Fermilab

Managed by Fermi Research Alliance, LLC for the U.S. Department of Energy Office of Science

SCD for the Users

Gabriel N. Perdue

49th Annual Fermilab Users Meeting

16 June 2016

SCD for the Users!

- The FNAL Computing Sector is built from two pillars - Core Computing and Scientific Computing.
  - SCD is the *Scientific Computing Division*
- What is scientific computing?
  - Software is “eating the world”*, and physics is not exempt. Computing is a central enabler and accelerant for modern science.
  - **Scientific computing is the suite of technologies that we use to do science, and it encompasses data acquisition, data handling and storage, simulation, analysis frameworks, distributed computation mechanisms, visualization tools, and more.**
- *Who* is SCD?
  - SCD is roughly 140** scientists, engineers, and other technical staff that work together to deliver these solutions at scale.

 *M. Andreessen; **Not a limit imposed by Twitter

Service catalog snapshot

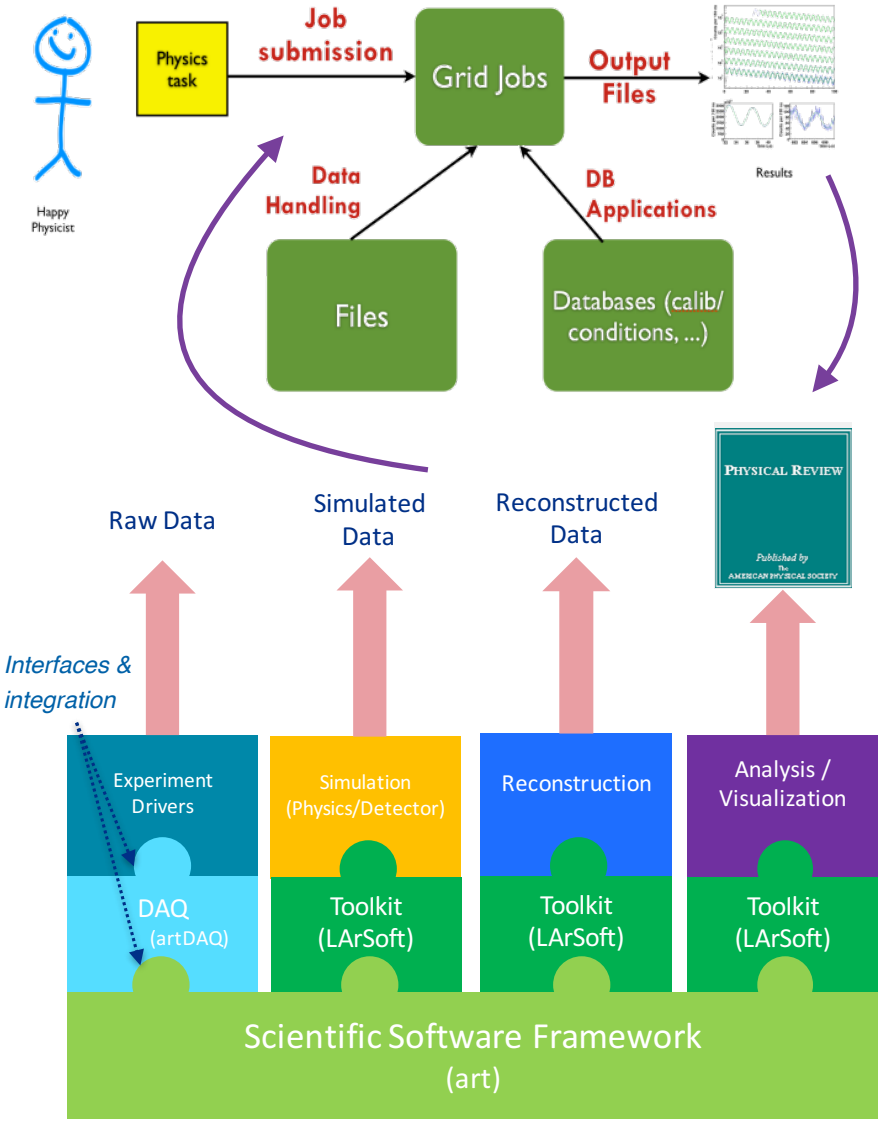
DAQ and Engineering	artdaq, DAQ and Engineering Consulting, Electronic Module Support
Distributed Computing	Batch Job Management (jobsub), Batch Job Management (jobsub) Enhanced, Community On-Boarding to Use Distributed Computing, Distributed Resource Accounting (Gratia), User Jobs Monitoring (fifemon)
High Performance Computing	USQCD Facility Application Support, USQCD Facility File-System Support, USQCD Facility Parallel and Tightly Coupled Batch Computing, USQCD Facility User Accounts, Wilson Facility Application Support, Wilson Facility File-System Support, Wilson Facility Parallel and Tightly Coupled Batch Computing, Wilson Facility User Accounts
High Throughput Computing	Batch Job Operations, Batch Job Operations - CMS, Data and Application Caching Operations, Distributed Resource Accounting Operations, Distributed Resource Provisioning Operations, Distributed Resource Provisioning Operations - CMS, Distributed Resource Provisioning Operations - CMS (enahnced), FermiCloud Operations, Scientific Authentication and Authorization Operations, Scientific Authentication and Authorization Operations - CMS
Physics and Detector Simulation	geant4, Genie, Pythia, Synergia
Scientific Collaboration Tools	DES Members, DES Publication Board, DES Speakers Bureau, Electronic Logbook, Projects, Shift Scheduler, Speakers Bureau, Git, CVS, Redmine, SVN, Electronic Logbook during Data Taking
Scientific Computing Systems	Continuous Integration Infrastructure, CVMFS Repository, Interactive Server Facility, Linux System Monitoring, Linux System Monitoring Enhanced, Scientific Server Management, Managed Scientific Workstation, Scientific Facilities Service, Scientific Server Management CMS Tier 1, Scientific Server Management Enhanced, Scientific Linux Distribution, Scientific Linux Engineering, Experiment Online System Management, Control Room System Management, System Administration and Engineering Consulting, Limited SLF Workstation and Scientific Test stand
Scientific Data Management	FTS (File Transfer Service), FTS (File Transfer Service) Enhanced, IFDHC (Intensity Frontier Data Handling Client), IFDHC (Intensity Frontier Data Handling Client) Enhanced, SAM (Sequential Access via Metadata) Enhanced, SAM4Users, UConDB, SAM (Sequential Access via Metadata)
Scientific Data Storage and Access	Active Archive Facility, dCache Disk Cache Storage, Enstore Tape Storage
Scientific Database Applications	Conditions Database, DES Alarms Viewer, DES Constants Database, DES Exposures Viewer, DES Telemetry Viewer, Hardware Database, IFBeam Conditions Database (IFBeamDB), IFBeam Conditions Database (IFBeamDB) (Enhanced), Query Engine
Scientific Production Processing	Calibration Processing, Keep Up Processing, Monte Carlo Generation, Monte Carlo Simulation, POMS
Scientific Software Infrastructure	art development and support, cetbuildtools/mrb, larsoft development and support, scisoft, ups development and support

How to make sense of all of it? Ask your liaison!

Every experiment has a representative to SCD to foster communication.

What do we do for scientists?

