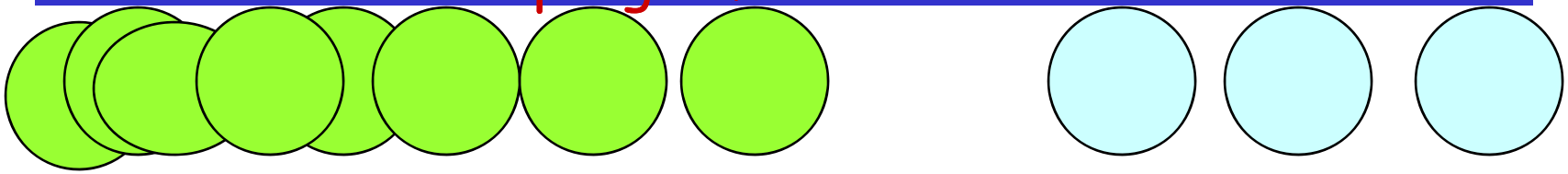
Computing Strategy
FRA Physics Visiting Committee
April 9, 2009

Victoria White
Assoc. Lab Dir. for Computing Science and Technology (acting)
CIO
Head, Computing Division

Computing in support of ALL scientific programs and the lab



Scientific Programs

Lab Divs & Sections & centers

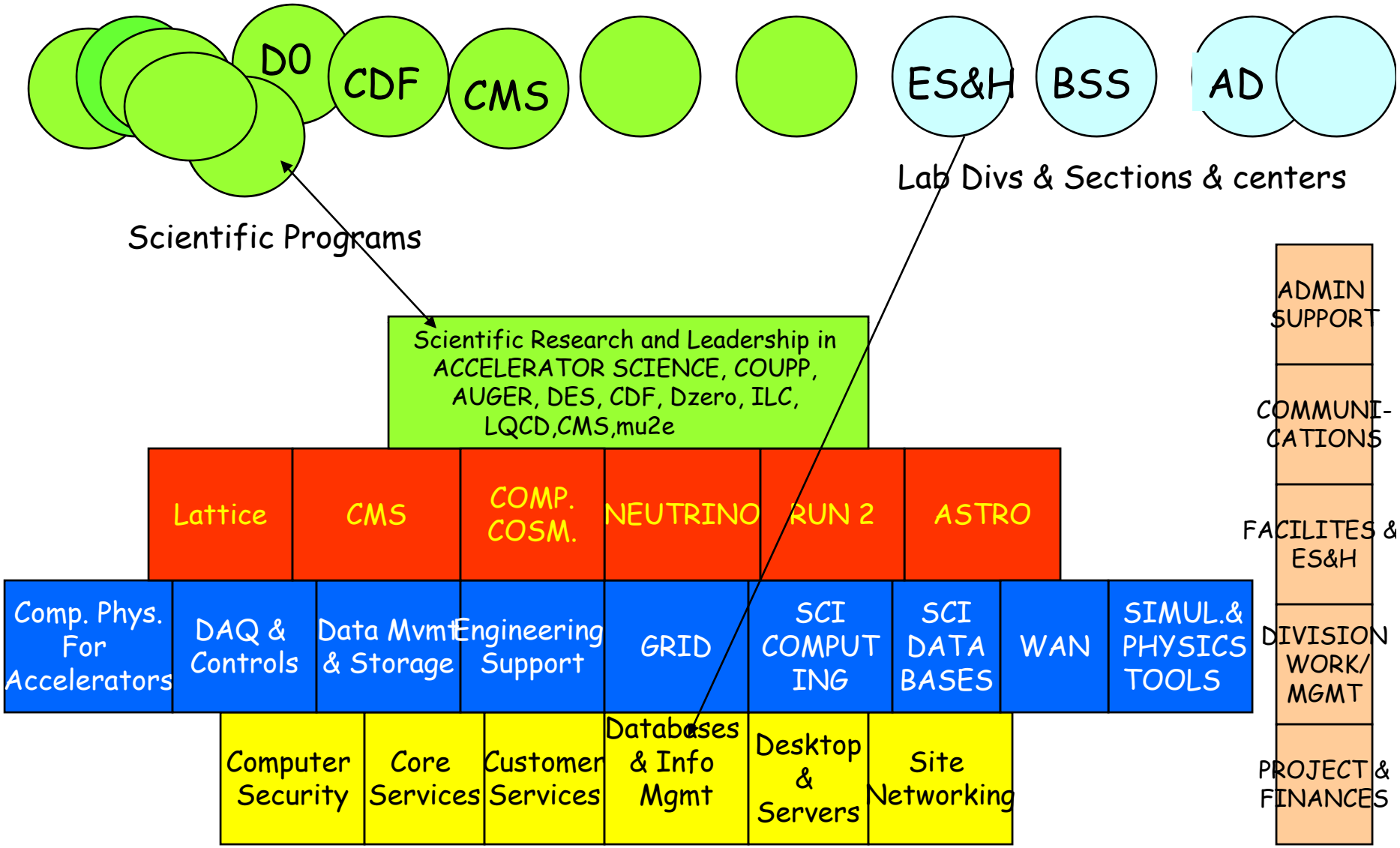
Science & Scientific Leadership 35 scientists (~ 9 FTEs)

Program specific operations, development and investigations (~72 FTEs of which ~ 30 on CMS, 15 on Run II, 12 ASTRO)

Tools, Services and Facilities development and support -for Scientific programs (~100 FTEs)

IT Infrastructure for the lab and scientific program (~65 FTEs) and Business Systems (~25 FTEs)

"Layers" of Computing



Scientific Stakeholders and customers

Scientific Program/Project	IT services	Computing	Storage and Disk	Significant use of GRID	Trigger and DAQ	Acclerator Modeling	Dedicated devel. or operations	Scientific
CDF	X	X	x	X	X		X	X
Dzero	X	X	X		X		X	X
CMS	X	X	X	X			X	X
MINOS	X	X	X				X	X
Lattice QCD Computation Facility	X	X	X				X	X
DES	X	X	X				x	X
JDEM	X							X
Auger	X	X	X		?			X
CDMS	X	X	X		X			
SDSS (finished)	X	X	X					X
MiniBooNE	X	X	X	X				
MINERvA	X	X			X		X	X
MicroBooNE	X	X	X		?			
SciBooNE, Argoneut	X				?			
Accelerator Science	X	X	X	X		X		X
ILC	X	X	X	X	X			X
Nova	X				X			
COUPP	X	X	X		X			X
Project X	X				?			X
mu2e	X				X		X	X
Dusel	X				?			

Key elements of Computing Strategy

- Scientists that provide bridge and requirements for scientific programs and who take leadership roles in development of future programs and operations of current programs and do research.
- Strong layer of core competencies that leverages expertise across programs and re-uses techniques and tools
- Encourage use of shared services and tools as much as possible
  - Have to do this for economies of scale with so many different scientific programs to take care of
- Strive for robust, efficient, well monitored, planned several years out: Facilities, networks and all bottom layer of IT services
- Continuously invest in R&D, technology refresh, and planning for the future in all layers
- Carefully redeploy people from current scientific program to future scientific program

Facilities, Power and Cooling

Feynman Computing Center (FCC) • 3 computer rooms • 17891 ft^2 24" raised floor (typical) • 750kVA UPS for all computing systems • 1.5 MWatt standby diesel generator • 270 tons*; Chilled Water cooling • <4kW/cabinet (typical) • 220 rack positions • Security system • Fire detection/suppression systems • Under floor water detection system	**Lattice Computing Center (LCC)** • 2 computer rooms • 2841 ft^2 12" raised floor (typical) • No UPS / ~850kVA total capacity for computing • Tap boxes to connect portable generators • 210 tons*; Chilled Water & air cooling mix • 4kW & 8kW/cabinet (typical) • 76 rack positions • Security system (access only) • Fire detection/suppression systems
Grid Computing Center (GCC) • 4 computer rooms / 2 communication rooms • 10384 ft^2 24" & 36" raised floor (typical) • 1000kVA * 3 UPS for all computing systems • Tap boxes to connect portable generators • 1000 tons*, air cooled units • 10kW & 14kW/cabinet (typical) • 235 rack positions • Security system • Fire detection/suppression systems (including VESDA)	**Fiber Central (FC)** • 1 communication room • 792 ft^2 12" raised floor (typical) • 20kVA UPS for computing systems • Standby generator • 10 tons*; Chilled Water cooling • <3kW/cabinet (typical) • 12 rack positions • Security system (access only) • Fire detection/suppression systems

