Scientific Computing At Jefferson Lab

Amber Boehnlein
Chief Information Officer
IT Division Director

Jefferson Lab

U.S. DEPARTMENT OF ENERGY | Office of Science

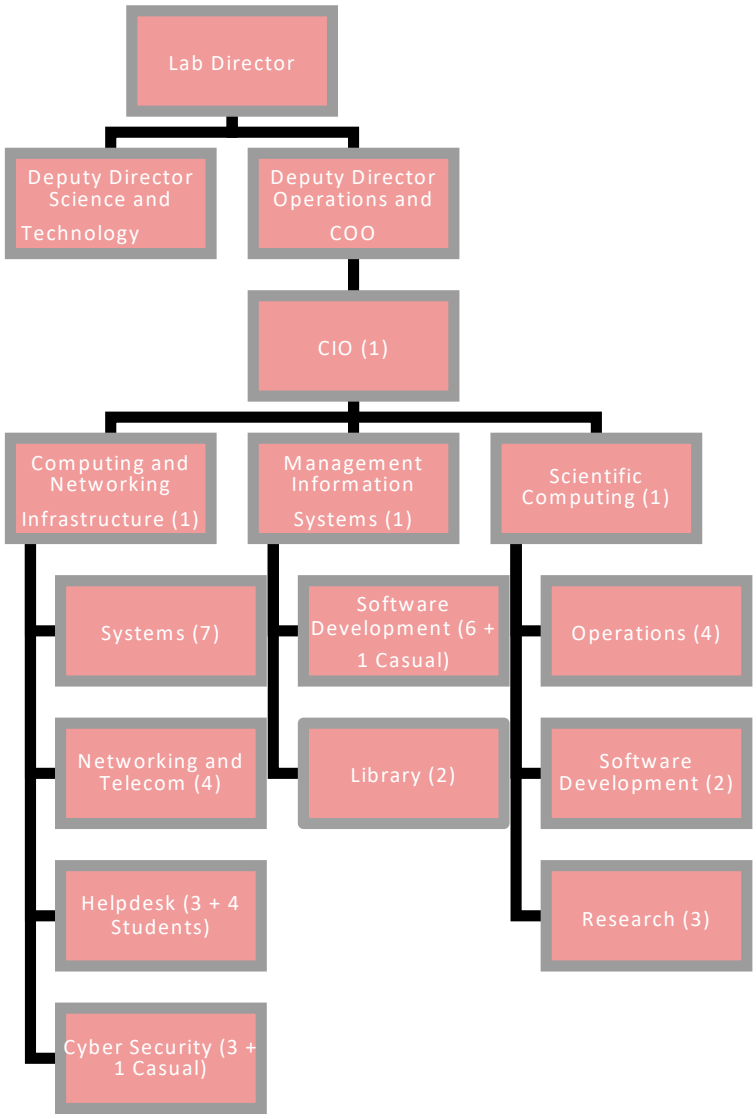
JSA

Outline

- The IT Division
- Computing Resources for 12 GeV
- The (nascent) Scientific Computing Strategy