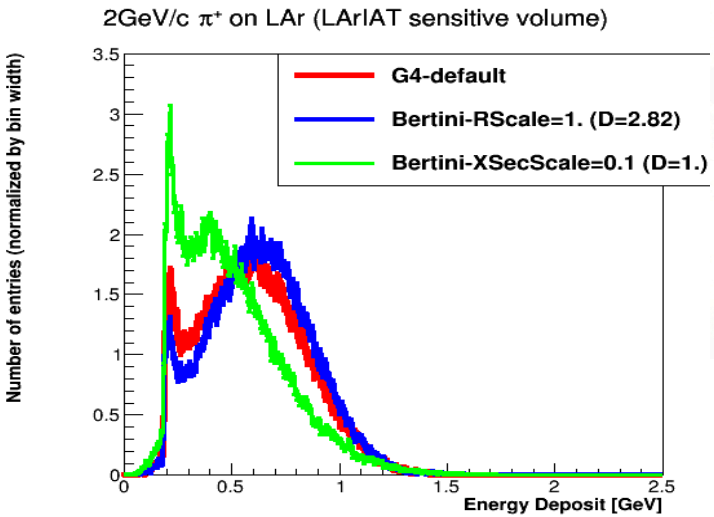
- We are involved in almost every stage of an experiment...
  - Designing detectors and beams
  - Simulation
  - Taking data
  - Data handling and storage
  - Analysis frameworks
  - Batch production
  - Reconstruction
  - Visualization tools
- In many of these categories we develop custom solutions in-house that leverage Fermilab infrastructure. In others, we partner with international collaborations.
- Let's look at a few in "detail"…

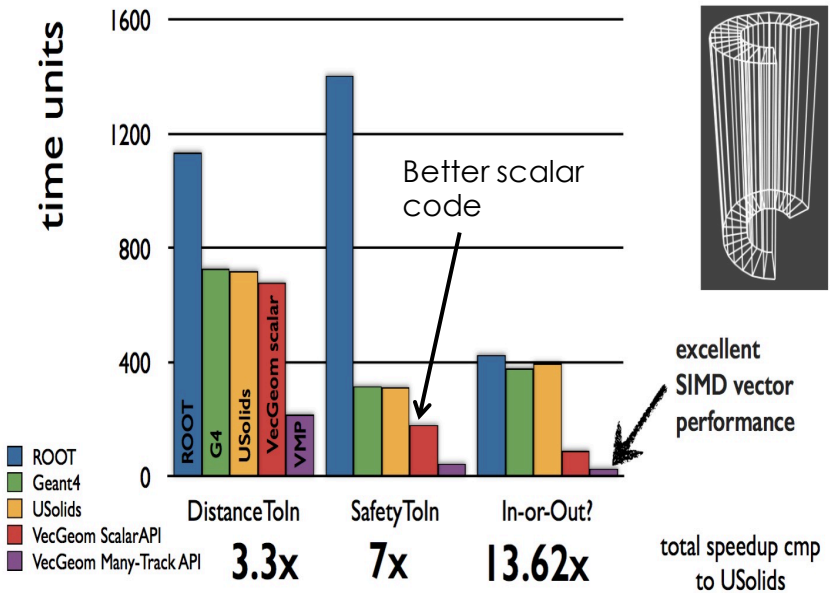


Planning your experiment and understanding your detector

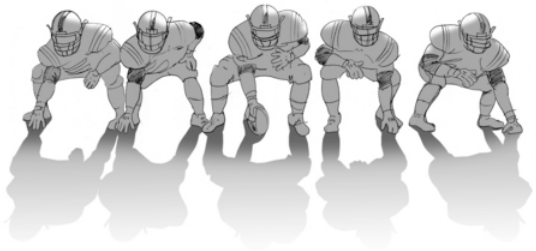
- Developing a new "varied model parameters" interface with a user-friendly API to study systematics due to the simulation model.
  - Beginning with the Bertini model.
  - API supported with robust art-based example applications, supporting documentation, model parameter documentation, presentation and analysis tools.

- Geant *Vectorized*
- Early success with Geometry (VecGeom)
  - SIMD parallelism brings some performance advantage to Geant4 also.
- Develop code to use abstract backends.
- Project on track to achieve 2-5x performance improvement.



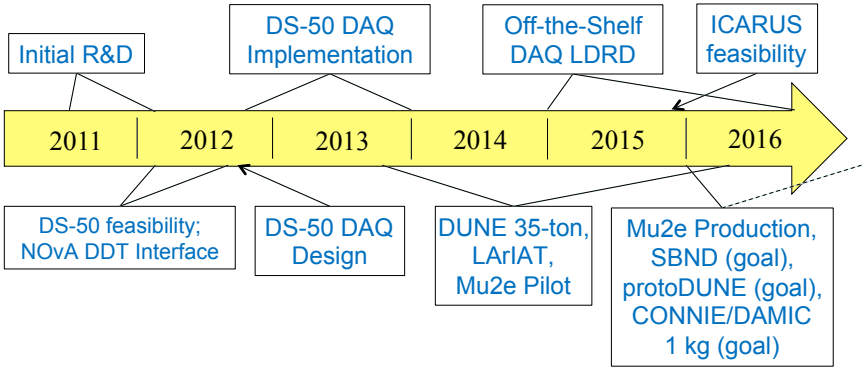


Frameworks - the unsung heroes



- art: an integrated framework for HEP experiments.
  - Started in 2010 as a fork of the CMS framework.
  - Rapid development in 2015-16: Improved build system, validation for configuration (FHICL), I/O enhancements, ROOT6 and Geant4.10 support, and much more.
    - https://cdcvs.fnal.gov/redmine/projects/art/roadmap
- LArSoft: common software for liquid Argon TPCs.
  - Also started in 2010; SCD joined in 2013.
  - Extensive progress in 2015-16: new reconstruction, more modularized services, improved memory footprint, weekly releases
    - https://cdcvs.fnal.gov/redmine/projects/larsoft/wiki/LArSoft_release_list
      - *Upcoming workshops (tomorrow)!:* art.fnal.gov/register & larsoft.org/register
- Heavy involvement in cmssw
  - Implemented much of the capability to make cmssw the first fully multi-threaded HEP software framework.

Data taking frameworks too!



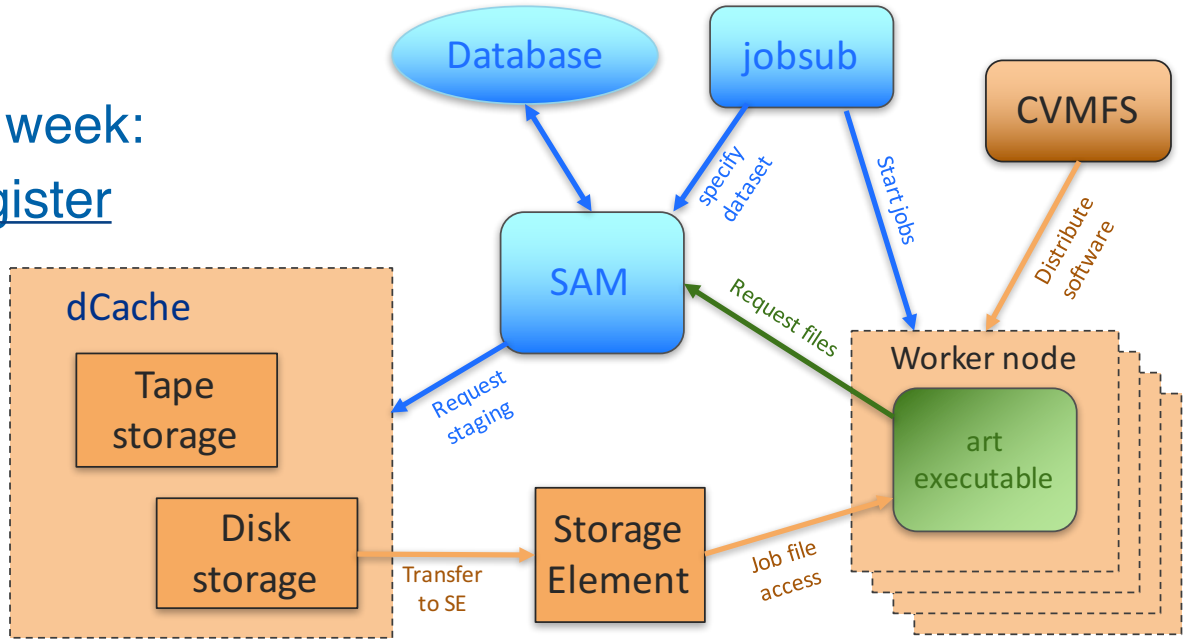
- What is artDAQ?
  - http://inspirehep.net/record/1229212
  - Core, composable DAQ software integrated with the art framework.
    - Same environment for on and offline algorithms
  - LArIAT turned on this past year with artDAQ - first artDAQ LAr experiment.
  - DAQ technology for ProtoDUNE
  - Off The Shelf DAQ (otsdaq)
    - IoT boards, JavaScript control GUI, host and embedded APIs for socket-based communication with artdaq, reusable firmware components

Batch processing

- JobSub
- Monitors, alarms, automated submission
- Software distribution (CVMFS)
- Authentication
- ECL, beam info databases

