of experimental results**
- **Working to make connections beyond JLAB**
- **Planning for EIC computing is underway**

BACK-UP