* One ton = 12,000 Btu/h; One Btu/h = 0.293 watts; One ton = 3.516kW

2008 Availability for each data center

	Availability Goals		Planned and Unplanned Outages		Unplanned Outages Only	
	Availability (goal %)	Downtime (hours)	Availability (actual %)	Downtime (hours)	Availability (actual %)	Downtime (hours)
FCC	**99.70**(1)	26.28	99.99	0.19	**99.99**	0.12
GCC	**99.66**(1)	30.00	99.78	19.21	**99.88**	10.16
LCC	**98.50**(1)	131.40	98.53	128.59	**98.58**	124.53

Full report on data center availability available in cd-docdb
https://cd-docdb.fnal.gov:440/cgi-bin/ShowDocument?docid=3123

CORE IT SERVICES

IT services on which the entire lab and all of the scientific programs rely. These include site network services; email; windows, mac and linux infrastructures; computer security; document databases; web servers; plone; indico; code repositories; file systems; backup and restore; service desk; database administration, video conferencing support and more….. (almost all overhead funded)

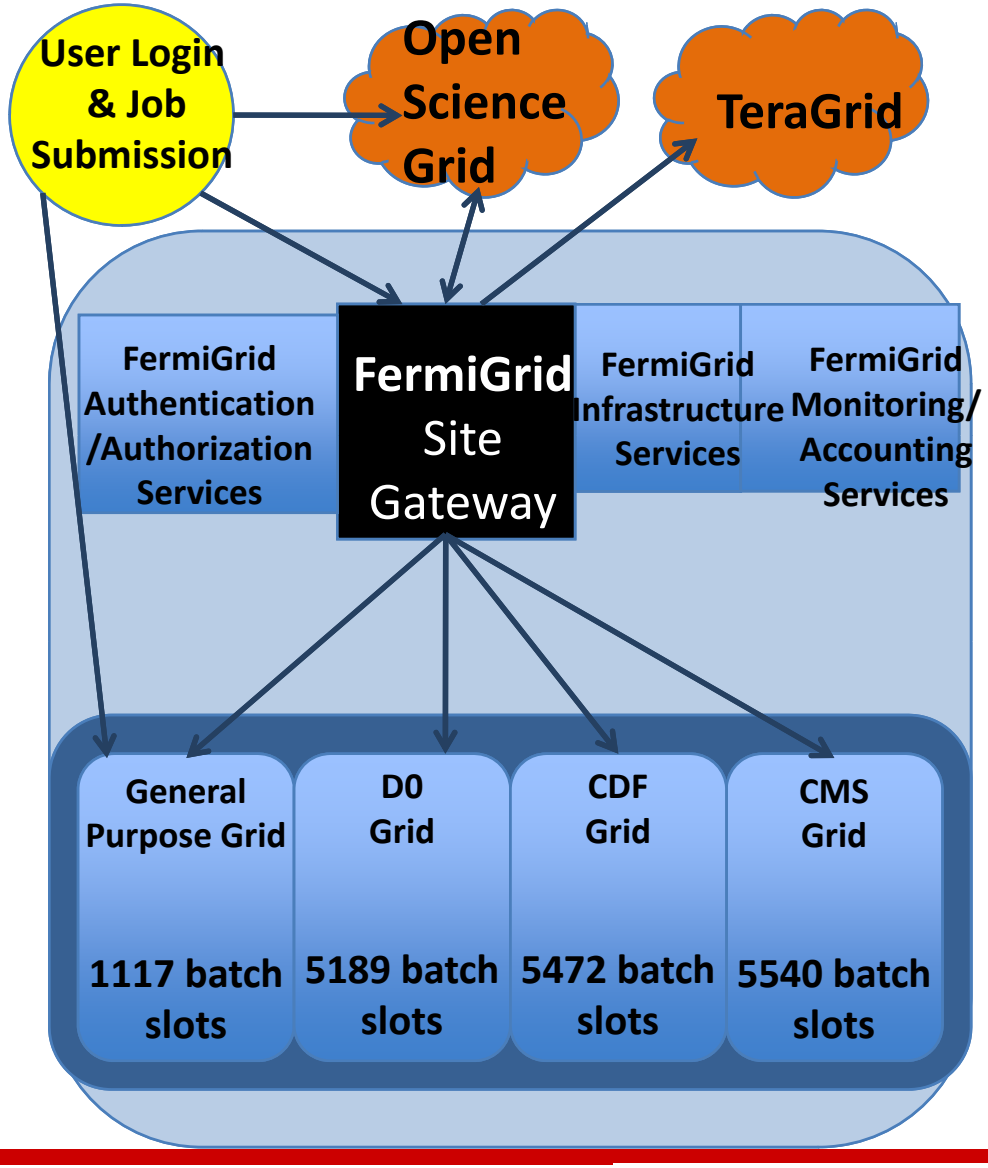
Core IT Services Strategies

- IT consolidation across the lab
  - Recent lab reorganization and CIO official role
- We have a roadmap for the lab to become ISO20000 certified in IT service delivery by FY2011
  - Invest now - in order to save later - in order to apply more resources to the scientific program and to modernization of the lab's information systems and services
  - Standardize, formalize processes, selectively outsource - in the areas where diversity and specialized solutions are not needed
    - save that for the scientific arena

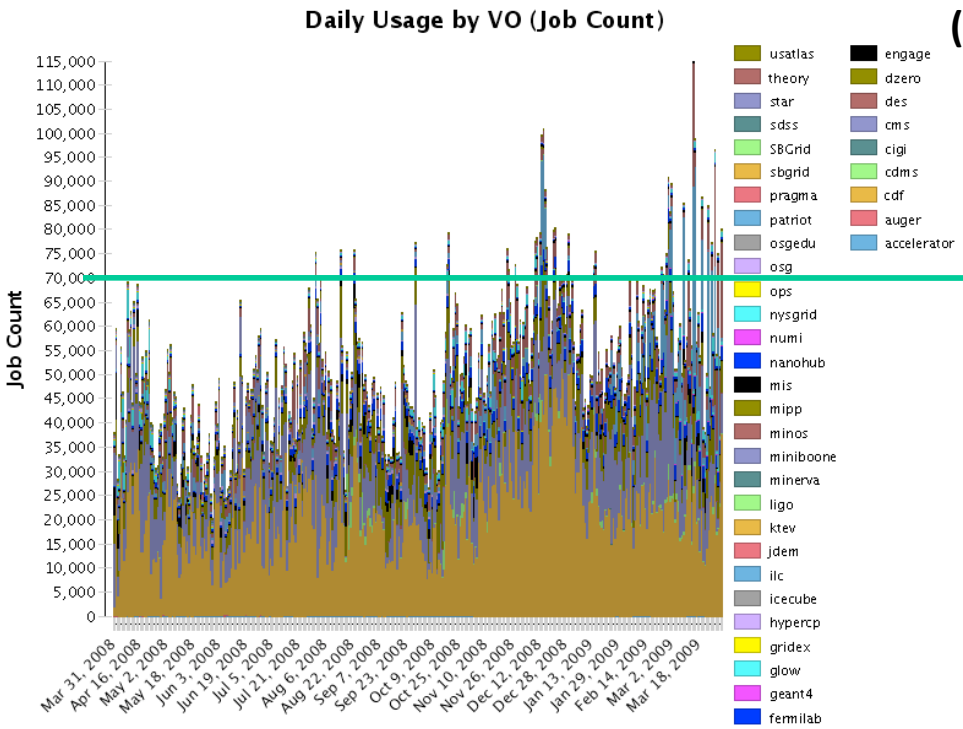
SHARED SCIENTIFIC SERVICES, TOOLS AND EXPERTISE

The shared services, tools and expertise on which <u>all the current experiments and programs are built</u> serve as a strong, reusable and extensible base of services and expertise for the future programs of the lab.

