Fermilab Site Report

Keith Chadwick

Grid & Cloud Computing Department Head

Fermilab

Work supported by the U.S. Department of Energy under contract No. DE-AC02-07CH11359

Fermilab

Acknowledgements

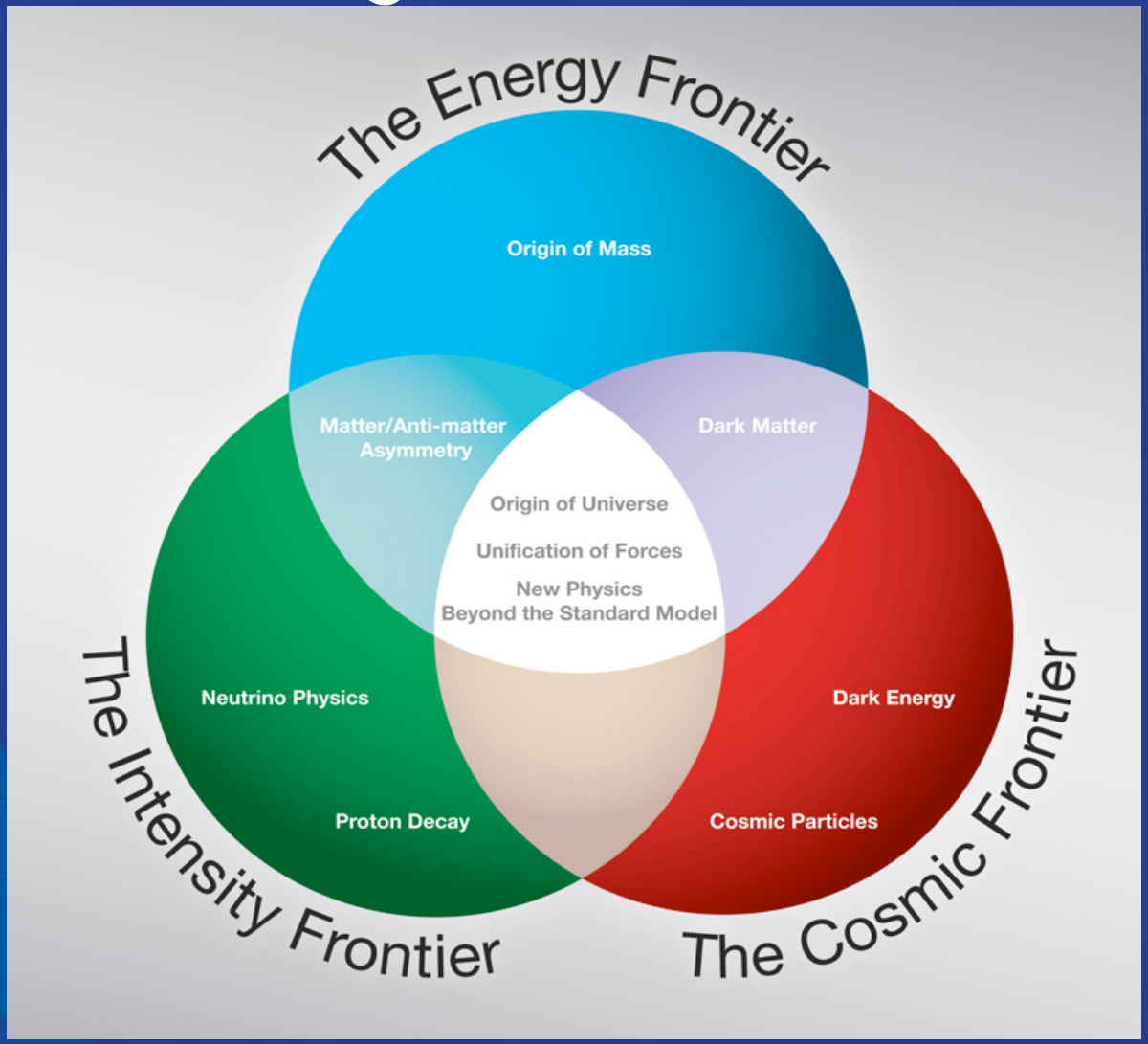
- I am reporting about the work that is being done by the many talented, hardworking and dedicated people at Fermilab.
- The vast majority of the information in this presentation has been shamelessly copied from the Fermilab Computing Division's contributions to the Feb 2011 DOE Scientific Computing.
- Please credit those individuals where credit is due and blame me for any mistakes or misunderstandings. ☺

