Fermilab

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

Scientific Computing - Supporting Experiments in their Scientific Endeavors

Oliver Gutsche

47th Annual Fermilab Users Meeting

12. June 2014

Fermilab's Computing Landscape

- Fermilab is a great place!

- We have an interesting and diverse science program:
  - Many different experiments from collider physics, neutrino physics, astro-particle physics, …
  - Vibrant theory community
  - Advanced accelerator science

- Everyone is using computing!
  - And scientific software, workflow tools, data handling, databases, …
  - We are supporting all experiments
    - With a diverse scientific computing services catalog
    - With direct or indirect support of a variety of software packages

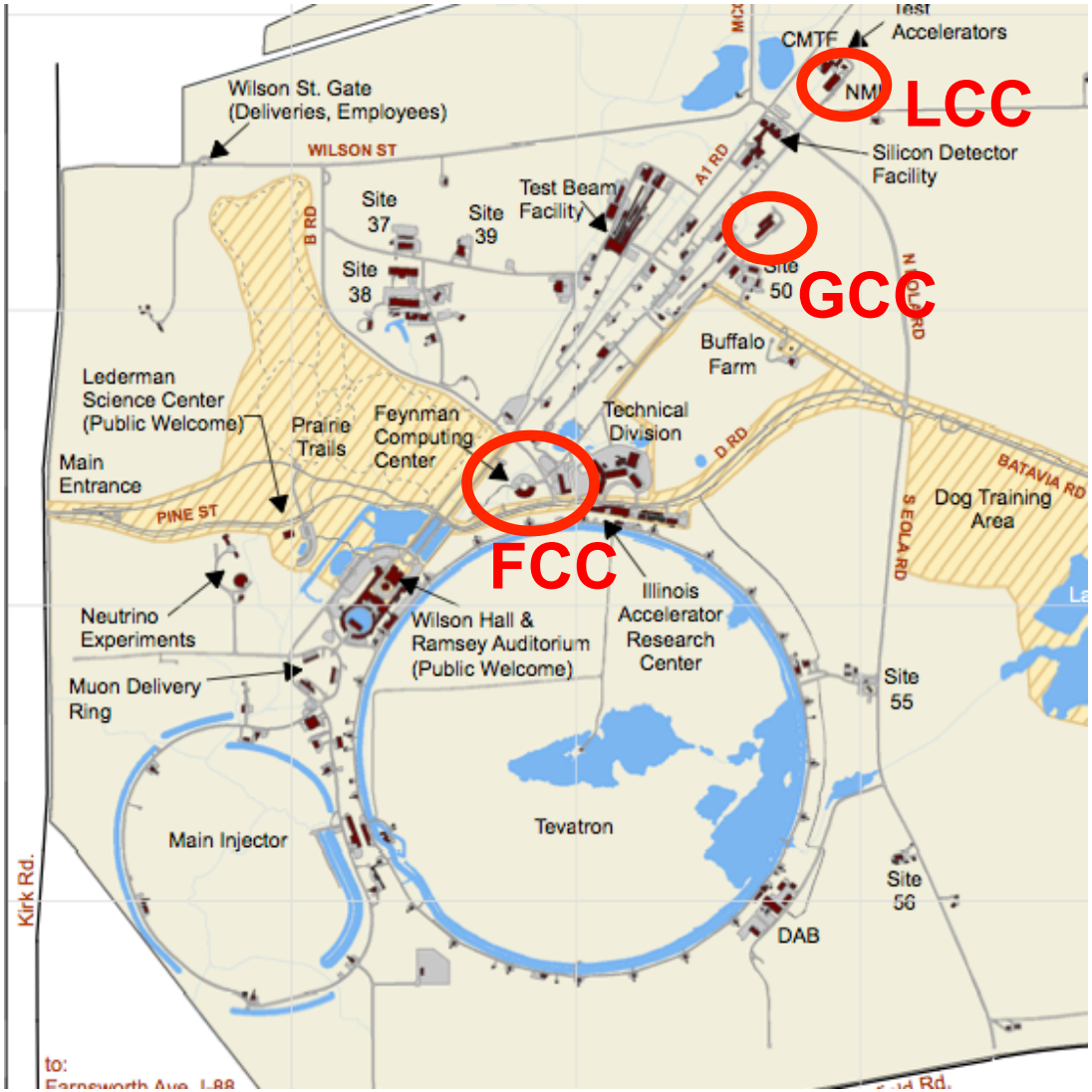
- Fermilab's Scientific Computing is a big part of everybody's daily scientific life!