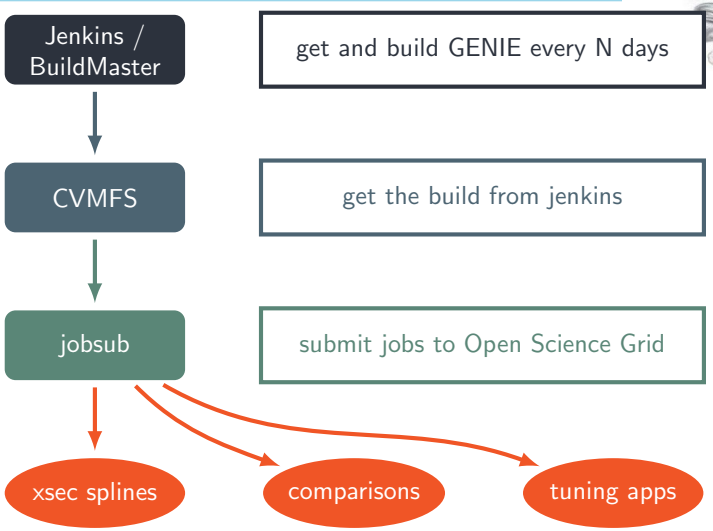
- Data handling & distribution
  - Sequential Access via Metadata (SAM)
  - dCache/Enstore
  - Fermi File Transfer Service (F-FTS)
  - IF Data Handling Client (IFDH)

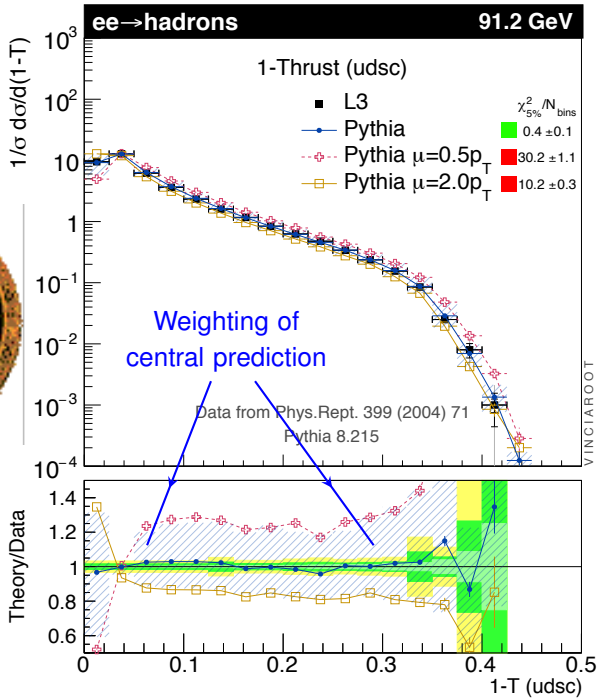
And another workshop next week: fife.fnal.gov/register



Understanding physics models with event generators

- GENIE is the most widely adopted neutrino MC event generator.
  - We help lab based experiments to meet needs for core event generation, e.g.
    - pushed special releases with new multi-nucleon interaction channels for NOvA,
    - pushed ROOT6 support for LArSoft team with single day turnaround time,
    - Simulation for Neutrinos working group.
  - Developing and implementing new physics models and automated validation and tuning.
- Pythia is an event generator for particle collisions (used by GENIE, astro, colliders). Under constant development:
  - Python API and interface to Madgraph5_aMC@NLO.
  - New framework for automated patron-shower uncertainty bands.
  - Weak merging implemented (W bosons may be produced as part of the hard matrix element or in the parton shower, and these possibilities must be consistently merged).
  - Improved interface to VINCIA and DIRE (external parton shower libraries).





Use other SCD tools heavily for GENIE automated validation: Build Service, CVMFS, jobsub, etc.

Facilities: the backbone of it all

- GPGrid: 21,108 cores
- CMS: 22,856 cores
- 28,240 cores (HPC / LQCD)
- 2,008 cores (DZero legacy)

- General: 292.6 YHDV
- CMS: 571.9 YHDV
- Legacy: 399 YHDV
- Capacity: 12,901 YHDV

- General purpose ~ 113 YHDV
- CMS ~ 292.6 YHDV
- Legacy CDF ~ 20 YHDV
- User disk ~ 93.1 YHDV
- HPC (Lustre) ~ 13.3 YHDV

- YHDV = Years of HD Video

What's next? How does SCD fit in the Strategic Plan?

- Many important themes for how users experience SCD:
  - Move to the cloud
  - Ubiquitous and uniform experiences on any device
  - Effective security that safeguards the data we manage
  - Stronger partnerships with industry and the Office of Science (with an emphasis on ASCR (Advanced Scientific Computing Research))
  - New onsite data center
  - Elastic resource delivery
- Much more in the plan:
  - cd-docdb.fnal.gov, doc 5733

Conclusions

- SCD is a core enabler for science at the lab.
  - Leveraging core tools and expertise across many experiments to get the largest possible impact from limited resources.
- We are doing exciting things right now (far too much to fit in this presentation - apologies to all the groups that weren't highlighted or discussed here!).
  - SCD is involved in every stage of an experiment from the moment data is created to the publication of results and beyond.
- Compelling vision in place for the future:
  - clear path to exascale computing,
  - tools to help users manage the complexity of a world with a mix of cloud-based, elastic resources as well as on-site computing facilities,
  - modern security and usage models.

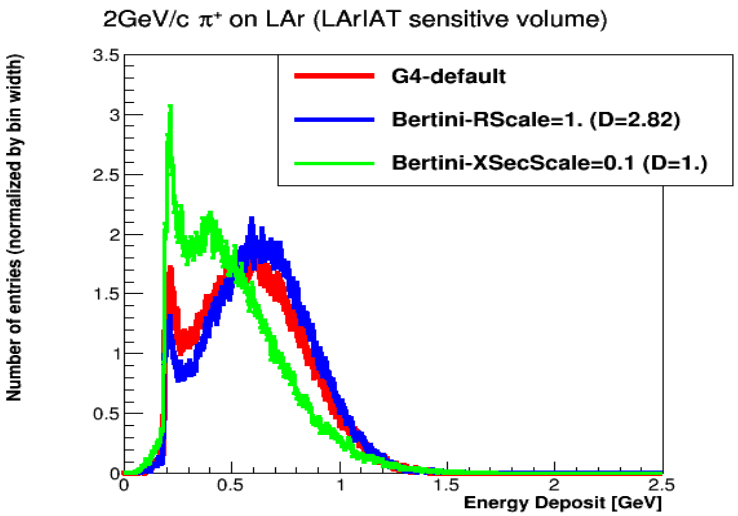
Thanks to K. Biery, S. Fuess, O. Gutsche, B. Holzman, M. Kirby, A. Lyon, P. Spentzouris, M. Votava for slides, figures, and comments…
Thanks for listening!

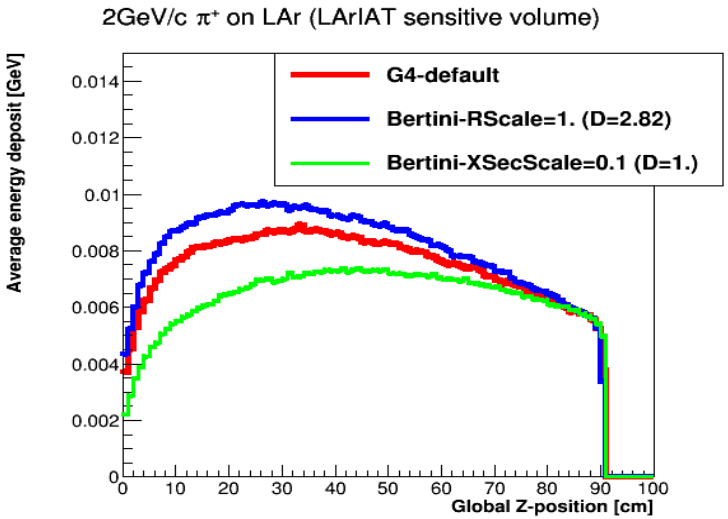
What is SCD?

- At Fermilab we organize scientific computing into four quadrants:
  - Systems for Scientific Applications
    - physics simulation programs, analysis frameworks, data acquisition
  - Scientific Programs
    - collaborators on your experiment
  - Scientific Computing Services
    - all things grid, data handling
  - Scientific Computing Facilities
    - the hardware to make it all work - disk, compute nodes, etc.
- We're involved in every step from getting data out of the detector all the way through analysis and publication.
  - And even long term data archival storage.