FermiGrid – shared computational resources



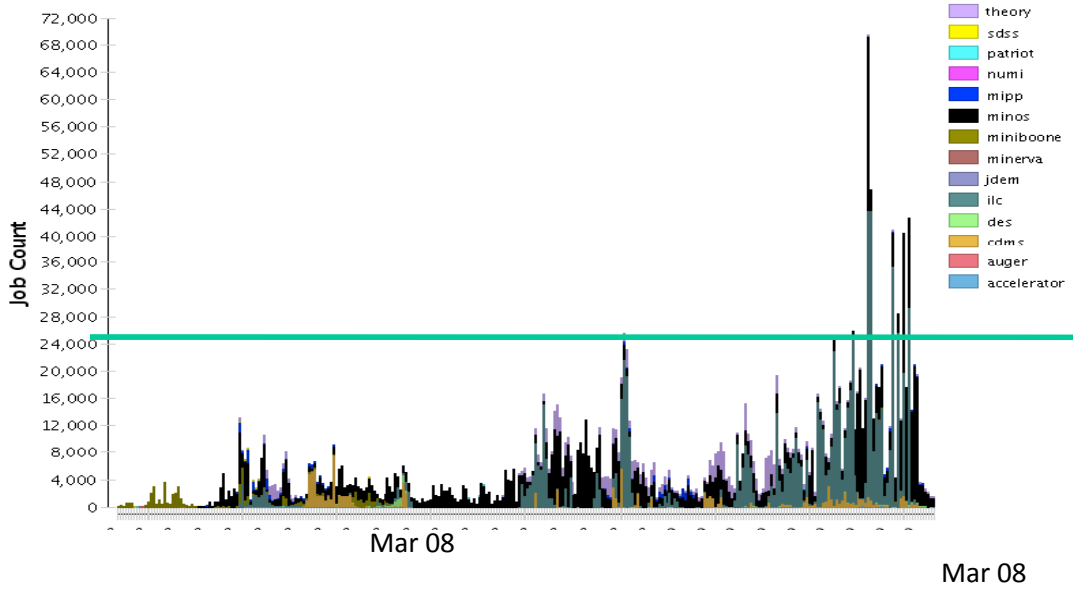
Total batch slots: 17,318



(all FermiGrid)

70,000 jobs per day

cdf
cms
d0

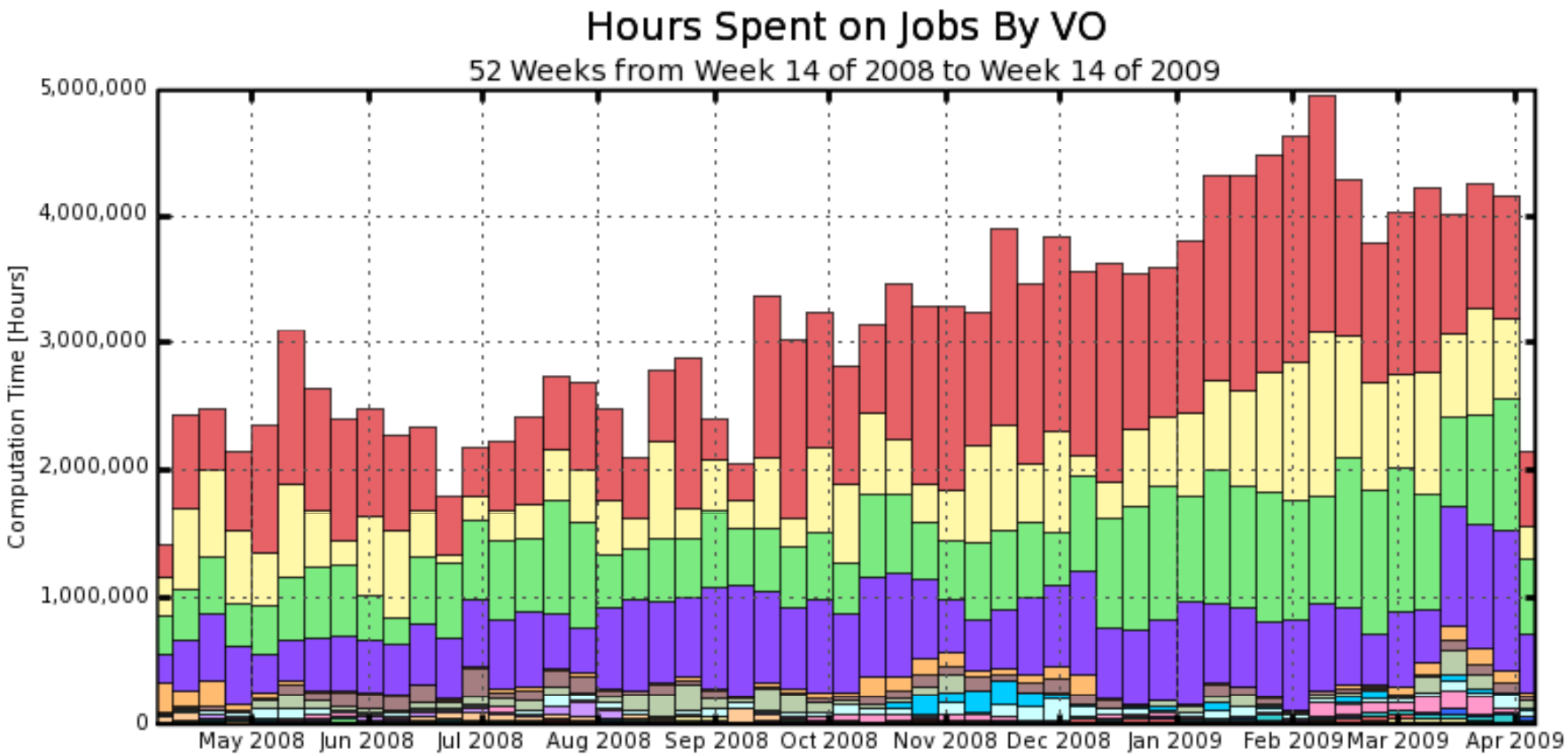


24,000 jobs per day

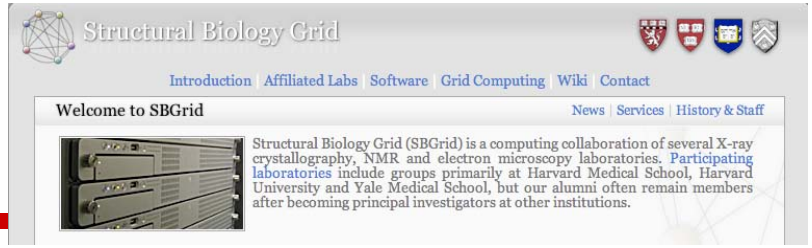
minos
ilc
cdms

Mar 08 Apr 09

National shared infrastructure and US Contribution to Worldwide LHC Computing Grid



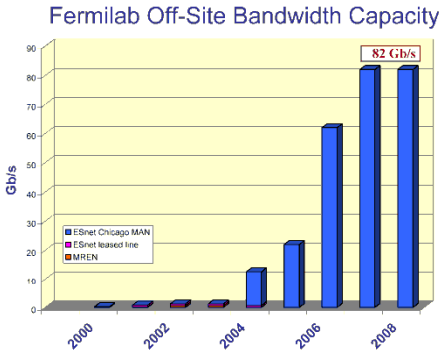
Adopted as Infrastructure for Multi-Institutional Structural Biology Community

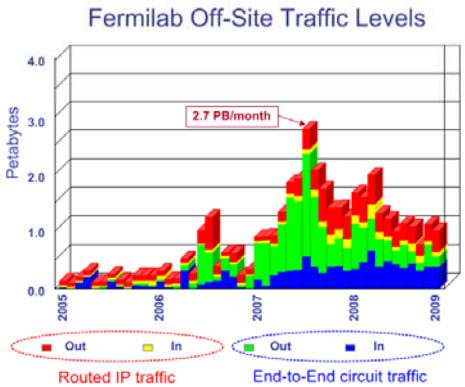


Wide-Area Network Strategies

1) Provision Scalable redundant WAN

- Off-site bandwidth increased by factor of 4 over past 3 years
  - Supporting high data movement requirements of experiments
  - Scalable MAN architecture provides cost-effective growth path for future needs