Outline

- Fermilab Program
  - The Energy Frontier
  - The Intensity Frontier
  - The Cosmic Frontier
- Fermilab Computing Infrastructure
  - FermiGrid
  - Storage
  - HPC
  - Computing Facilities
  - Significant Infrastructure Events (since the Fall 2010 HEPiX)
- Windows
- Scientific Linux
- FermiCloud
- ITIL
- Structure
- Summary

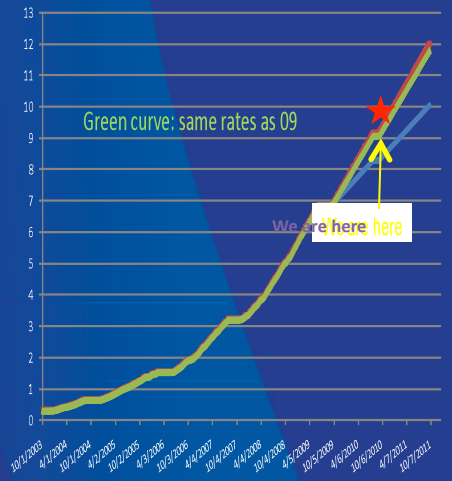
Fermilab Program

Fermilab Program - Three Frontiers



Fermilab – The Energy Frontier

Tevatron **LHC**	**LHC**	LHC Upgrades ILC??	LHC ILC, CLIC or Muon Collider
Now	2013	2016	2019	2022



Tevatron RunII (CDF and D0)

- Computing has been key to the enormous scientific output from these experiments and has grown in importance during the 25 years of the Tevatron and will continue throughout the LHC era.
  - The availability of computing resources (dedicated and shared) has allowed quick and timely processing of the data and Monte Carlo simulations and supported 100's of Tevatron users to complete their complex analyses.
- CDF and D0 physics analysis will continue well past the end of Tevatron operations.
  - We have a commitment to maintain the Tevatron data for 10 years.
  - This means we need to examine data preservation techniques so that the stored dataset remain useable.

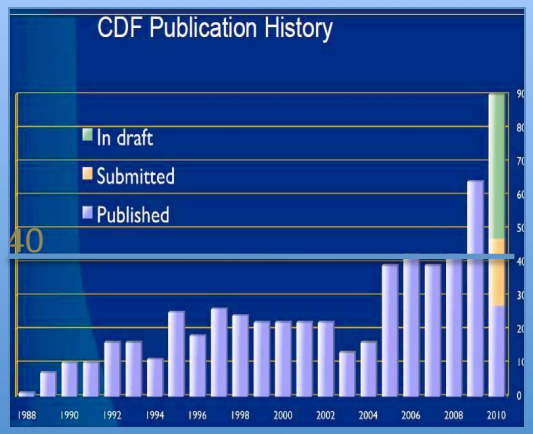
Computing required for Accelerator-based experiments

- Reconstruct Raw data -> physics summary data
- Analyze reconstructed data
- Create MC data needed for analysis
- Reprocess data, skim and regroup processed data
- Store data and distribute to collaborators worldwide
- Software tools & services and expert help at times (e.g. detector simulation, generators, code performance)
- Long-term curation of data and preservation of analysis capabilities after experiment ends

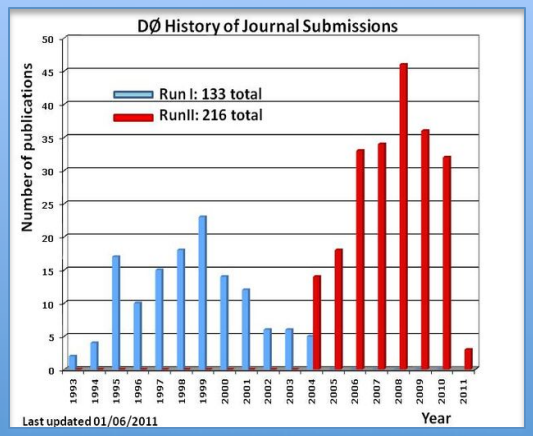
Run II Computing Plans

- Provision for continuation of some Run II production processing and Monte-Carlo production capability
  - Reprocessing efforts in 2011/2012 aimed at the Higgs
  - Monte Carlo production at the current rate through mid-2013
- Provision for Analysis computing capability for at least 5 years
  - Modernization of code and hardware needs to be done now to reduce support later
  - Push for 2012 conferences for many results –no large drop in computing requirements through this period
- Curation of the data: > 10 years
  - Some support for continuing analyses
- Continued support for
  - Code management and science software infrastructure
  - Data handling for production (+MC) and Analysis Operations