Supporting Experiments in their Scientific Endeavors

- Facilities
- People
- Services
- Software
- Experiments

- Enabling the Future

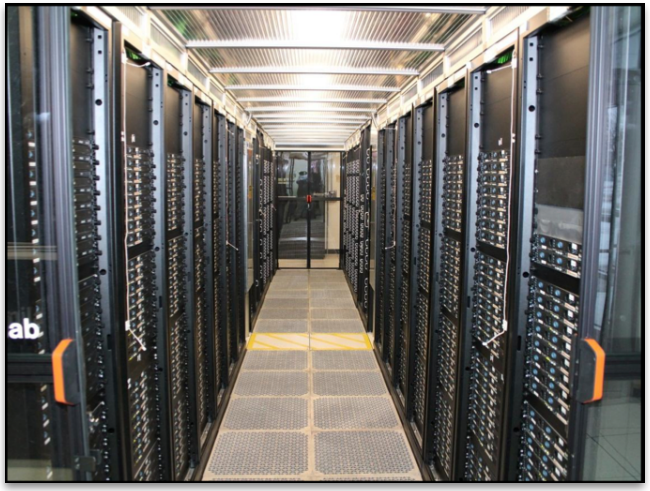
  - Offline Processing Operations Group
  - Data Archival Facility
  - New hardware platforms



Fermilab

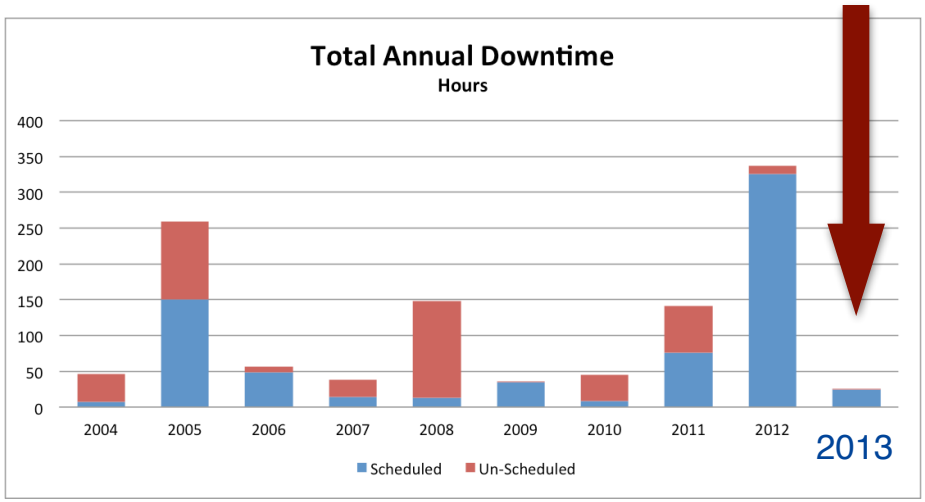
Fermilab's Computing Facilities: Large Numbers

- CPU: 65,000 cores
  - ½ experiments - ½ lattice QCD (numerical theoretical calculations)
- Disk: 25 PetaBytes
  - Technologies: dCache, EOS, BlueArc, Lustre
- Tapes: 375 PetaBytes capacity
  - 70,000 tape slots in robots (not all used)
  - Current usage: 88 PetaBytes
- Network:
  - Internal: ~32,000 network ports
    - connected through network fabric from copper and fiber links
  - External: 130 Gb/s
    - 1 x 100 Gbps and 3x 10 Gbps fiber links