2) Work with high bandwidth expts Keep moving forward with technology

- Pioneering in use of data circuits:
  - 14 end-to-end circuits deployed
  - Isolates high impact (LHC) data movement from general internet
  - Provides predictive bandwidth
- WAN-related R&D efforts:
  - Lambda Station path selection service
  - Structured WAN performance analysis methodology
  - perfSONAR network performance monitoring service

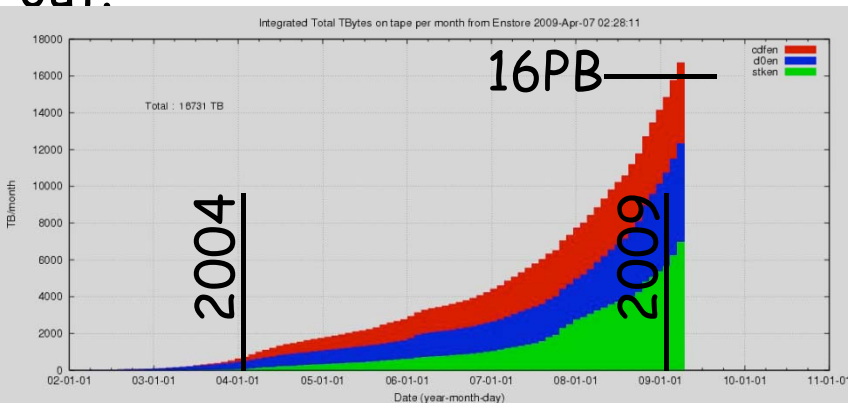
Shared Data Movement and Storage Services

Tape

Three "instances" of Enstore tape system: CDF, D0, CMS+ others, all operated by Storage group across 10 shared tape robots

Current data stored: 16PB

Current capacity: 40PB, plus some older robots being phased out.



Disk

Four "instances" of dCache disk storage system:
CMS, CDF, Lattice QCD, all others.

Largest has 3PB of usable disk.

Fermilab is a core member of the dCache collaboration; dCache is used by many LHC physics centers, including all USCMS Tier 2 sites.

Storage Futures and Strategies

Tape

Data is migrated to new media every 2nd technology generation.

- Reduces slot & cartridge count for old data by approx. 75%/6yr.

Enstore collaboration is expanding.

Scaling up performance for higher request rates.

Adding support for small files. "Eventually, all files are small."

Disk

Investigating new high-throughput file systems, such as GPFS, Lustre, Hadoop, QFS.

Deployed Lustre for HPC: Computational Cosmology & LQCD.

Scaling up dCache's Storage Resource Manager (SRM) interface for higher request rates.

Scientific SoftwareTools and Services

- **Performance Tools: SimpleProfiler**
  - Non-intrusive profiling with graphical analysis tools
  - Database for organized collection of results from multiple runs
  - Geant4 is initial customer
- **ROOT** – integrated software tool for analysis and display of HEP data
- **Geant4** – Hadronics improvement – Performance and QA
- **Pythia** – General-purpose Event Generator tool
- **Generic Framework** – leveraging existing frameworks
  - Reasonably full-featured; initial customer mu2e
- **Workflow tools** **(for Lattice QCD, JDEM?)**
- **Coding for experiments**
  - CMS Framework,
  - NoVA DAQ
  - Scientific Databases and applications (DBAs are a core IT service)
- **C++** – Reviews, Consulting, Courses, Standard Committee

Trigger and DAQ: Mission Need

- **Versatile Data Links**
  Increase data rates while reducing detector mass.

  Joint ATLAS/CMS long term R&D for SLHC.

- **Computing Intensive Triggers and DAQ**
  "Trigger-less architectures": merging of computing and network fabrics.

  mu2e, LAr, test-beam future collider readout systems.

- **Computing Intensive Front-ends**

  Estimator-based front end signal processing: ultra-low noise readouts.

- **Computing Intensive Controls**

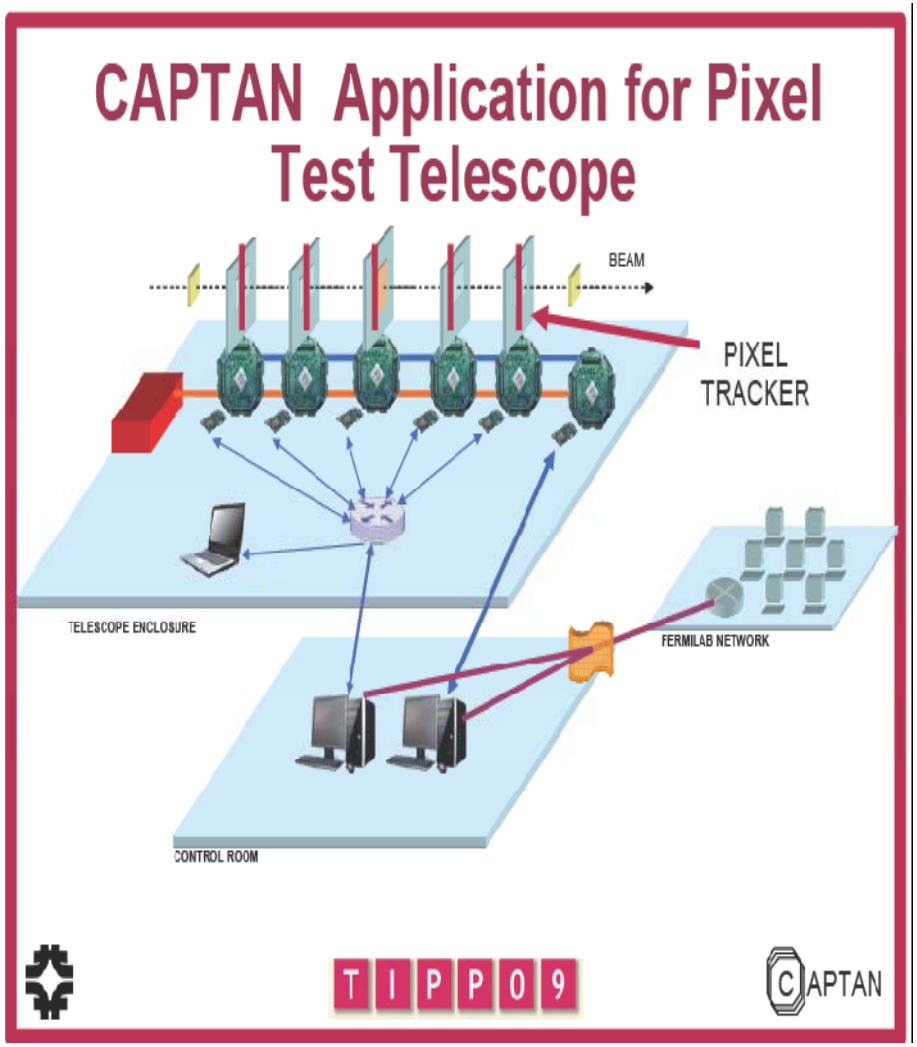
  High performance LLRF: ILC & Project-X.

Powerful Data Driven DAQ Architectures

CAPTAN: **C**ompact **A**nd **P**rogrammable da**T**a **A**cquisition **N**ode

Events are "built" and streamed through now standard network fabric (GBit Ethernet)

Many interested customers : MTest, INFN, IIT, Brown,….



Trigger and DAQ: Opportunities and Challenges

- **Versatile data links:**
  R&D on low-power and low-mass industry standards that are sufficiently tolerant of collider environments.

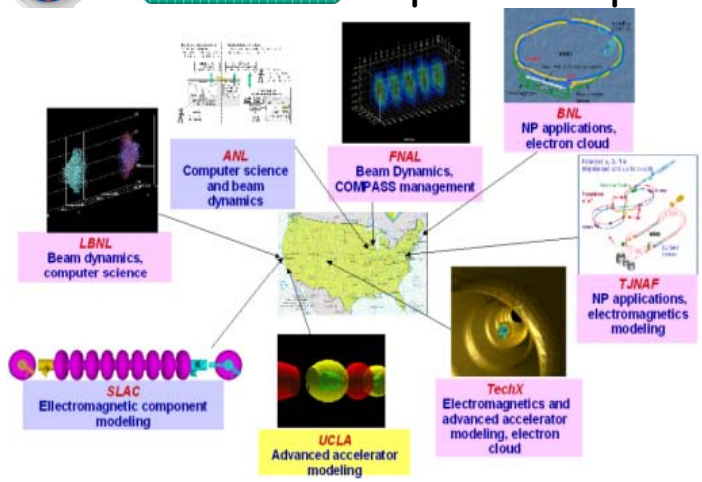
- **Computing Intensive Triggers and DAQ:**
  Many-core paradigms, Graphical Processors. Strong synergies with HPC and "offline".