Tevatron – looking ahead



CDF Publications each year



D0 Publications each year

The collaborations expect the publication rate to remain stable for several years.

Analysis activity:

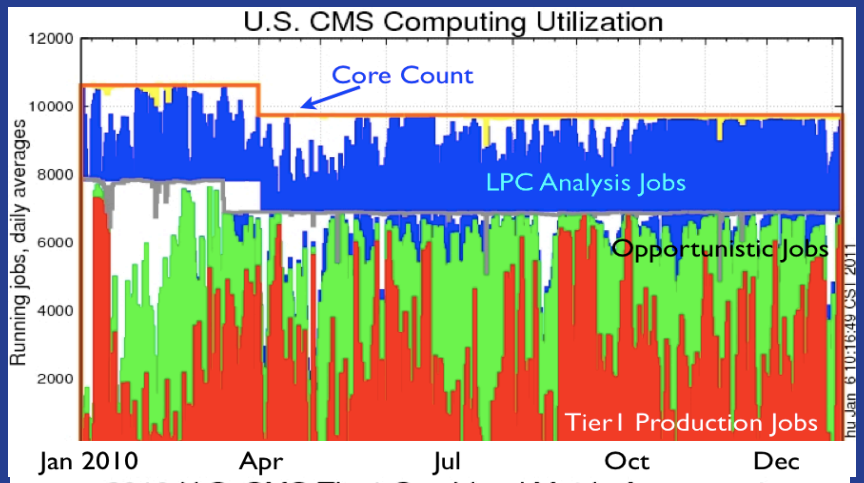
- Expect > 100 (students+ postdocs) actively doing analysis in each experiment through 2012.
- Expect this number to be much smaller in 2015 though data analysis will still be on-going.

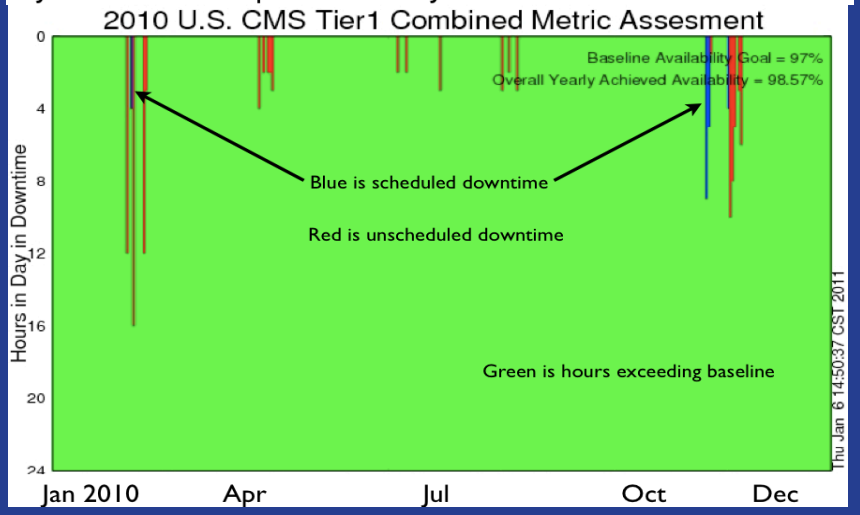
CMS Offline and Computing

- Fermilab is the host institution for U.S. CMS Research Program
- Fermilab is a hub for CMS Offline and Computing
  - 41 Senior Scientists (6 in CD)
  - 14 Research Associates
  - Leadership roles in many areas in CMS Offline and Computing: Frameworks, Simulations, Data Quality Monitoring, Workload Management and Data Management, Data Operations, Integration and User Support.
- Fermilab hosts the U.S. CMS Tier-1 facility.
  - Largest of the 7 CMS Tier-1 Centers
- Fermilab Remote Operations Center allows US physicists to participate in monitoring shifts for CMS.