Fermilab's Computing Facilities: Saving money for the Future

- Top facilities ➜ Top science
  - Phenomenal availability in 2013:
    - Less downtime ➜ less hassle for everyone doing science

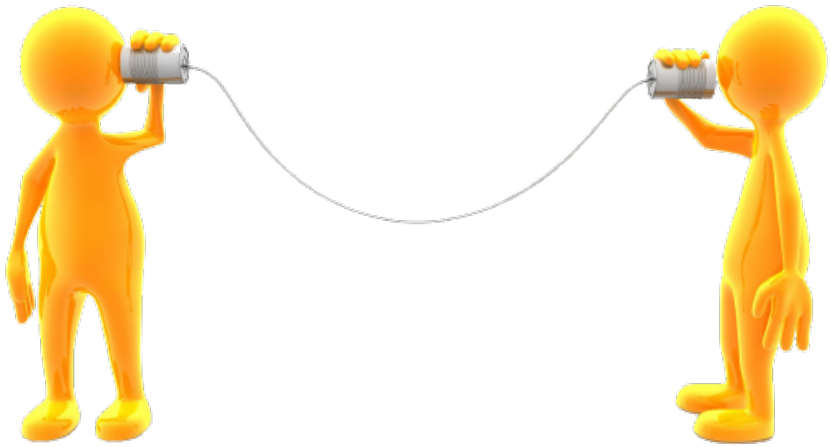


- Continuously improving to save energy
  - Latest: cold aisle containment in computing rooms ➜ reduces cooling costs
  - Cost will be recovered in 21 months
    - After 21 months, we are saving money ➜ can be used for science

Fermilab

Scientific Computing Division (SCD): People & Communication

- SCD is working closely with the scientific community and all other laboratory functions to support the scientific mission



- Good communication is the key to efficient use of facilities and services and to help everyone succeed
  - Approach: Liaisons and Portfolio Management

  - Liaisons: communication between experiments and SCD
    - Every experiment, scientific project and collaboration has SCD liaison
    - Liaisons are embedded in Scientific Computing and vice versa ➜ ensuring smooth operation and communication
    - Bi-weekly liaison meetings ➜ Experiments included in management of scientific services
    - Works well!
      - Important: FIFE workshop 16. - 17. June 2014

  - Portfolio Management
    - Annual review of computing plans, requirements and concerns of experiments
      - Last review was 1. - 2. May 2014
      - Successful model: most problems from 2013 review solved, solutions are available

Scientific Services Catalog available to Experiments

	DAQ	Fermigrid	Jobsub	Fifemon	Redmine	CVS/Subversion/ Git	ECL	ups/upd	CVMFS	Interactive machine in GPCF	continuous integration - middleware	sam/ifdh	file transfer service (FTS)	IFBeam	custom databases
	DAQ and Controls	**Grid and Cloud**			**scientific collaboration tools**				**scientific computing systems**			**scientific data management**		**scientific databases**	
CDF	N/A	plan	plan	plan		yes		yes				SAM			
D0	N/A	test	test	test		yes	yes	yes	test		plan	SAM			
ArgoNeut		yes	yes	yes	yes	yes		yes	no	yes		no			
Lariat	artdaq?	yes	yes	yes	yes	yes	yes	yes	In Design	yes	yes	SAMweb	plan	yes	plan
LBNF	artdaq	yes	g-wms	yes	yes	yes		yes	progress	yes	yes	SAMweb	plan		progress
LBNF 35T	artdaq	yes	yes	yes	yes	yes	yes	yes	yes		yes	plan	plan		plan
uBoone	custom	yes	yes	yes	yes	yes	yes	yes	progress	yes	yes	SAM	yes	yes	
Minerva	Custom	yes	yes	yes	yes	yes	yes	yes		yes	buildbot	SAMweb	yes	yes	yes
MiniBoone	custom	yes	?	yes	yes	yes	yes	yes	no	yes	no	no			
Minos+	Custom/artdaq?	yes	yes	yes	yes	yes	yes	yes	yes	yes		SAMweb		yes	
Mu2e	artdaq	yes	yes	yes	yes	yes	plan	yes	plan	yes	yes				
Muon g-2	midas	yes	yes	yes	yes	yes		yes	yes	yes	yes	SAMweb	yes		progress
NoVA	novadaq	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	SAMweb	yes	yes	yes
Darkside	artdaq	yes	yes	yes	yes	yes		yes	plan`	yes		SAMweb	yes		
seaquest	PREP HW	yes	yes	plan	yes	yes	shift cal	plan	plan	yes	no	no	no		
coupp	custom	yes	yes	interested	no	interested	yes	yes	no	yes	no	yes	yes	no	