Physics and Detector Simulations - GeantIV

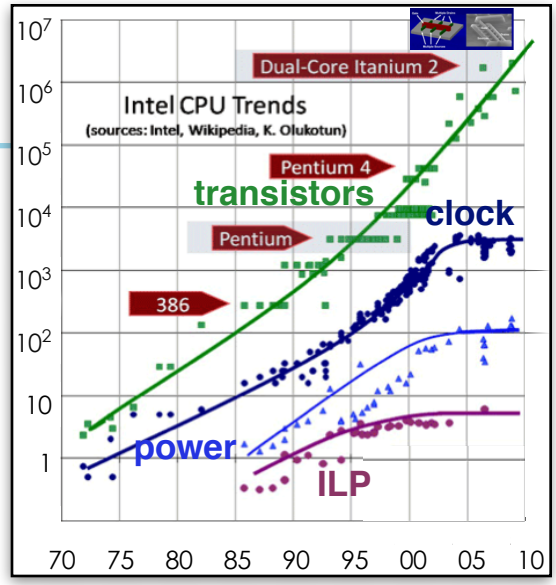
- Developing a new database application (DOSSIER) to make G4 validation easier and more transparent.
  - Web interface and API for users.
  - Supports other MCs (e.g. GENIE).
- Developing a new "varied model parameters" interface to make it possible for users to modify Geant4 model parameters with user-friendly API to study systematics due to the simulation model.
  - Beginning with Bertini model for low energy hadronic physics.
  - API supported with robust art-based example applications, supporting documentation, model parameter documentation, presentation and analysis tools.
  - Shown here: radius scale and internal cross section scale modifications (large!) on LArIAT geometry.

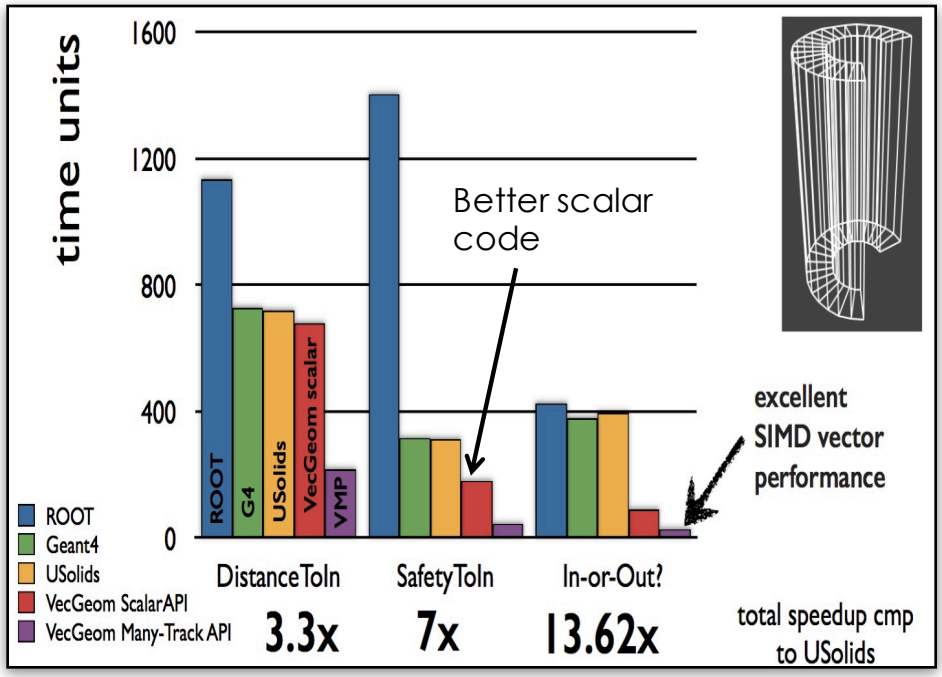




Physics and Detector Simulations - GeantV

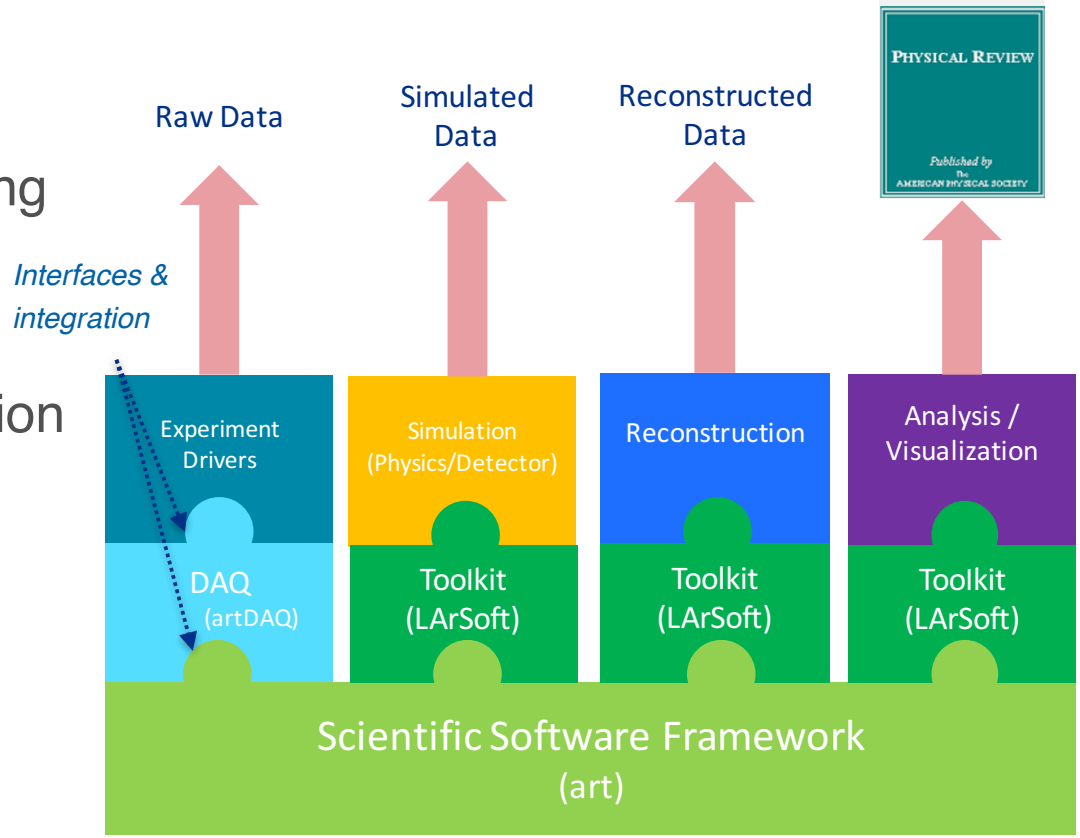
- Geant *Vectorized*
- Project goal - develop a simulation 2-5 times faster than Geant4, with a good understanding of the limitations for a x10 speed increase.
- Early success with Geometry (VecGeom)
  - SIMD parallelism brings some performance advantage to Geant4 also.
- Develop code to use abstract backends for different physical vectorized hardware implementations.
- Project on track to achieve 2-5x performance improvement.





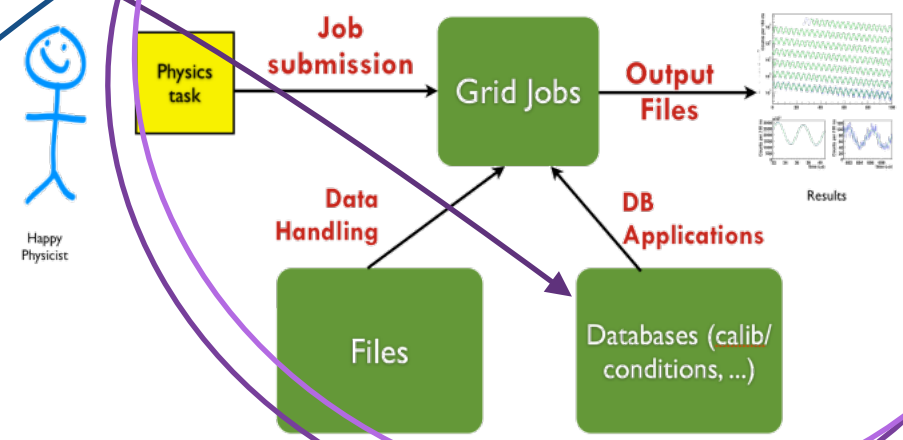
Scientific Applications

- Software Infrastructure:
  - Frameworks
  - Reconstruction
  - Tools and Advanced Computing
  - Build and Release
- Simulation
  - Physics and Detector Simulation
  - Accelerator Simulation
- Real-Time Systems
  - DAQ Controls and Detectors
  - Real-time Software Infrastructure
  - Detector Electronics
  - Physics Research Equipment