CMS Tier 1 at Fermilab

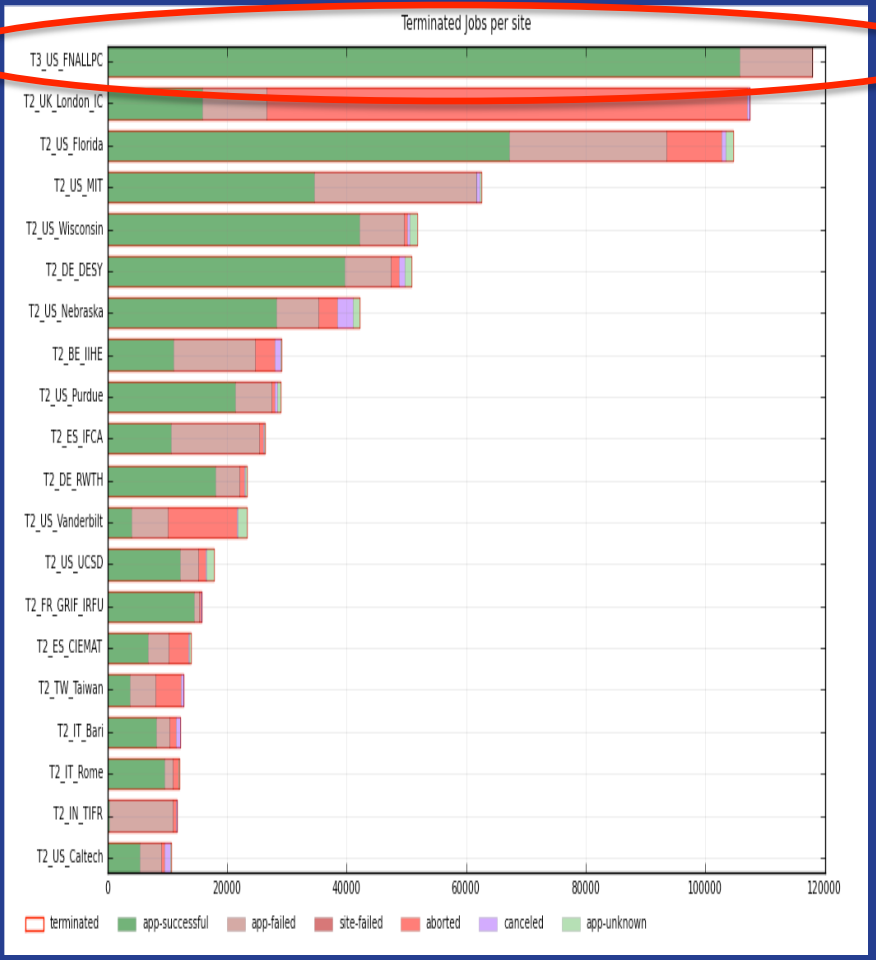
- The CMS Tier-1 facility at Fermilab and the experienced team who operate it enable CMS to reprocess data quickly and to distribute the data reliably to the user community around the world.
- > 2 PB Data transferred to the Fermilab Tier-1 in 2010
-