- **Computing Intensive Front-end and Controls:**
  Complex algorithms, configurations and conditions.

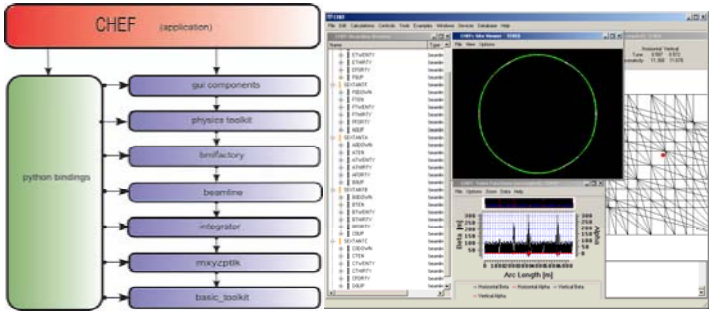
Accelerator modeling tools

- Leading ComPASS: development of High Performance Computing (HPC) accelerator modeling tools
  - Multi-physics, multi-scale for beam dynamics; "*virtual accelerator*"
  - Thermal, mechanical, and electromagnetic; "*virtual prototyping*"
- Development and support of CHEF
  - general framework for single particle dynamics

SciDAC Scientific Discovery through Advanced Computing https://compass.fnal.gov/



Collaborative **H**ierarchical **E**xpandable **F**ramework



Accelerator modeling tools development

- Strategy: apply state-of-the-art computing to enable state-of-the-art accelerator modeling
  - Move beyond the age of one code per one effect
  - Develop modeling capabilities to be integrated in realistic applications
    - Advance to muti-physics, multi-scale models
  - Develop and apply expertise in parallel computing from clusters with O(100) processors to capability machines with O(100,000) processors
    - Synergistic with Lattice QCD effort
  - Leverage experience with providing support for user-oriented software

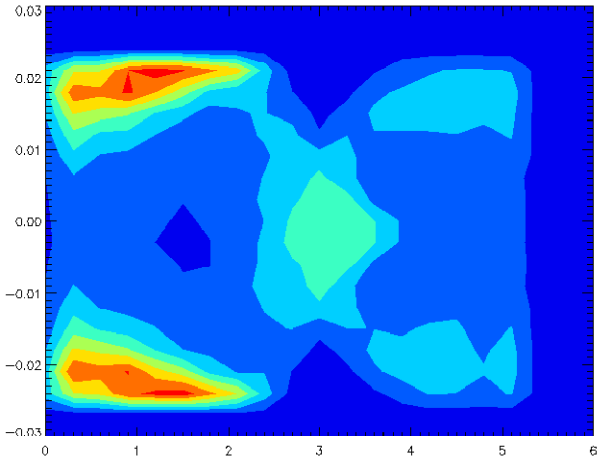
LAB'S FUTURE SCIENTIFIC PROGRAMS

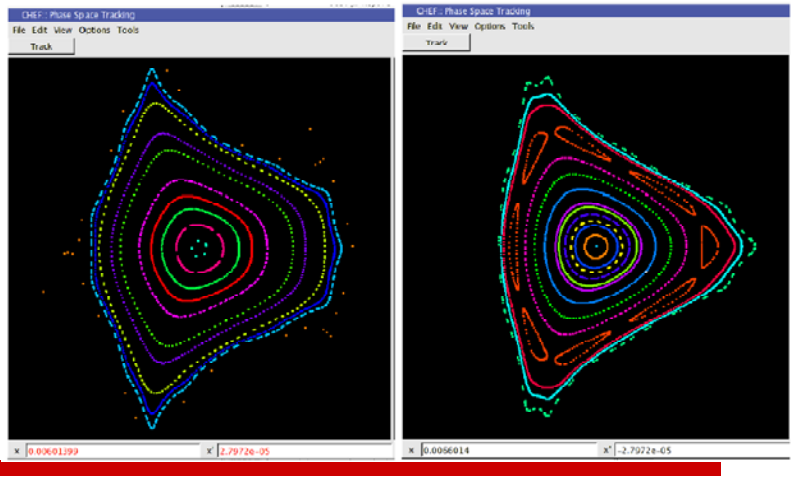
LHC is a big driver for computing, as well as Lattice QCD and Computational Cosmology (and of course Run II expts)

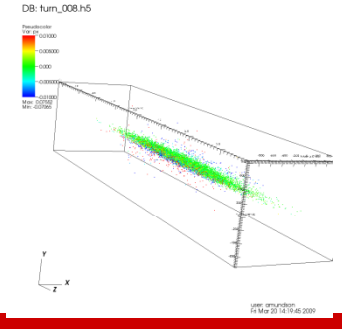
But in line with the focus of this FRA review we look at computing support for a few specific future scientific programs

Accelerator Modeling Applications

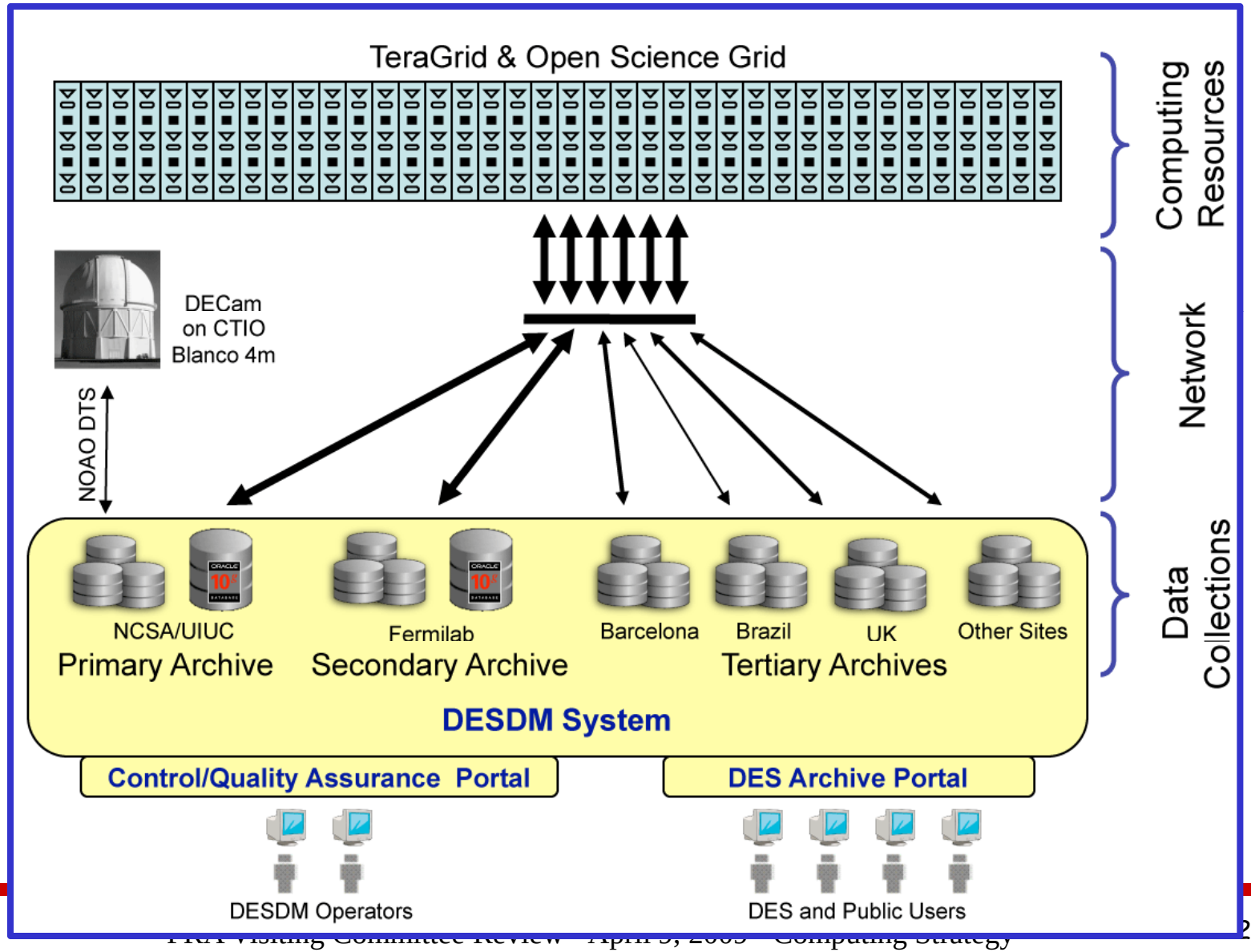
- Model electron cloud detector in the Main Injector (microwave propagation)
  - Multi-scale, multi-physics
- Model resonant extraction including space-charge at the debuncher for the proposed Mu2e experiment