- In the past, fewer experiments were competing for resources ➜ SCD was able to work with individual experiments and help them succeed
- Now, even more experiments are competing for resources ➜ Necessary to look for communality and leverage sharing wherever possible
- We are now providing a Catalog of Scientific Services that experiments can choose from
  - Same services can be used by different experiments

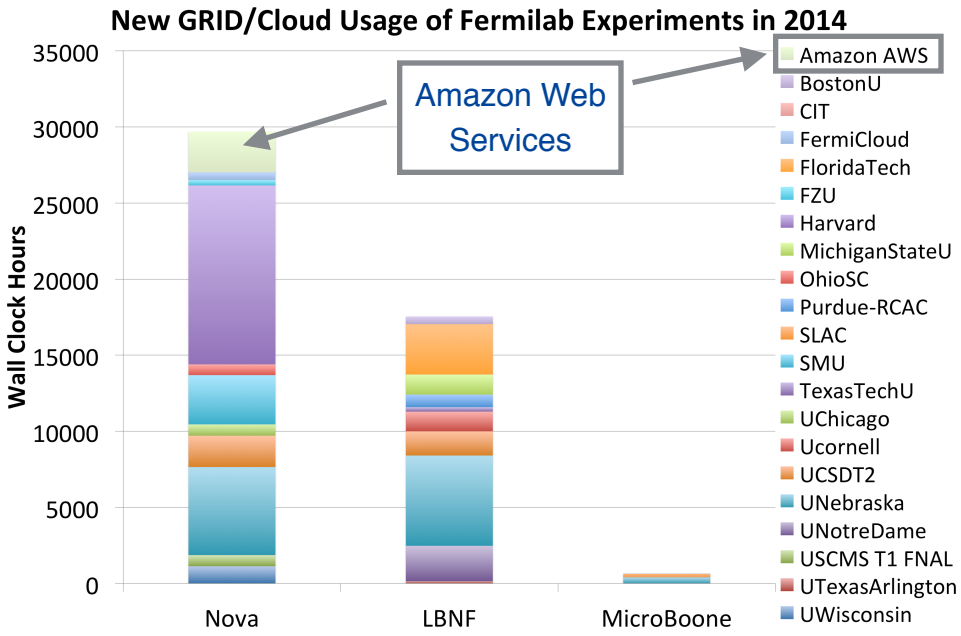
- Big success!
  - We are providing solutions that experiments find useful!

- New: Effort to be more transparent to Experiments' needs
  - Running many services for many experiments needs higher level of organization that ensures success ➜ working on service level agreements

Fermilab

Fermilab and the OSG: Enabling GRID/Cloud usage

- Fermilab's facilities well used, utilization in the last year was around 80%
  - But peak demand is > 100% (everybody wants to produce results for the same conferences)



- SCD is enabling experiments and users to use the OSG for peak demands
  - Easier for the experiments if they are already setup to run on FermiGrid

- CMS, OSG and FNAL experiments are working together more closely
  - NOvA is now first class citizen on the GRID, MicroBoone is on its way, Muon g-2 is coming soon too