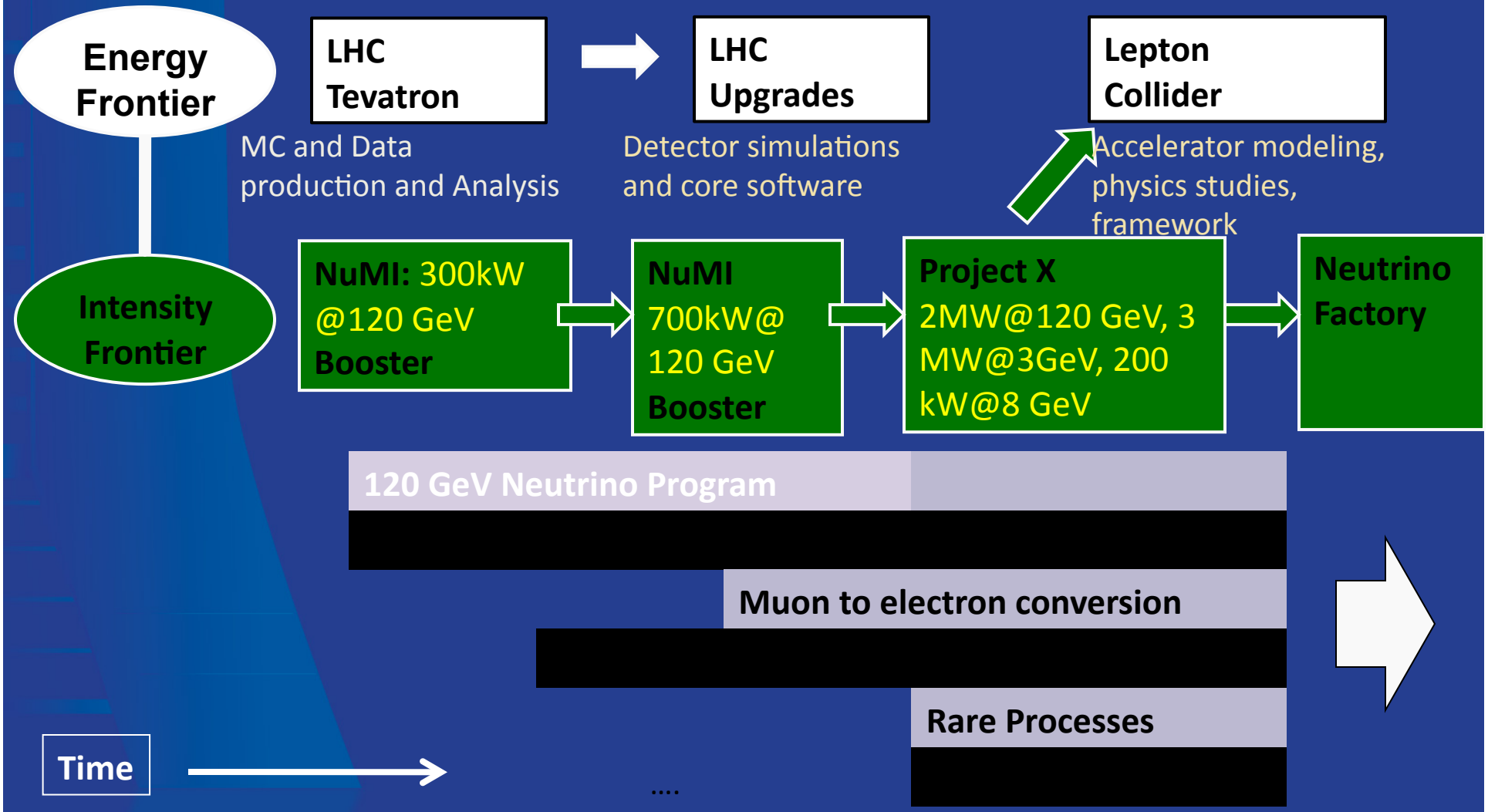


CMS Physics Analysis at Fermilab

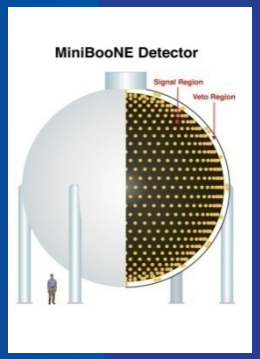
- U.S.CMS analysis facility at Fermilab has the highest number of CMS analysis jobs - typically ~120k-200k jobs/week.
- The analysis job success rate is also very high - typically >85% success rate!
- Fermilab is an excellent place to work on CMS physics analysis.