Scientific Computing Services

- Scientific Data Processing
  - Scientific Data Management
  - Workflow Management and Provisioning
  - Scientific Database Applications
- Scientific Distributed Computing Solutions
  - Scientific Computing Information Security
  - User Support for Distributed Computing
  - Experimental Astrophysics
  - Workflow Operations for Distributed Computing

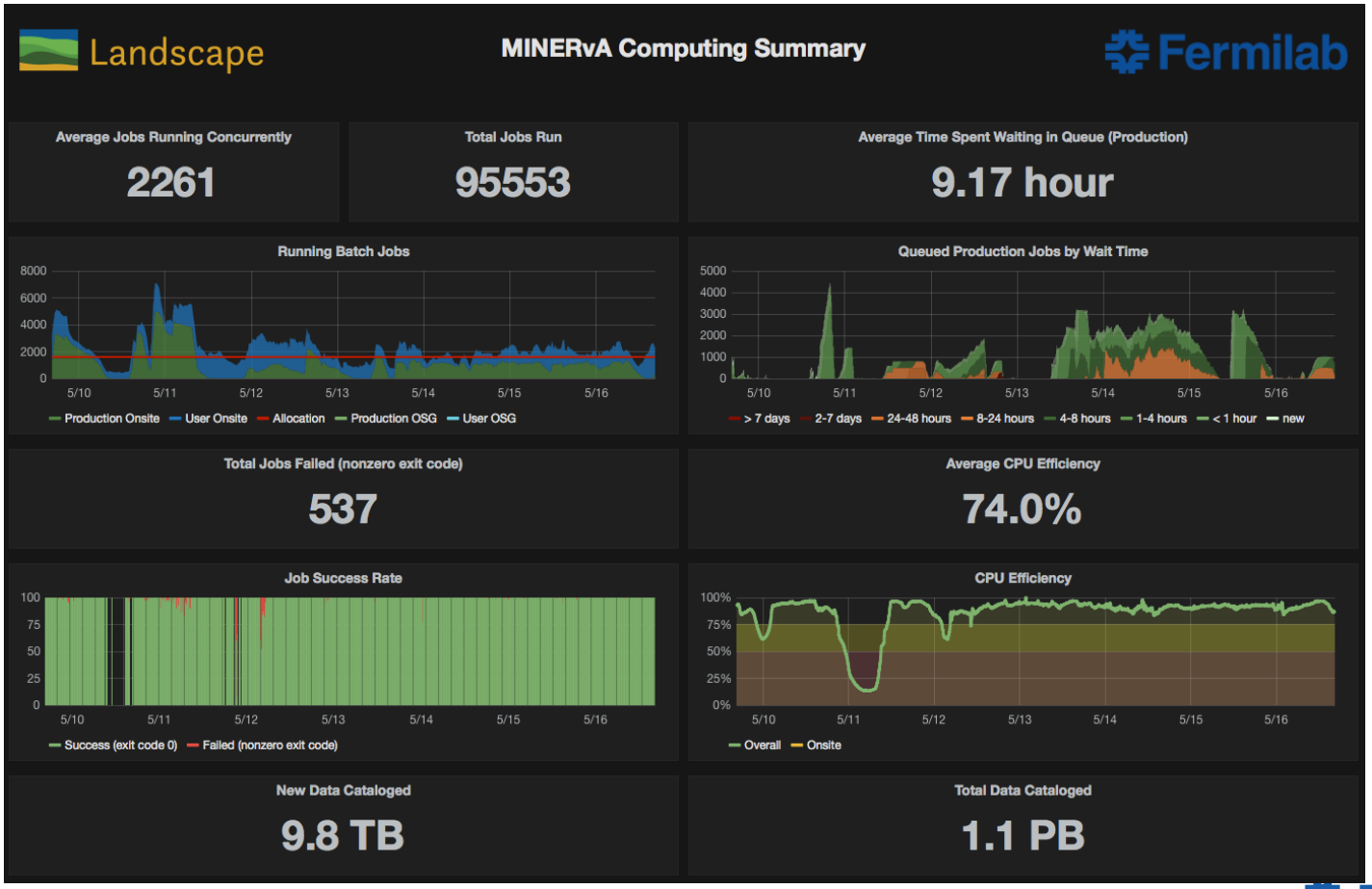


FABRIC FOR FRONTIER EXPERIMENTS

Fermilab

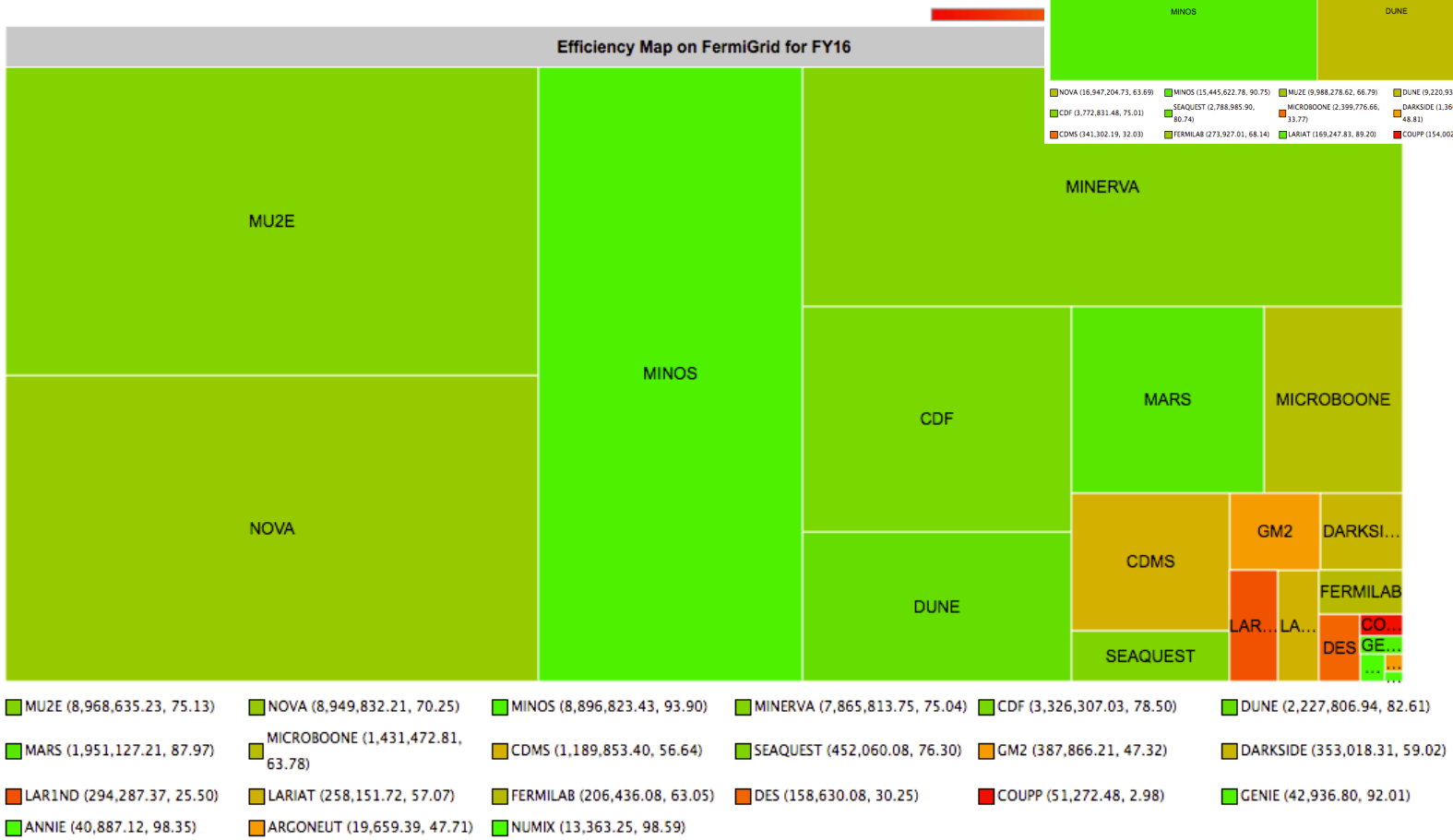
Computing Services - Monitoring

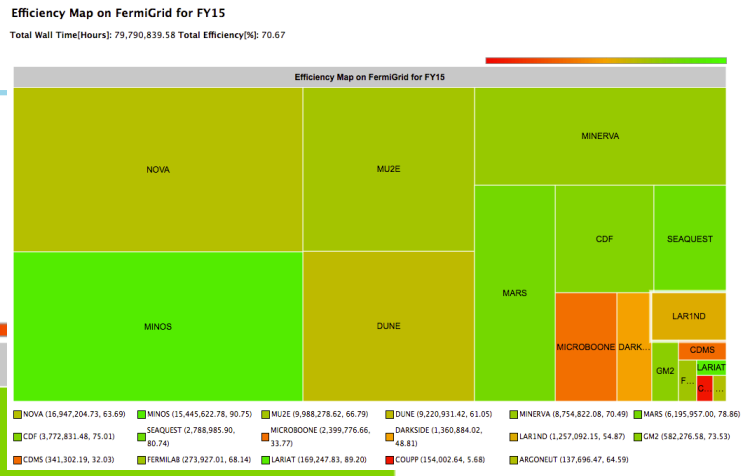
- Lots of progress in many areas, but visually arresting example: Monitoring!
  - Powerful new capabilities: https://fifemon.fnal.gov/monitor/