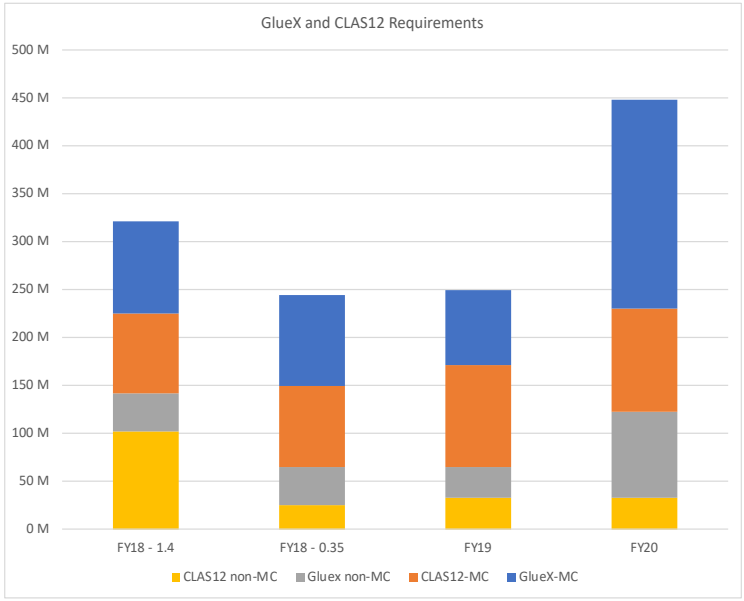
IT Organization

- Overall IT Staff
  - 38 Full Time Staff
  - 1 Casual
  - 4 Students
- Responsible for
  - Networking and Telcom
  - Computing Infrastructure
  - Business Systems
  - Cyber
  - Scientific Computing
- Within Jefferson Lab close collaborations with
  - Theory on LQCD
  - Physics with Experimental Computing
  - Accelerator Division



Experimental Computing Planning

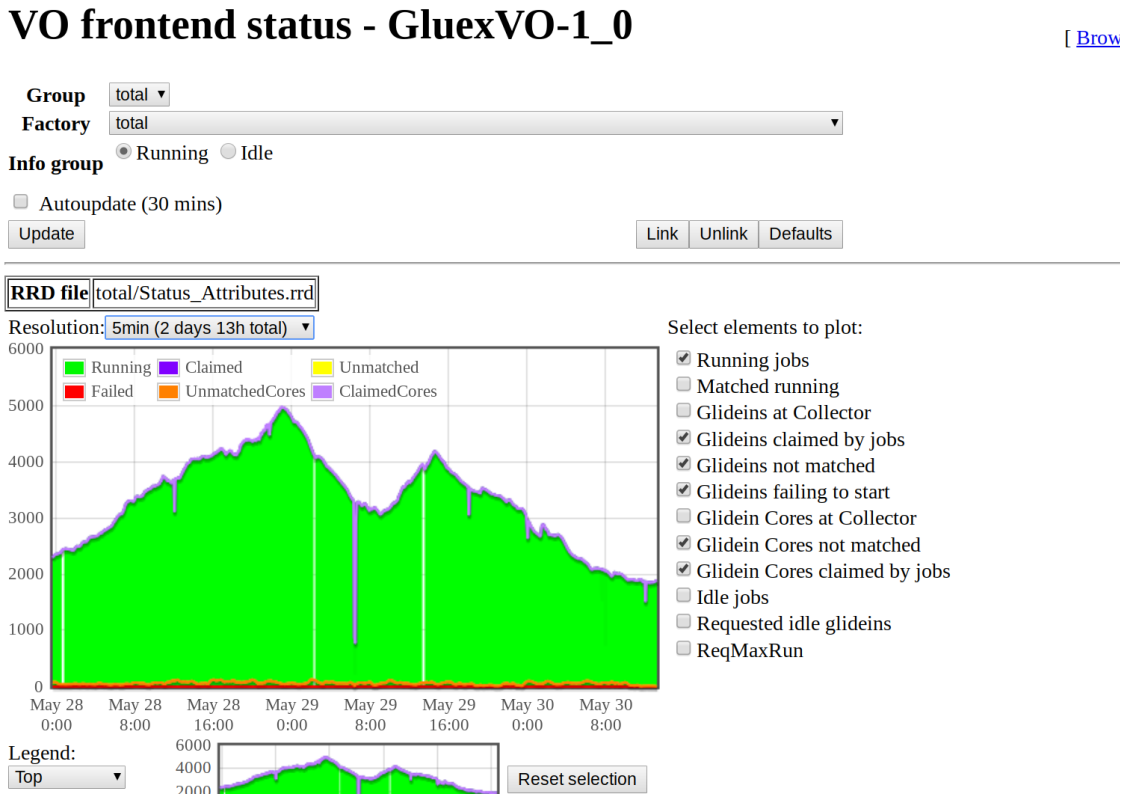
- Computing necessary to support the the experimental program
  - Estimates based on spreadsheets of key parameters benchmarked against actual performance
  - Local farm resources
  - Disk, Tape and Networking to support distributed resources
- Making major investments in compute, tape drives, disk including SSD buffer