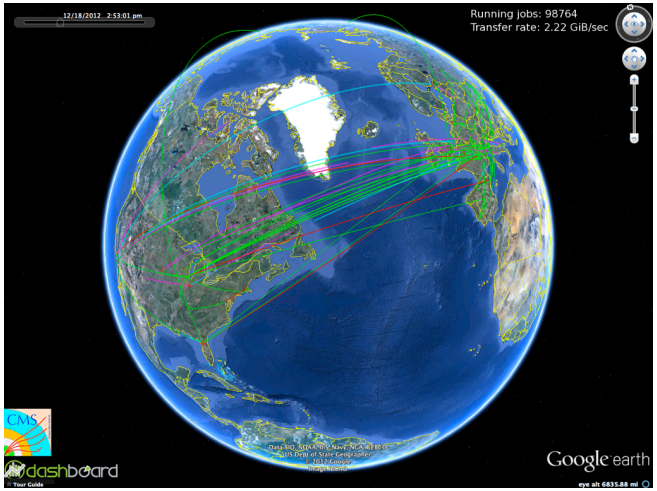
Fermilab's biggest experiment: CMS

- Fermilab is CMS' largest Tier-1 center
  - Half of the CPU, disk and tape available to experiments at FNAL is for CMS

- CMS is preparing for LHC Run 2
  - Higher center-of-mass energy and higher inst. luminosity leads to significantly higher event complexity
    - Re-engineering of software and computing components is nearing completion
    - Capacity of facilities will be ramped up by a factor of two

- Re-engineering done during shutdown:
  - Software CMSSW was optimized for higher number of PileUp events
  - Computing model was adapted to increase efficiency and flexibility
    - Dynamic and automated data placement on GRID sites
    - Increased usage of access to files through the wide area network (xrootd, AAA)
    - Opportunistic resource usage on non-CMS GRIDs/Clouds/etc. based on glideinWMS
      - Mutually beneficial for CMS and other Fermilab experiments
  - Data challenge starting in a few weeks

CMS worldwide:
62 computing centers



Level	US	Non-US	Total
Tier-0		1	1
Tier-1	1	7	8
Tier-2	8	45	53

SCD Software Support: Experiment Software

	scientific frameworks: Software Framework	scientific frameworks: software processing
CDF		
D0		
ArgoNeut	art	larsoft
Lariat	art	larsoft
LBNF	art	larsoft
LBNF 35T	art	larsoft
uBoone	art	larsoft
Minerva	gaudi	
MiniBoone		
Minos+	root/SRT	
Mu2e	art	
Muon g-2	art	
NoVA	art	
Darkside	art	
seaquest	custom	custom
coupp	custom	custom