Computing Services - Monitoring

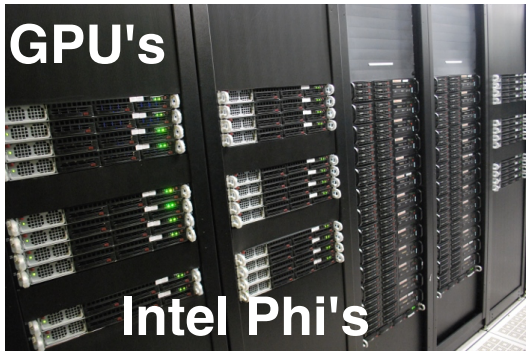
FY16: 7% integrated CPU efficiency gained

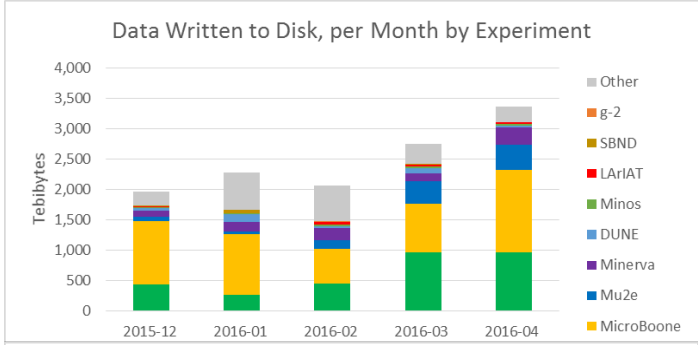


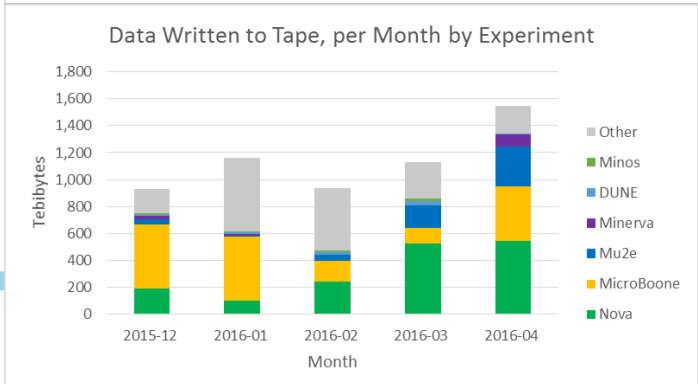


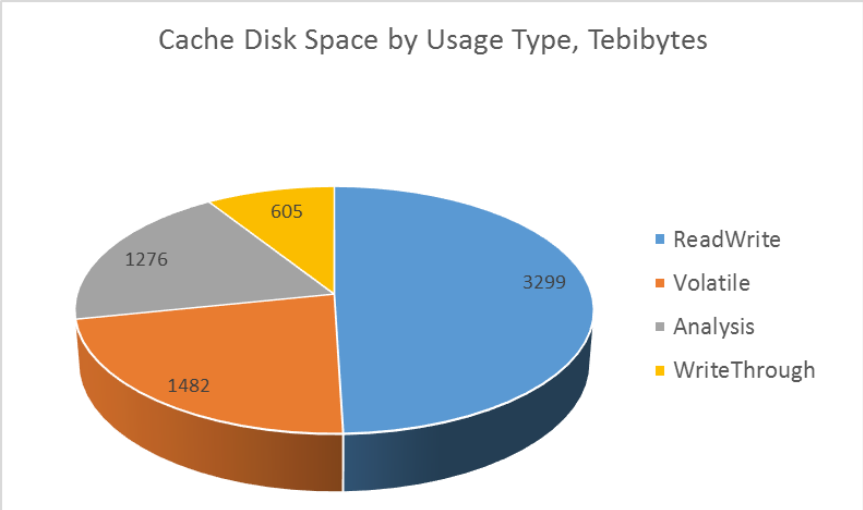
Scientific Computing Facilities

- High Performance Parallel Computing Facilities
- Data Movement and Storage
  - Data Movement Development
  - Storage Services Administration
- Experiment Computing Facilities
  - Scientific Linux and Architecture Management
  - Scientific Server Infrastructure
  - Grid and Cloud Operations







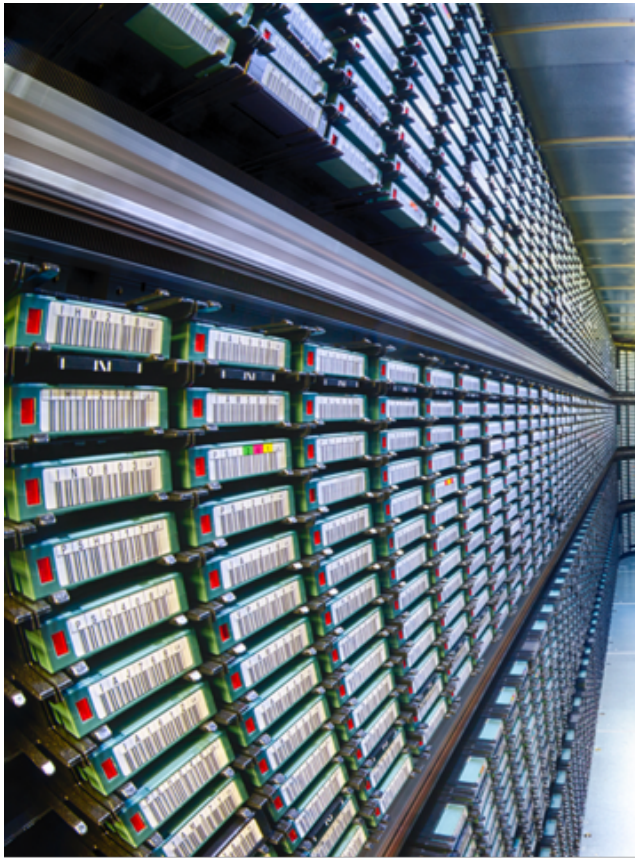


Facilties - Compute

- "Worker node" core counts:
  - General purpose (GPGrid):
    - 16,608 cores (GP Grid)
- Also:
  - 17,872 cores (CMS Tier-1)
  - 4,984 cores (CMS LPC)
  - 28,240 cores (HPC / LQCD)
  - 2,008 cores (DZero legacy)
- Finally, currently re-allocating 4,500 decommissioned cores from LQCD to GP Grid