	Current	FY19	FY20
CPU (M-core-hours/year)	37	87	90
Scratch Disk & Cache Disk (PB)	0.65	1.1	2
Tape (GB/s)	3	5	7
WAN bandwidth (Gbps)	10	10	10

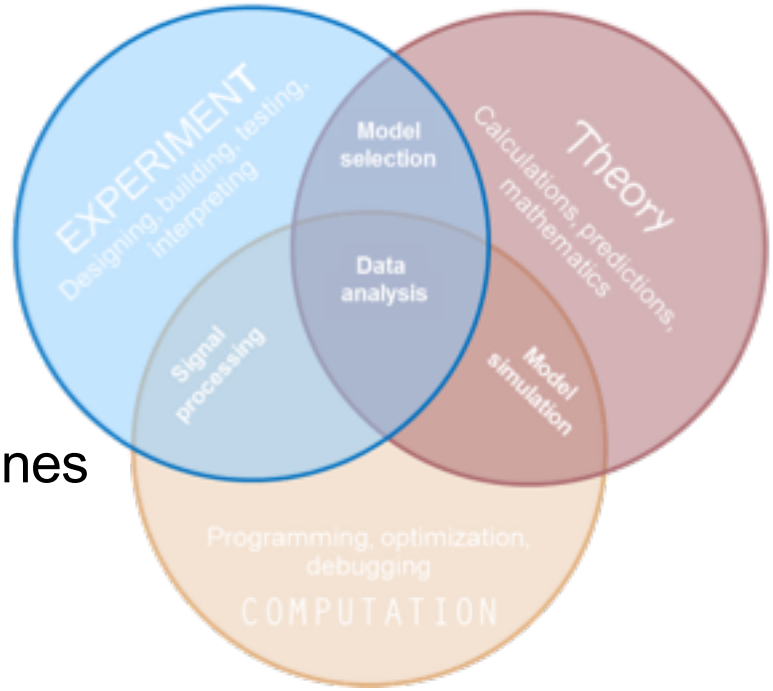
Distributed Computing Resources

- Computing at Jefferson Lab Users' institutions have always been essential.
- For 12 GeV, Distributed Computing is part of the planning
- GlueX Monte Carlo Production on the Open Science Grid (OSG)
  - Led by Richard Jones (UCONN)
  - Have seen resources comparable to local farm
- Tests of GlueX reconstruction at the National Energy Research Science Computing Center (NERSC)
- Prototyping and beta users on Amazon Web Services