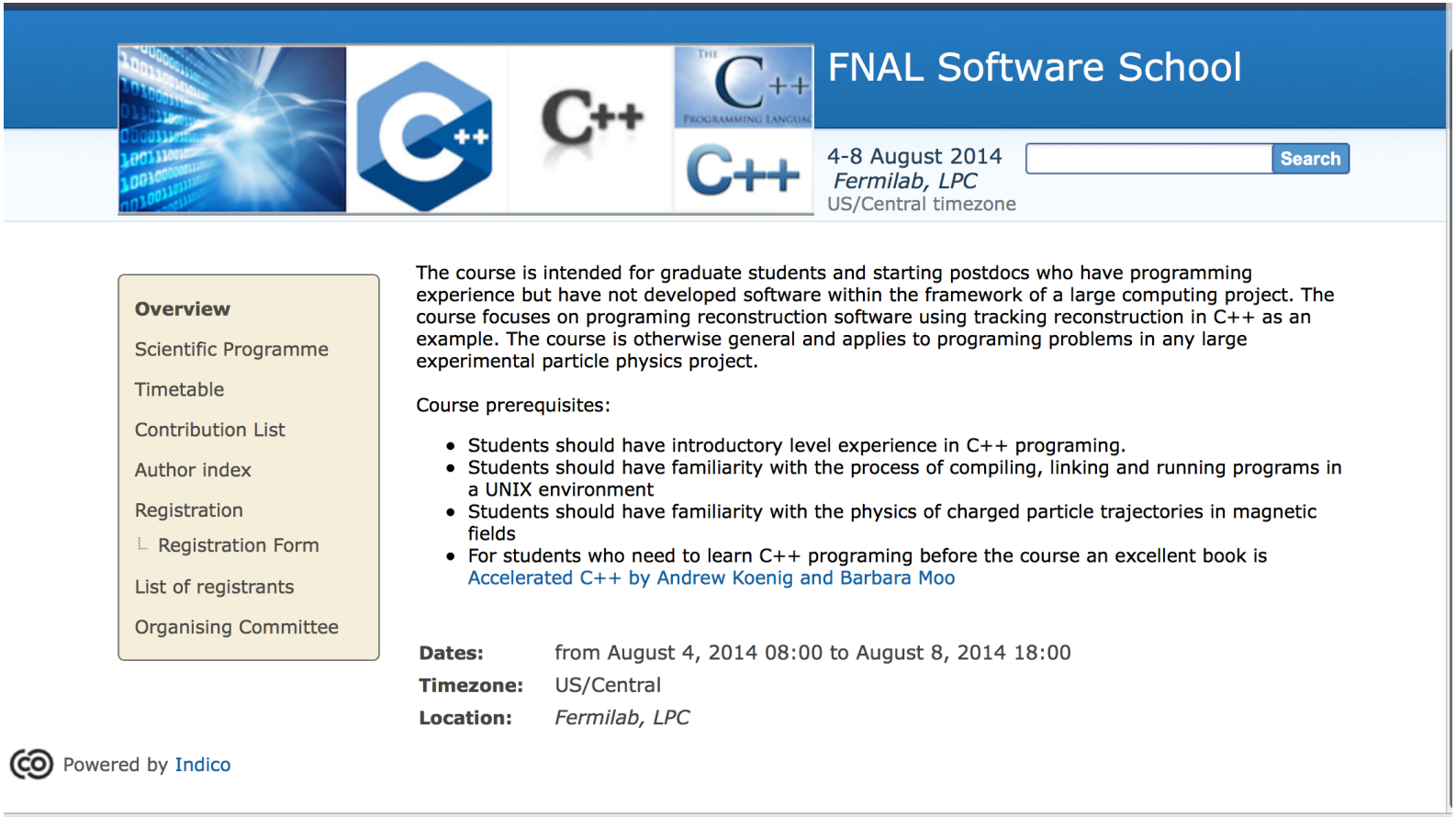
- SCD supports the core software frameworks for the Fermilab experiments: CMS and Intensity Frontier
  - CMSSW and Art are key ingredients for the success of the Fermilab experiments and continue to be very successful
  - Reason for success of both:
    - Quality of the framework code (advanced C++ that physicists use - not write)
    - In-house expertise
    - Truly collaborative effort between the framework authors and the stakeholder experiment(s)
  - Art was forked from CMSSW ➜ common concepts and base ideas

- New:
  - Multi-threading: one executable uses several cores in parallel to speed up processing and reduce memory requirements per core
  - Concrete help for the experiments:
    - ART workbook (documentation)
    - C++ courses (also see next slide about training)

- On top of ART: LarSoft
  - Specialized on reconstructing data from Liquid Argon detectors
  - Developed in international collaboration with contributors from many universities/ labs, amongst those Fermilab

Training - FNAL Software School - C++

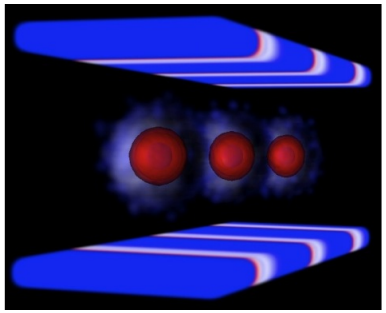
- Quote from OHEP task force on workforce development, “Lab-university partnerships can provide necessary research infrastructure, expert support, and direct training…”



SCD Software Support: Simulation & Modeling

Generators	
CDF	
D0	Pythia, Geant, Alpgen
ArgoNeut	Genie, Fluka
Lariat	Genie, Cry
LBNF	Genie, geant4
LBNF 35T	Genie, Cry
uBoone	Genie
Minerva	Genie
MiniBoone	
Minos+	Genie
Mu2e	
Muon g-2	
NoVA	Genie, Flugg, Cry and Coriska
Darkside	g4ds
seaquest	pythia 8, g4, custom
coupp	srim, geant4, mcnp