Slide from S. Fuess

Facilities - Storage: Tape



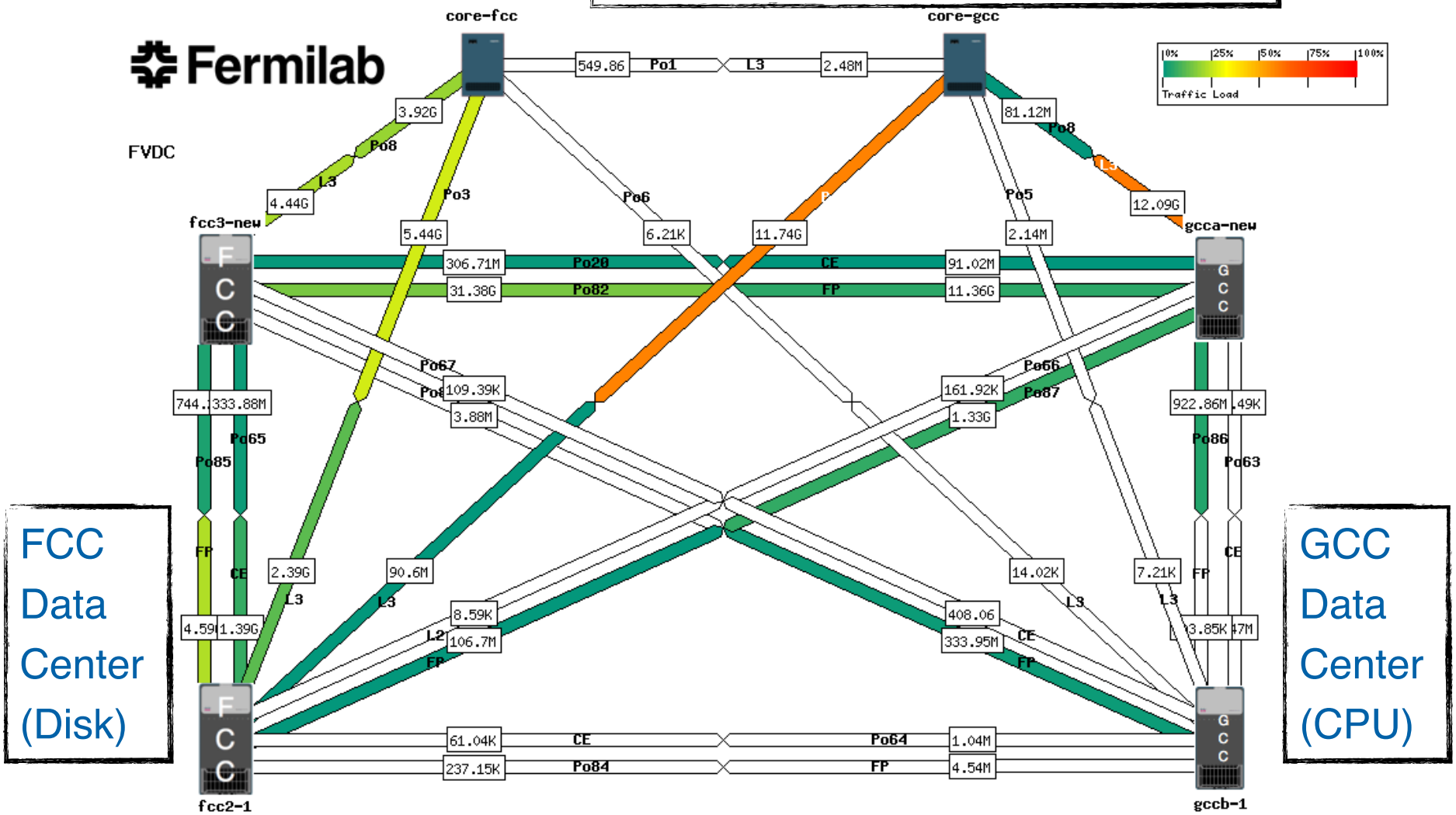
- Tape Capacity:
  - 7x 10,000 slot libraries
    - With 5.4 TB T10000C ~ 375 PB
    - With 8.5 TB T10000D ~ 595 PB
  - Allocation:
    - General purpose: 4 libraries
    - CMS: 3 libraries
- Current usage
  - General purpose ~ 22 PB
  - CMS ~ 43 PB
  - Legacy CDF, DZero ~ 30 PB
  - (includes migration duplicates)
- Fun fact: 1 PB is 13.3 *years* of HD video.

Facilities - Storage: Disk

- Cache disk (dCache):
  - General purpose ~ 8.5 PB
  - CMS ~ 22 PB
  - Legacy CDF ~ 1.5 PB
- Project/user disk (NAS):
  - General purpose ~ 2 PB
  - CMS (EOS) ~ 4 PB
  - Legacy CDF, DZero ~ 1 PB
- HPC disk (Lustre):
  - LQCD ~ 1 PB

Facilities - Network

Offsite network: 2x 100Gb, 3x 10Gb



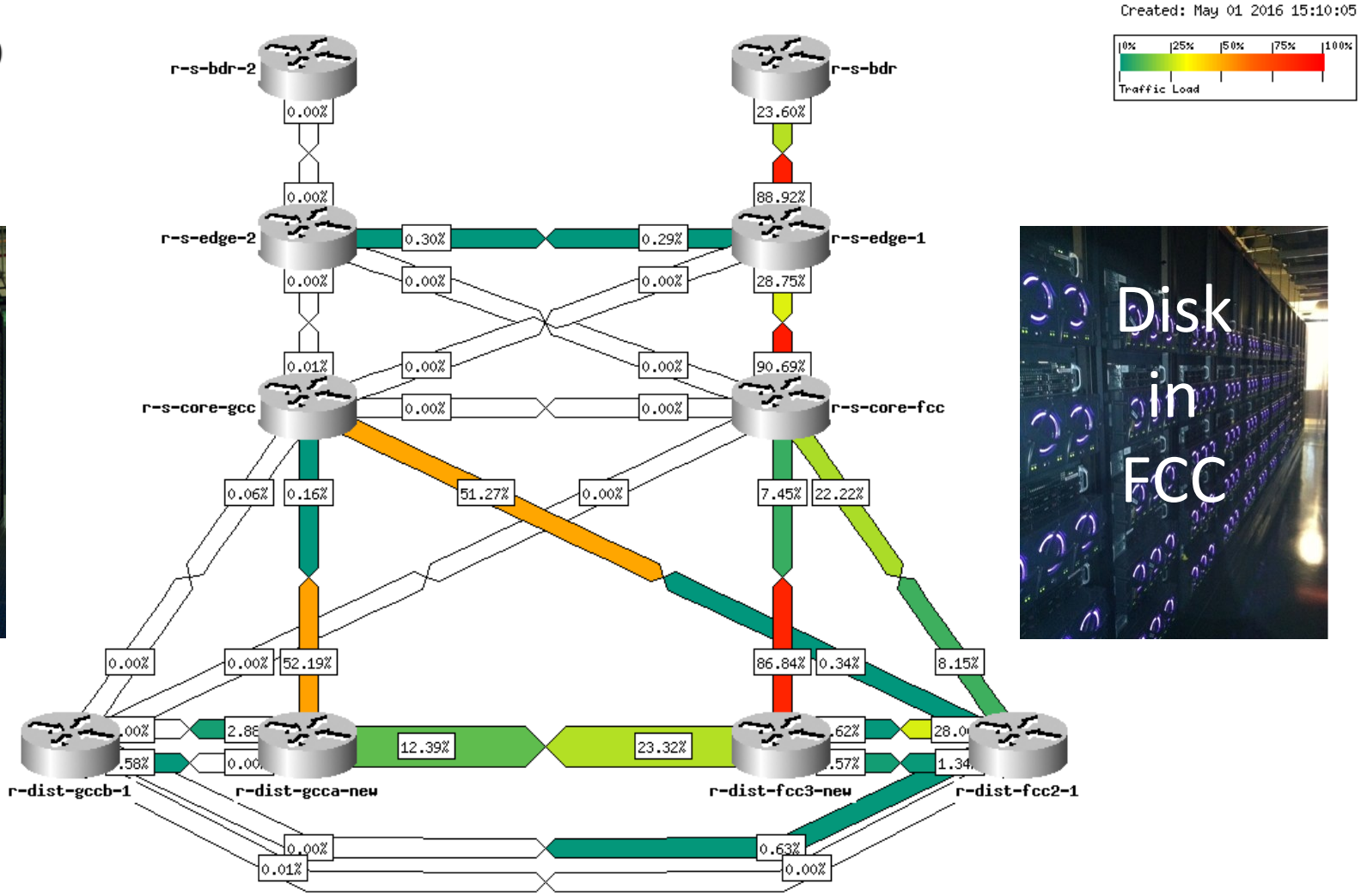
FCC Data Center (Disk)

GCC Data Center (CPU)

- Network support from Core computing Divsion.

Facilities - Network Monitoring

Fermilab
Core Network



Strategic Plan: Active Archival Facility

- Developed as a component of the HEP Cloud Facility to enable external scientific collaborations to utilize large scale storage at Fermilab.
- Requires new development in:
  - monitoring and accounting tools to support many customers,
  - enhanced cataloging and management tools,
  - access strategies for non-HEP partners.
- The AAF will provide a valuable service to scientific efforts that cannot otherwise access a massive storage facility.