DES Computing Architecture



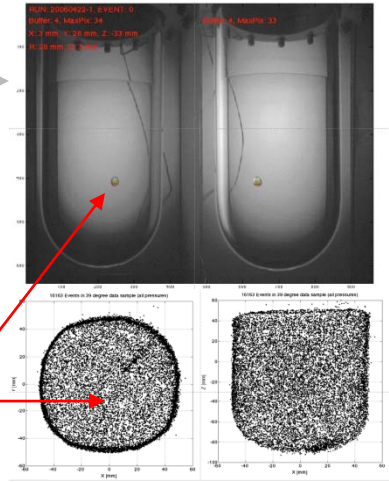
DES Archive support at FNAL

- Fermilab will host a copy of the raw and processed data and a live standby of the database.
  - Currently have complete copy of the Data Challenge 4 (DC4) raw and processed data in our Blue Arc storage space.
  - have set up a standby database that is a copy of the development system at UIUC
  - Will soon have a standby of the entire DC4 catalog
- Several CD scientists are involved in the DECam project and the Data management project (2.5 FTEs).
- Computational Physics Developer support (1+ FTEs -> 2+ FTEs) for
  - the DES simulation effort
  - Data management pipeline
- Programming for the online database development (0.5 FTE).
- DES computing systems are supported by the same system admins who handle all the CDF, D0 and general purpose farm computational resources

➔ Strategy – gradually extracting dedicated resources from Run II and MINOS to provide dedicated help for future expts

COUPP

A Dark Matter Search using a Bubble Chamber. Best in the world limits with an engineering prototype. New chambers deploying soon.



Experimental Program

- 2 Kg prototype chamber run in 2006 yielding Science paper Feb 2008.
- Chamber rebuilt as 4 Kg with synthetic quartz; deployed ~June.
- New 60 Kg chamber ready by fall.
  - Initial commissioning in MINOS hall here at Fermilab
  - Move deep underground (Soudan? / SNOLAB?) ASAP
- The goal is a 500 Kg chamber operating deep underground yielding world's best Dark Matter limits
- The key is control of all backgrounds.

Computing Division Contributions

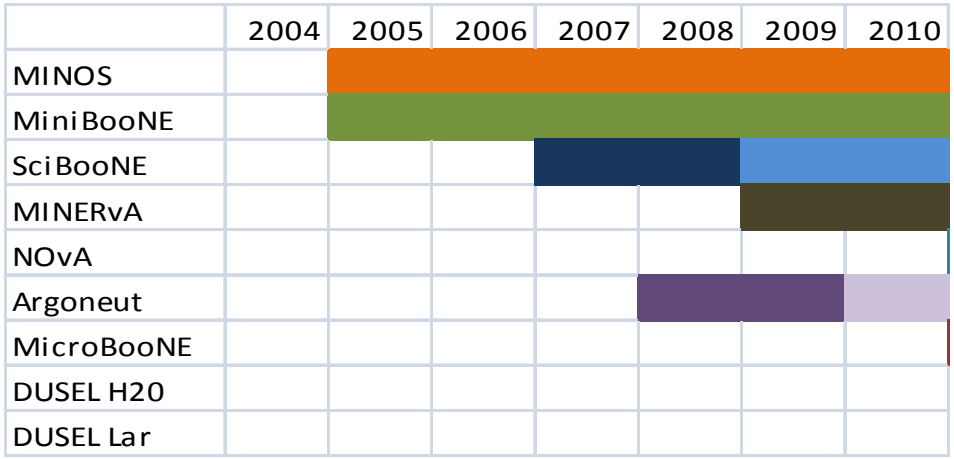
- CD scientists and staff are responsible for the engineering and science for Trigger, Data-acquisition, Storage, and Analysis for all of these chambers.
- The challenge is indefinite operations in unattended, remote, deep underground sites (hostile: hot, wet, electrical power most of the time).
- We are engineering dedicated trigger electronics and a system with a minimum number of components for maximum robustness.
- Present efforts are directly extendible to 500 Kg and/or multiple chambers.

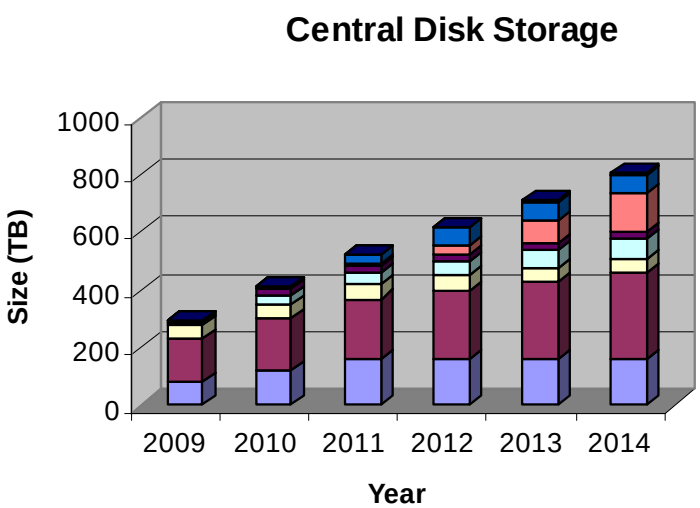
Recent Computing for Neutrino Workshop : Conclusions

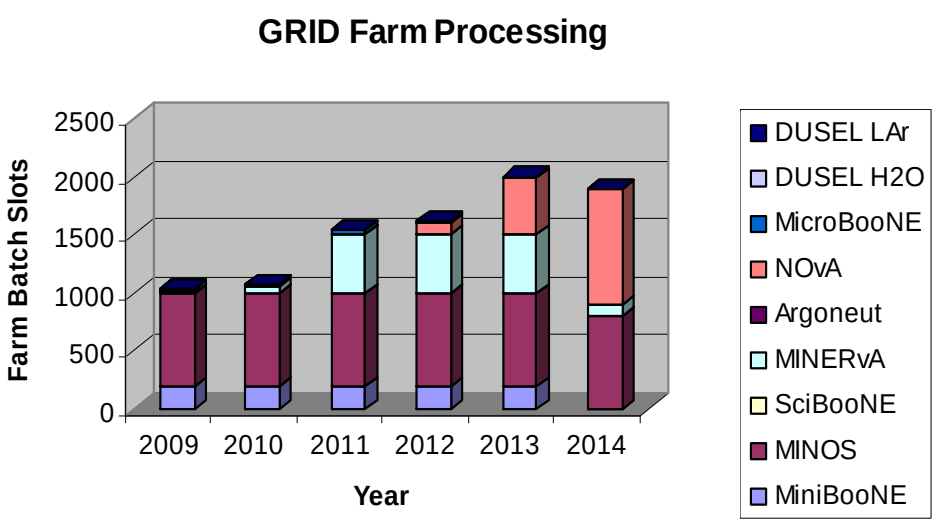
- Technology sets data size
  - Phototube based detectors (MINOS, SciBooNE, MINIBooNE, MINERvA, NOvA , DUSEL H_2O): 10 – 100TB/yr
  - Liquid Argon detectors (Argoneut, MicroBooNE, DUSEL): 1 - 100 PB/yr.
- Experiments within each class (PMT vs. LAr) are pretty similar in terms of computing needs
- All use common tools
  - ROOT
  - GEANT
  - GENIE/NUGEN
  - SL4/5
- Each experiment's core software is unique
- Goal to work together and share more in the future