Scientific Computing Strategy For Jefferson Lab

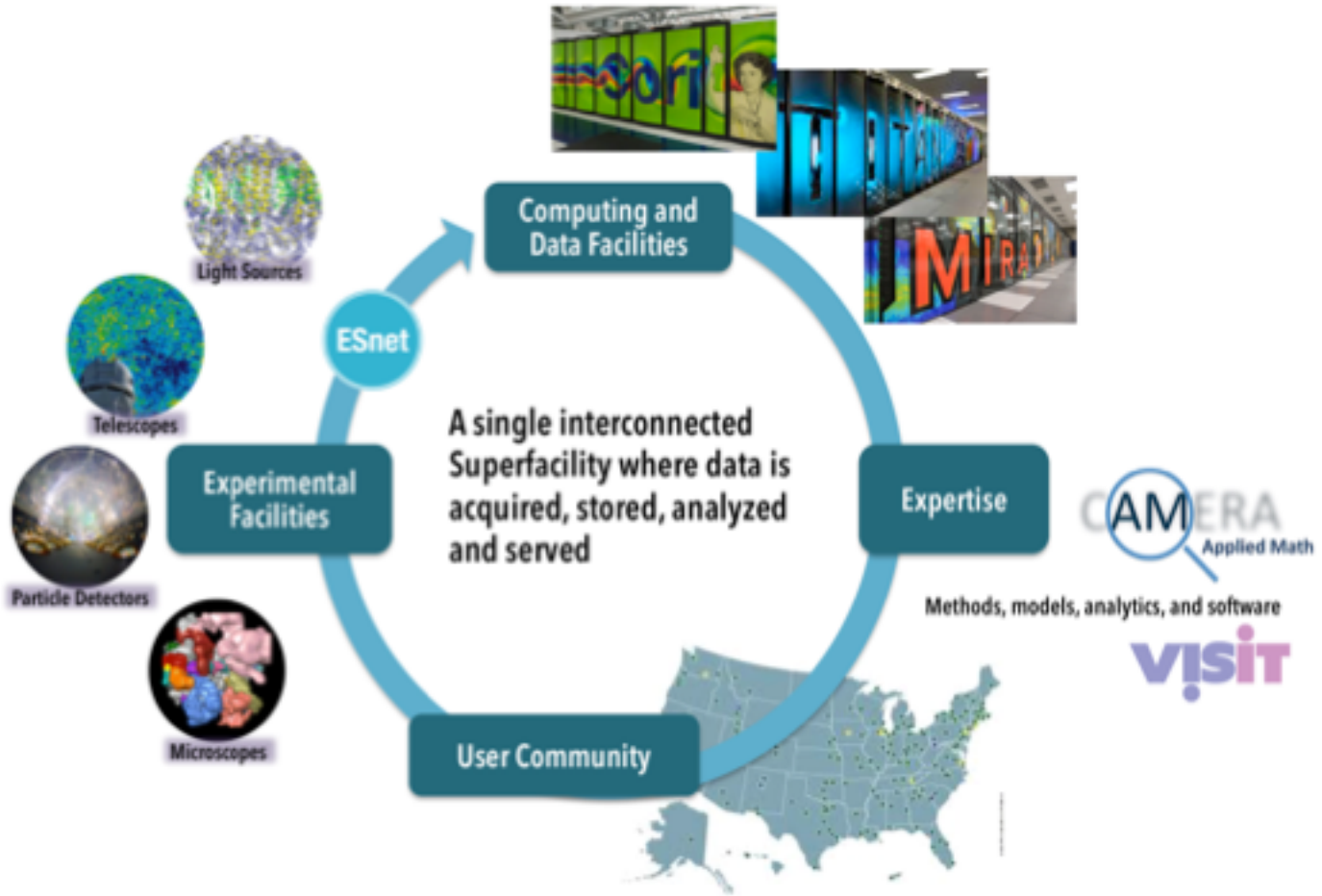
- The richness of the Jefferson Lab science program will spur development in theory, experiment and phenomenology
  - 12 GeV era: Multi Channel & Multi Dimensional Science
- Goal: Develop computing and computation for the success of the 12 GeV Physics Program that transitions toward the EIC era with computational science as a pillar
  - The era of 12 GeV Science is coinciding with a revolution in computational techniques and disciplines
    - Machine Learning
    - Data Science
    - Exascale



- Strategy: Contribute to Jefferson Lab Science by expanding and developing unique computational science through *collaborations*.