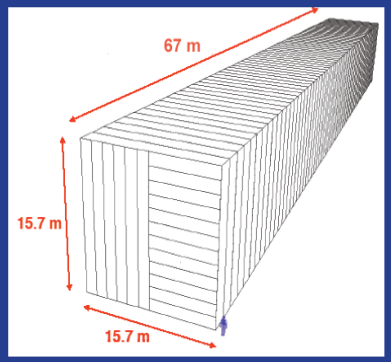


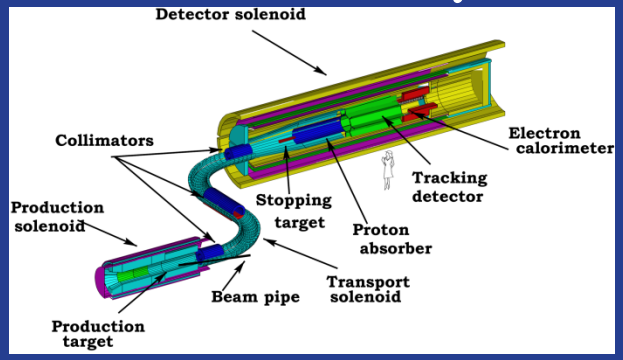
Preparing for the future

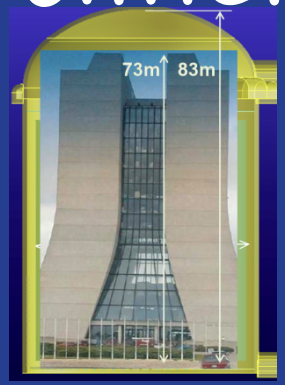


Fermilab – The Intensity Frontier

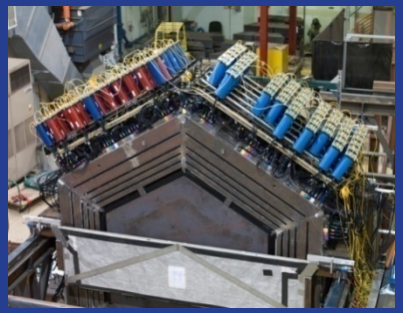


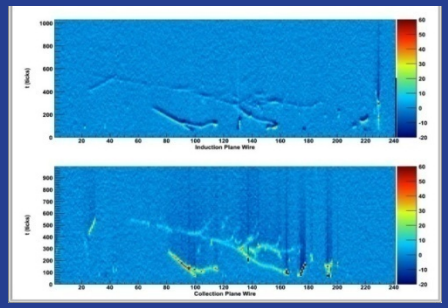


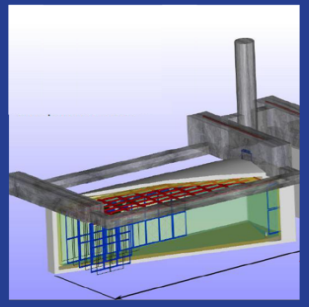




Now	2013	2016	2019	2022
MINOS MiniBooNE MINERvA SeaQuest	NOvA MicroBooNE Muon g-2 SeaQuest	LBNE Mu2e	Project X+LBNE μ, K, nuclear, … ν Factory ?	







Fermilab Program Planning

Draft 2010-13 Fermilab Accelerator Experiments' Run Schedule

Typically Revised Annually - This Version from June, 2010

Calendar Year		2010	2011	2012	2013
Tevatron Collider		CDF & DZero	CDF & DZero	OPEN	OPEN
Neutrino Program	B	MiniBooNE	MiniBooNE #		MicroBooNE
	MI	MINOS	MINOS		OPEN
		MINERvA	MINERvA		MINERvA
		ArgoNeuT			
			NOvA		NOvA
SY 120	MT	Test Beam	Test Beam		Test Beam
	MC	OPEN	OPEN		OPEN
	NM4	E-906/SeaQuest	E-906/SeaQuest		E-906/SeaQuest

This draft schedule is meant to show the general outline of the Fermilab accelerator experiments schedule, including unscheduled periods.

Major components of the schedule include shutdowns:

In Calendar 2010, a 4 week shutdown for maintenance scheduled to begin July 19.

In Calendar 2011, no shutdown for maintenance is shown.

A 2012-3 11-month shutdown is shown to upgrade the proton source and change the NuMI beam to the Medium Energy (ME) config.

Duration of the MiniBooNE run will depend on preparations for MicroBooNE.