Neutrino program: Computing Resource Estimates

- Schedule for each Project
- Projections for Central Disk Storage and GRID Farm usage

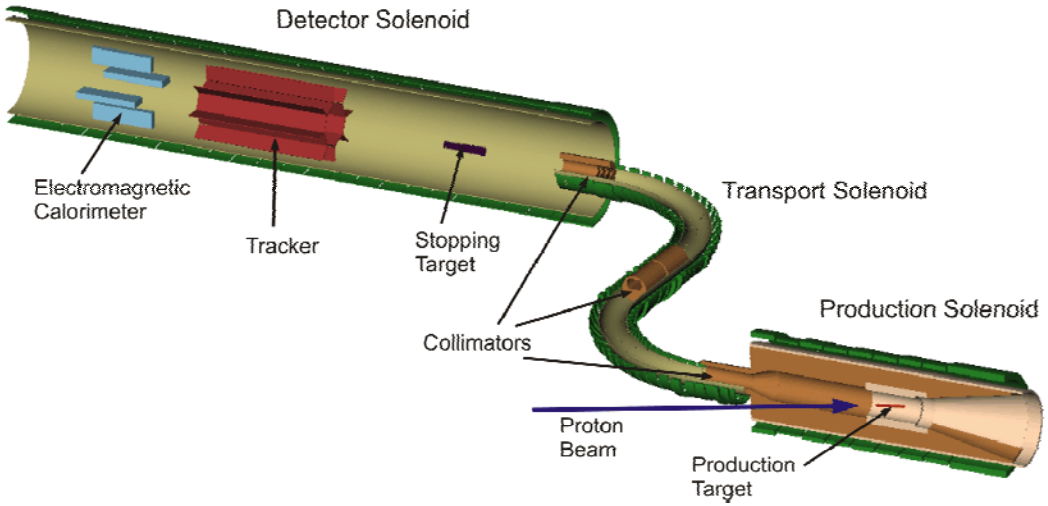






The Mu2e Experiment

Search for neutrino-less conversion of a muon to an electron in the Coulomb field of a nucleus.

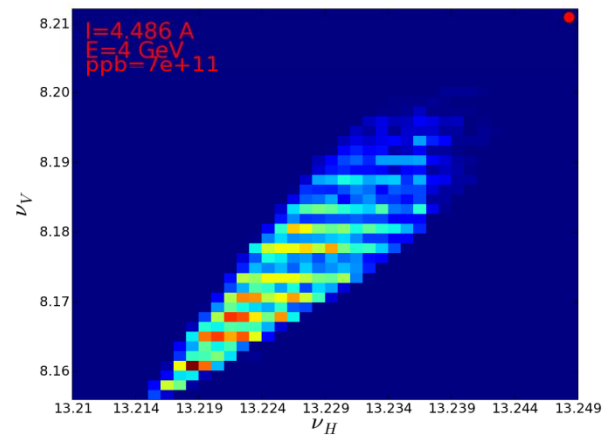


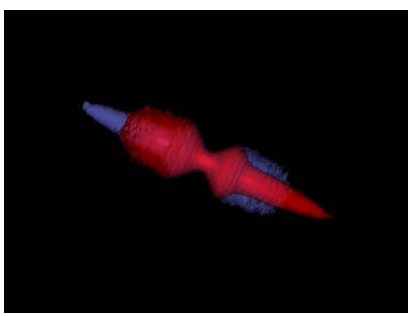
- Timeline: CD0 (≈summer 2009); CD3(≈Q4 2012); CD4(≈Q4 2016).
- CD providing infrastructure software (based on CMS framework)
  - Alpha release delivered and in use; ongoing development.
- CD providing Geant4 assistance: getting underway now
- Will use CD resources: GRID, tape robots, NAS disk, networks, docdb, code management, and more
- CD scientists active members of Mu2e (Kutschke, Tschirhart).
- Will request CD support for Trigger and DAQ.

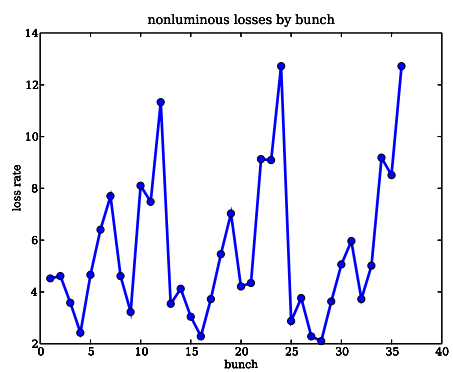
EXTRA SLIDES

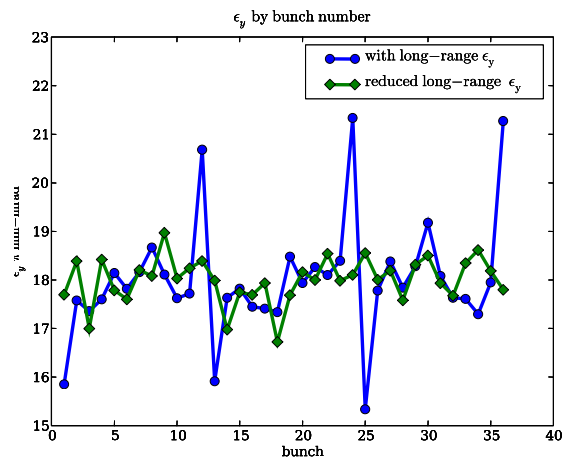
Accelerator Modeling Applications

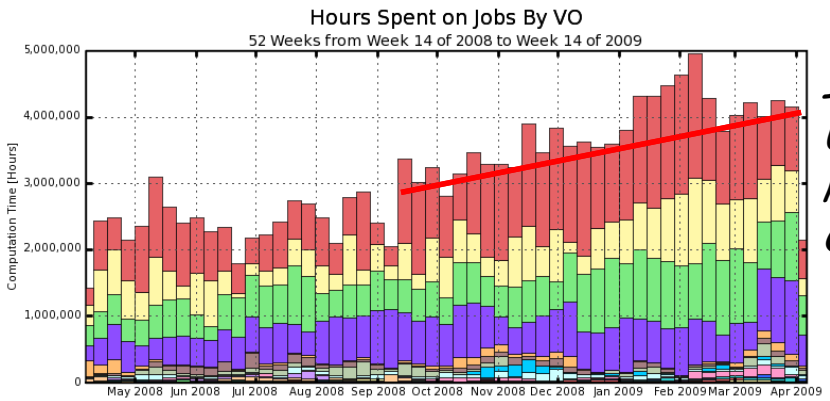
- Model space charge effects at the CERN PS2
- Model beam-beam & impedance effects with 36 on 36 bunches at the Tevatron