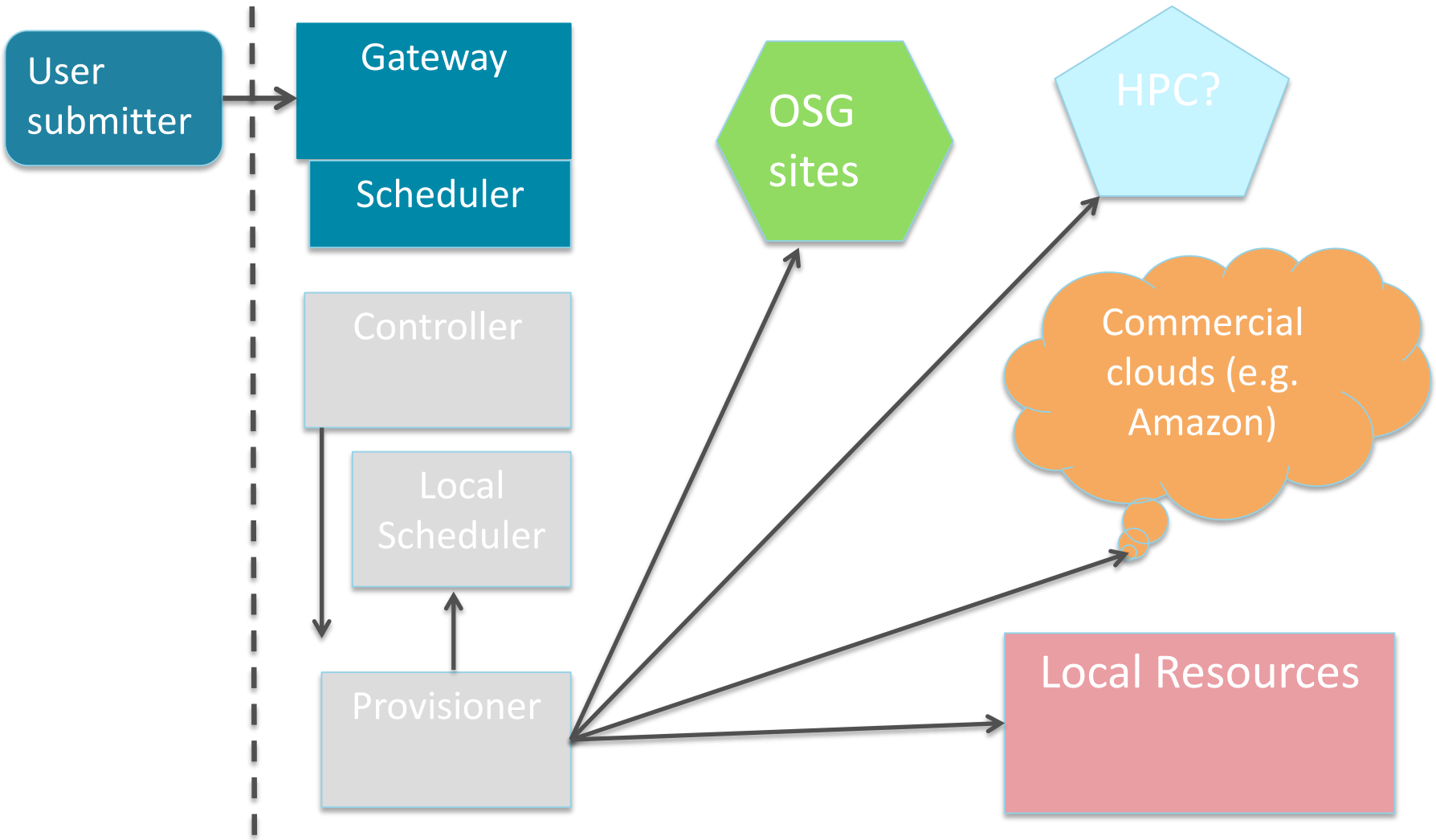
Strategic Plan: computation frontier

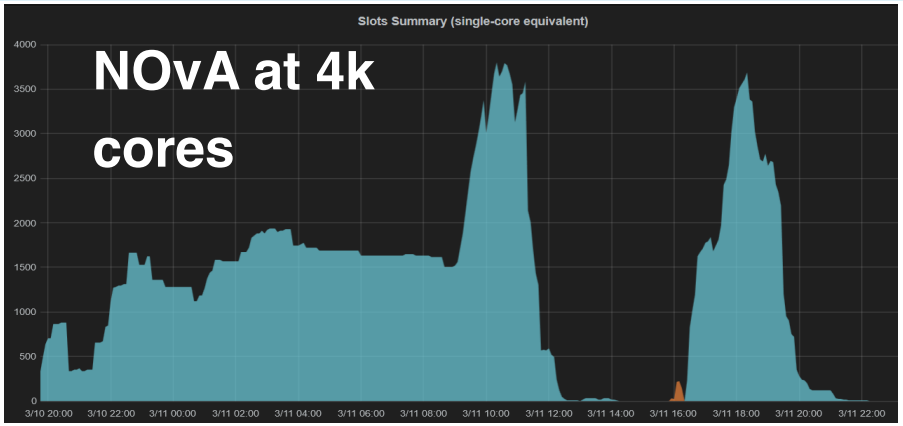
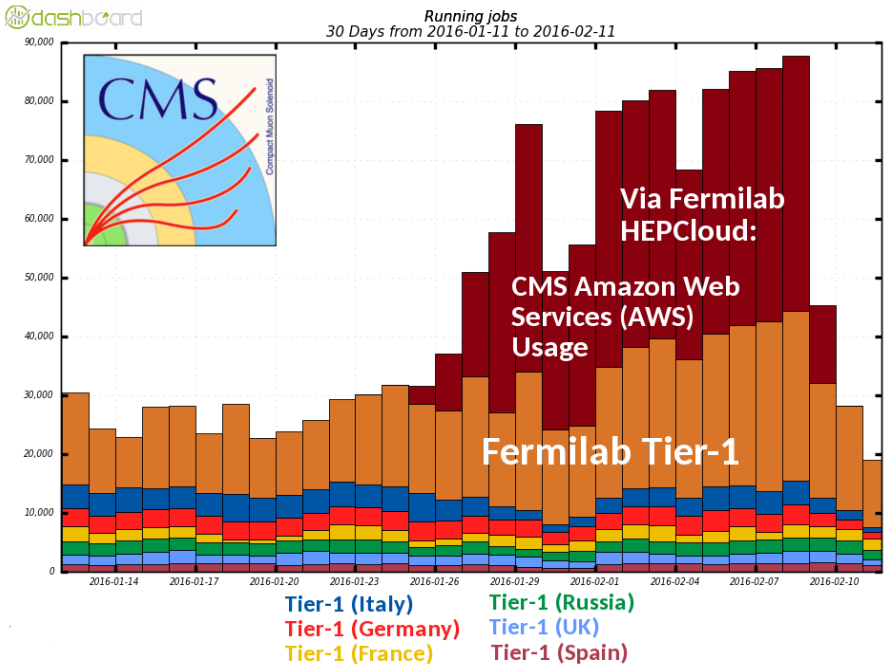
- Evolve software toolkits to utilize multithreading and vectorization.
- Seed and coordinate cross-cutting development efforts that span multiple HEP experimental programs.
- Create a new Center for Numerical Computation in HEP.
  - Establish a terabit network link between FNAL and Argonne to leverage respective expertise,
  - Include applied mathematics and computer science colleagues in lattice QCD and simulation activities,
  - participate as full partners in DOE ASCR high-performance computing roadmaps.

Strategic Plan: HEPCloud Facility

- The “exabyte era” is coming: in the next decade, data volumes from the HL-LHC will be 100 times larger than current HEP capacity.
- The HEP Cloud Facility concept leverages industry trends and requires significant development in:
  - efficient use of commercial cloud resources (Google, Amazon, Microsoft),
  - algorithms to route specific computation tasks to the most cost-effective resources,
  - seamless data handling protocols capable of abstracting away underlying resource particulars,
  - uniform and user-friendly interfaces, accounting, monitoring, and error reporting.

HEPCloud


User submitter
Gateway
Scheduler
Controller
Local Scheduler
Provisioner
OSG sites
HPC?
Commercial clouds (e.g. Amazon)
Local Resources

HEPCloud

- Elastic resource model
- Manage onsite, offsite OSG, and commercial cloud resources under one HTCondor pool - make a bigger cloud!
- Automatic, cost-aware decision engine under development, deployment scheduled for 2018.
- Already testing with limited, specific use cases now:
  - NOvA, CMS, DES
  - Added support for AWS storage in SAM, IFDH, FTS