Application Instituted & Super Facility

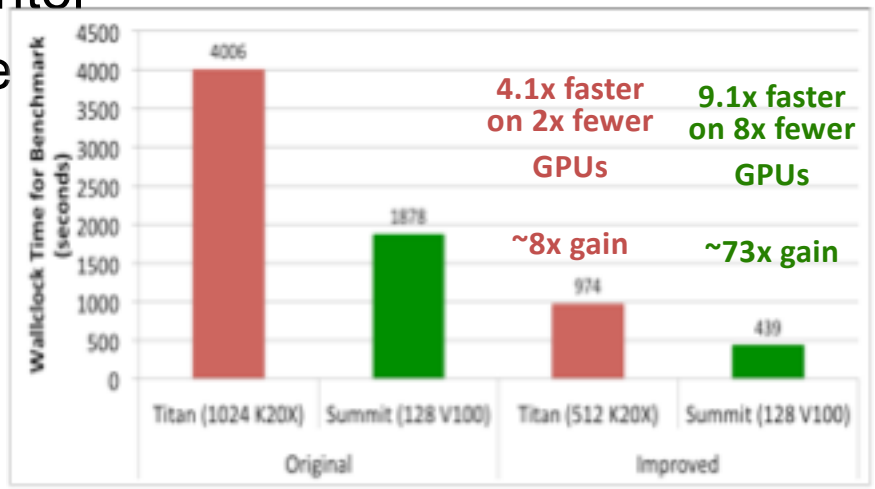
- Science Application Institutes
  - Cross Disciplinary and often University lead
  - Scientific Discovery Through Advanced Computing (SciDAC)
- Super Facilities are two or more coupled DOE Scientific User Facilities



Lattice Quantum Chromodynamics

- Discovery of the properties of hadronic and nuclear matter now possible in the Exascale Computing Project Era
  - Maximal exploitation of advanced leadership hardware capabilities allow previously unaffordable calculations

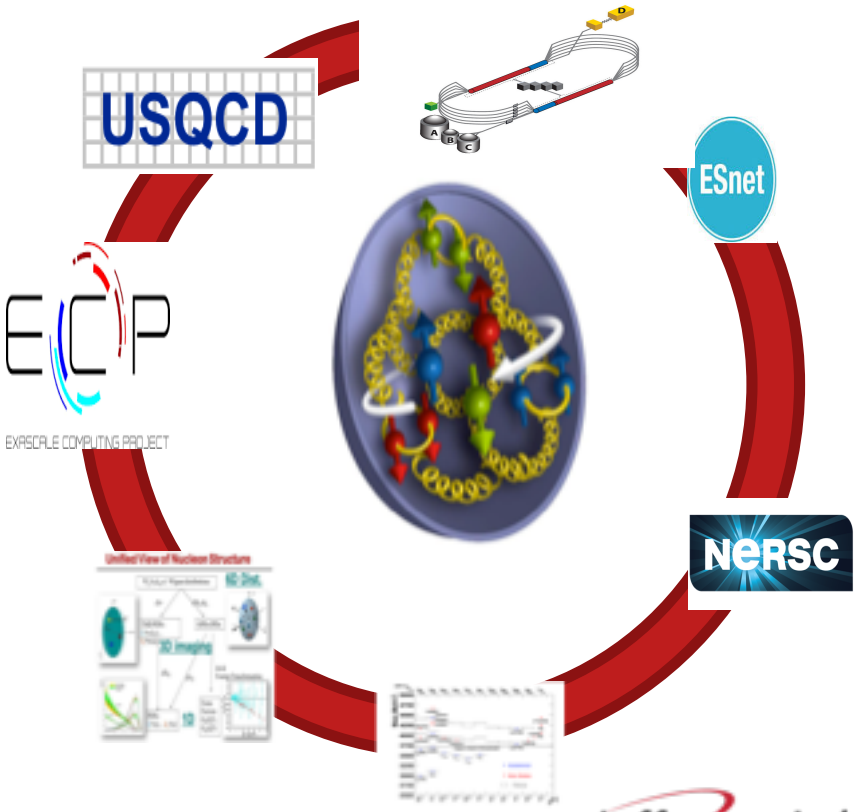
- World Class Program in Theory and Computational Science
  - Partnerships with NVIDIA and Intel
  - New Nuclear Physics Hardware
    - Extended SciPhy XVI cluster Xeon Phi KNL by 12,240 cores
  - SciDAC4
  - Exascale Computing Project
    - Oakridge Leadership Facility: Summit