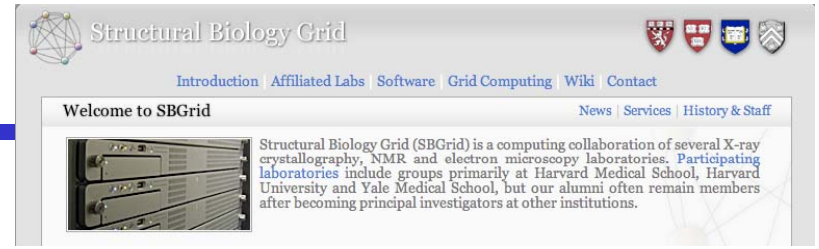








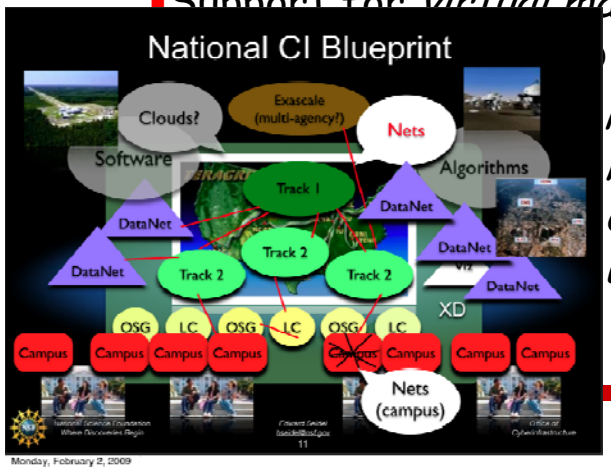
Increased Usage by HEP and Others



Adopted as Infrastructure for Multi-Institutional Structural Biology Community

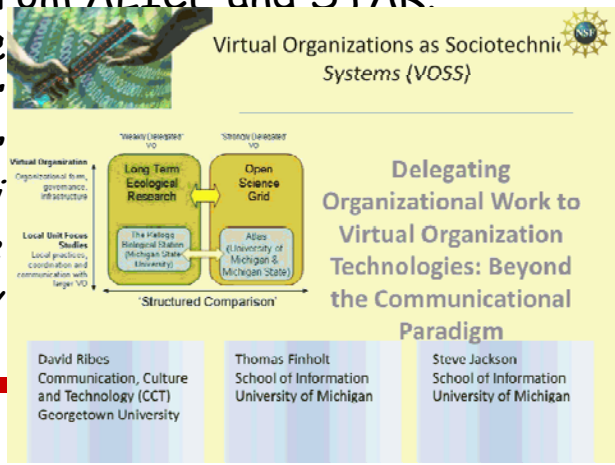
Efforts Continue for Future Needs:

- Widespread deployment of *more efficient "pilot" job technologies*.
- *Ramp up of LHC Tier-3s:* ~25 registered of the ~70 expected in the next year.
- *Security technologies:* Shibboleth, OpenID, Bridge-CA; CS/Math analysis of risk/ response.
- *Storage & data management:* currently adding Hadoop; network management/monitoring.
- *Adapt/adopt new services:* authorization policy, CREAM compute element.
- *Portals for small communities:* task force for Nanohub, Engaging new communities.
- *Ad-hoc, dynamic commuity collaboratories* needed: e.g. by SNS
- Support for *virtual machine technologies*: see reports from ALICE and STAR.
- Blue-Gene, commercial clouds, ...



Part of the NSF National CyberInfrastructure Blueprint

Being Studied as Exemplar Collaborative Community

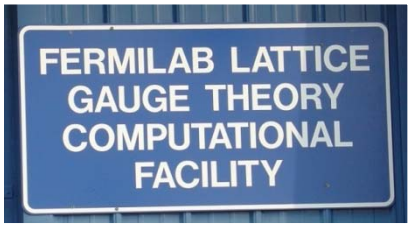


Open Science Grid

The Architected Model for Locating Computing in the Data Centers

- Feynman Computing Center (FCC)
- High Availability Services
- Networking, Computer Security, BSS, Email, Web, Databases, Core Site Services
- Tape Robotic Storage
- UPS & Standby Power Generation

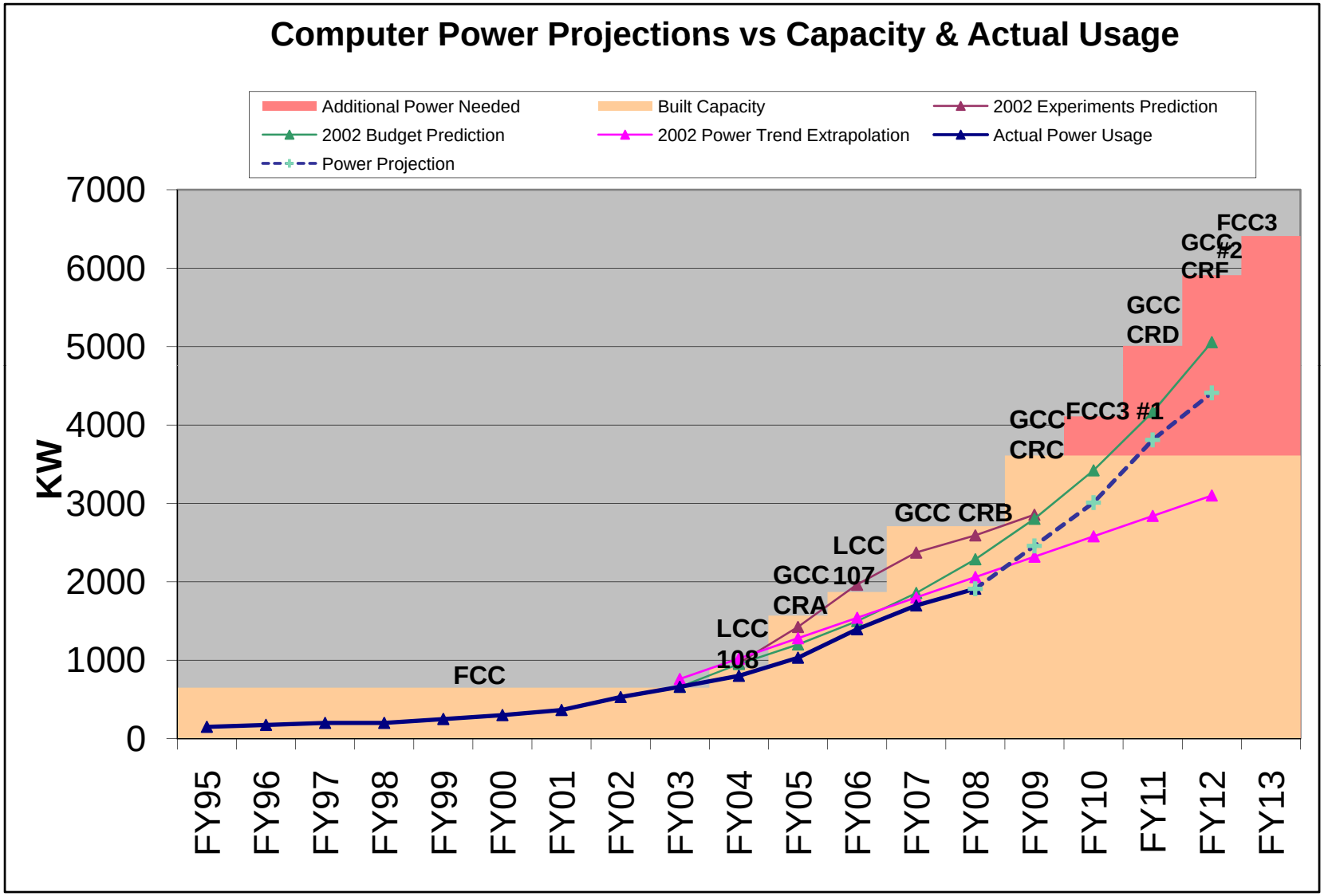
- Grid Computing Center (GCC)
- High Density Computational Computing
- CMS, RUNII, GP batch worker nodes
- Lattice HPC nodes
- Tape Robotic Storage
- UPS & taps for portable generators



- Lattice Computing Center (LCC)
- High Performance Computing (HPC)
- Accel. Sim, Cosmology nodes
- No UPS; taps for portable generators

Facilities Strategy

- Refurbish computer rooms
- Planning for new High Availability computing room in Feynman starting this summer
- Feynman Computing Center – generator backed high availability computing room – soon to be two.
- Grid Computing Center – for compute intense resources – 3 computing rooms + tape room
- Lattice Computing Center – less reliable
- Pay attention to "green" in purchase of computer servers
- Keep planning several years out
- Work like mad to keep it all up and running


Computer Power Projections vs Capacity & Actual Usage
Additional Power Needed
Built Capacity
2002 Experiments Prediction
2002 Budget Prediction
2002 Power Trend Extrapolation
Actual Power Usage
Power Projection
KW
7000
6000
5000
4000
3000
2000
1000
0
FY95
FY96
FY97
FY98
FY99
FY00
FY01
FY02
FY03
FY04
FY05
FY06
FY07
FY08
FY09
FY10
FY11
FY12
FY13
FCC
LCC 108
GCC CRA
LCC 107
GCC CRB
GCC CRC
FCC3 #1
GCC CRD
GCC CRF
GCC #2
FCC3