- Generators:
  - Strong involvement in Pythia
  - Genie (neutrino generator) ➜ began participation in the development team
- Simulation:
  - Strong Geant4 team working with experimenters and experiments
  - Working on the evolution of Geant4 to enable multi-threading



Wakefield simulation

- Accelerator modeling tools:
  - Development and support of:
    - single-particle beam dynamics code CHEF
    - multi-particle beam dynamics framework Synergia

Fermilab

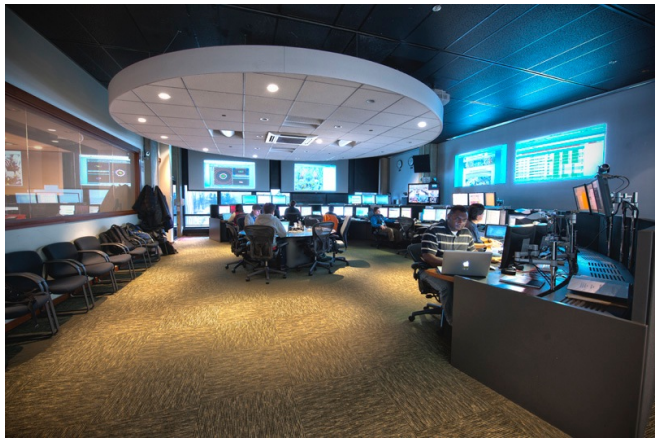
New: Fermilab's Offline Processing Operations Group

- Large scale workflow execution on GRIDs/ Clouds and other opportunistic resources is not easy!
  - Not every experiment has the required people and skills

- SCD is planning to support Fermilab's experiments also in executing workflows
  - Help the experiments execute their workflows on GRIDs/Clouds/etc.
  - Modeled after the CMS operations group

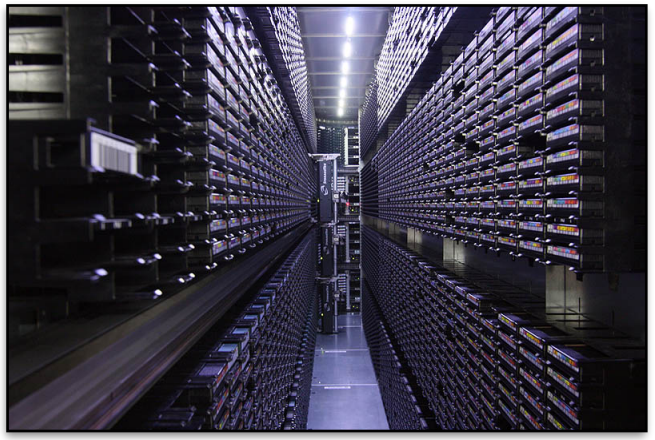
- Effort is ramping up, Fermilab Application Physicist position was opened to lead effort

Fermilab ROC



New: Scientific Data Archival Facility

- Develop and provide service to store data of DOE Office of Science experiments
  - Based on large experience on storing data on tape at Fermilab
- Designed as a pay for service model
- Pilot project with the University of Nebraska and their Biology department
  - Several other experiments/ Universities have expressed interest



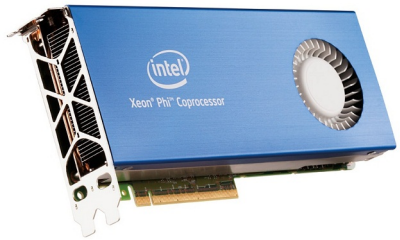
New: Preparing for the future: new hardware architectures

- Traditionally we are using Intel compatible hardware for processing, production and analysis of data

- New hardware architectures are emerging
  - Significantly increasing the speed of specific calculations
    - Systems with Graphics Processor Units (GPGPU)
    - Intel Phi coprocessors
  - Reducing the power consumption and getting more cycles per watt
    - ARM architectures
  - All the new architectures will have significantly less memory per core ➜ need to prepare software and hardware for transition