https://www.olcf.ornl.gov/2018/06/08/summit-by-the-numbers/

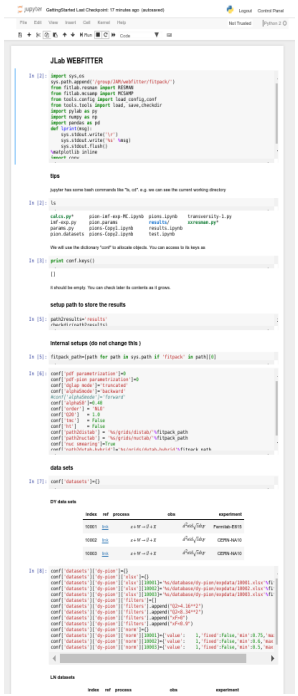
Positioning towards a Strategy

- Building towards software institutes is a long process
- In the meantime, we have jobs to do.
- Foster current activities that build towards strategy
- Experimental Computing:
  - Stream Readout Laboratory
    - Flexibility to reduce temporal and geometric detector constraints
    - Explore analysis implications
  - Foster 'Super Facility' mindset: GlueX production at NERSC
  - Data repositories and attribution mechanisms for data and software—needs attention
- 3D Tomography Parameter Extraction
  - Using 1D fits as exemplar
  - Improving physics simulations
  - Incubate cross-disciplinary science
    - Institute' approach to develop math, data science and visualization

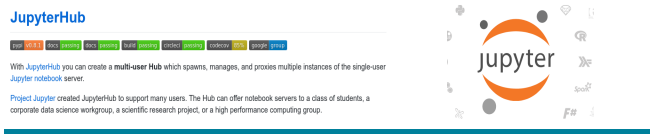


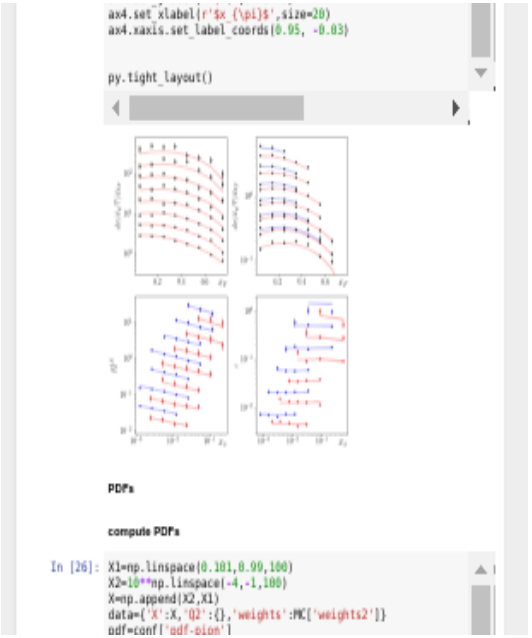
Parton Distribution Function ‘Self-Serve’

JLab webfitter



JupyterHub





User friendly setups

Easy data visualization tools

- Enable experimentalists to fit 1D Nuclear Structure
- Using theory to motivate experimental parameters
- Introduce Jupyter Notebooks for analysis
- Broad collaboration—Jefferson Lab Physics, Theory and IT + Users

Conclusions

- Significant investments in onsite computing for the 12 GeV program
- Also making investments to facilitate distributed computing

- Developing a Scientific Computing Strategy that integrates advances in computing and computation and relies on explicit collaboration
  - Gentle positioning has begun with 12 GeV Science.

- Questions?

Super Facility – Courtesy of NERSC

https://people.eecs.berkeley.edu/~yelick/talks/data/Superfacility-TechX17.pdf