- Fermilab provides access to these new hardware architectures to experiments for evaluation and tests:
  - Intel Phi systems
  - GPU systems with the latest cards
  - Setting an ARM system now

Fermilab's Phi Cluster



Fermilab

Summary & Conclusions

- Fermilab's Scientific Computing is supporting all its experiments' computing needs:
  - The Catalog of Scientific Services is well received and used throughout the experiments
  - The core and simulation/modeling software efforts are integral parts of the experiments' strategies

- Looking into the future, Fermilab's Scientific Computing is planning to
  - Help experiments execute their workflows at large scales on GRIDs/Clouds and other resources
  - Enable other DOE experiments to store data on tape, building on the large experience of Fermilab
  - Help experiments evaluate and test new hardware architectures to prepare for the future

Glossary

- **GPGPU**: General-Purpose computing on Graphics Processing Unit
- **Intel Phi**: Brand name for Intel's Many Integrated Core Architecture or Intel MIC (pronounced Mike): multiprocessor coprocessor architecture developed by Intel
- **LT04**: Linear Tape-Open (or LTO) is a magnetic tape data storage technology, LT04 capacity: 800 GB
- **T10k**: Magnetic tape data storage format developed by the Storage Technology Corporation owned by Oracle; T10kc capacity: 5 TB
- **dCache**: is a system for storing and retrieving huge amounts of data, distributed among a large number of heterogeneous server nodes, under a single virtual file system tree with a variety of standard access methods.
- **EOS**: is a Xroot-managed disk pool for analysis-style data access
- **Xrootd**: is a fully generic suite for fast, low latency and scalable data access, which can serve natively any kind of data, organized as a hierarchical file system-like namespace, based on the concept of directory
- **BlueArc**: network-attached storage (NAS) systems that are sold either as appliances bundled with storage, or as "NAS heads" supporting third-party storage area network connected storage
- **Lustre**: is a type of parallel distributed file system, generally used for large-scale cluster computing
- **Gbps**: giga bits per second
- **Cold aisle containment**: Workernodes are arranged in rows. The aisles between the rows are alternating hot and cold, where the workernodes fronts point to the cold aisles to suck in cold air. Cold aisle containment consists of a physical barrier that allows the cold air to pool inside the aisle between the worker node racks.
- **ISO20k**: first international standard for IT service management to reflect best practice guidance contained within the ITIL framework
- **ITIL**: The Information Technology Infrastructure Library (ITIL) is a set of practices for IT service management (ITSM) that focuses on aligning IT services with the needs of business.
- **GENIE**: (Generates Events for Neutrino Interaction Experiments) is a universal object-oriented neutrino MC generator supported and developed by an international collaboration of scientists whose expertise covers a very broad range of neutrino physics aspects, both phenomenological and experimental.
- **CRY**: cosmic ray generator used by neutrino experiments to study backgrounds
- **GEANT4**: (for GEometry ANd Tracking) is a platform for "the simulation of the passage of particles through matter," using Monte Carlo methods.
- **SAM**: SAM is a data handling system organized as a set of servers which work together to store and retrieve files and associated metadata, including a complete record of the processing which has used the files.
- **SVN**: Apache Subversion (often abbreviated SVN, after the command name svn) is a software versioning and revision control system distributed as free software under the Apache License.
- **GIT**: is a distributed revision control and source code management (SCM) system with an emphasis on speed.
- **docDB**: is a powerful and flexible collaborative document server.
- **redmine**: Redmine is a free and open source, web-based project management and bug-tracking tool.
- **daq**: Data Acquisition
- **FIFE**: FabrIc for Frontier Experiments, provides collaborative scientific-data processing solutions for Frontier Experiments
- **OSG**: Open Science Grid: provides common service and support for resource providers and scientific institutions using a distributed fabric of high throughput computational services.
- **AAA**: "Any data, anywhere, any time": CMS global data federation based on Xrootd