RUN/DATA

STARTUP/COMMISSIONING

INSTALLATION

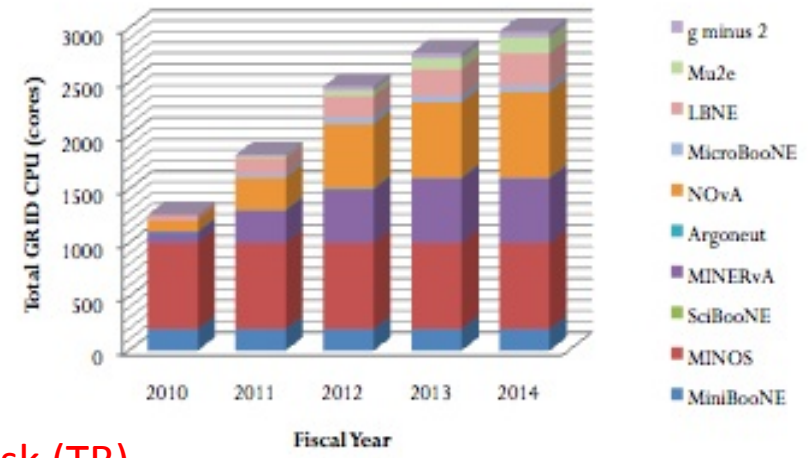
M&D (SHUTDOWN)

15-Jun-10

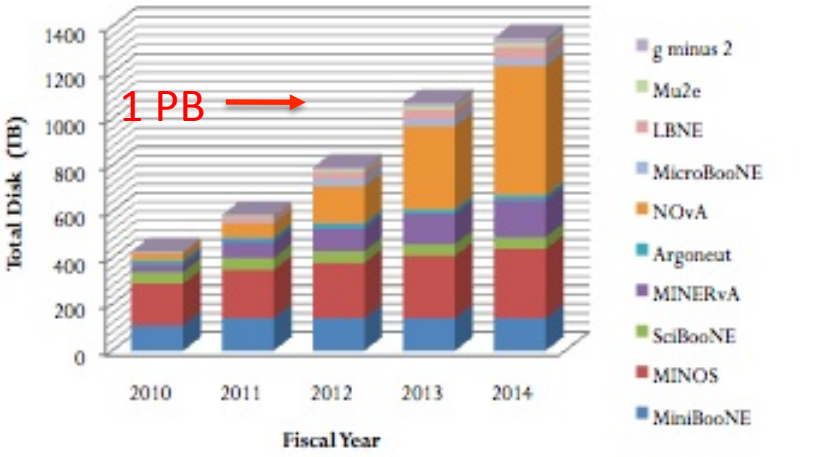
Intensity Frontier program needs

- Many experiments in many different phases of development/ operations.
- Projections for Central Disk Storage and GRID Farm usage are from requests from the experiments. ➔
- We meet monthly with the experiments to plan, review and discuss progress.
- Application support for databases and simulations (GENIE, GEANT4) has been requested.

CPU (cores)



Disk (TB)

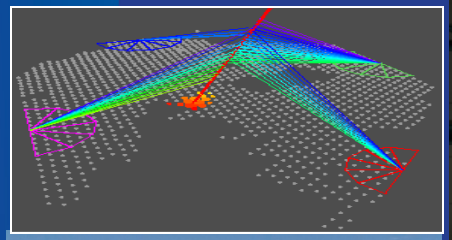


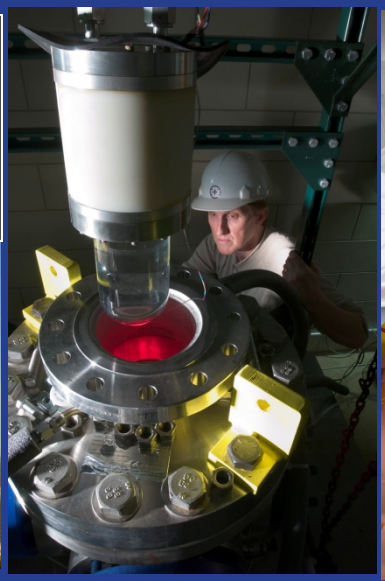
Future Fermilab accelerator based program

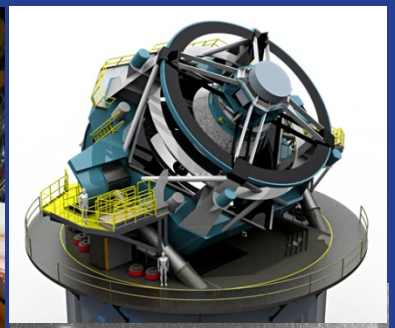
- Intensity Frontier
  - Mu2e
  - LBNE (Water Cherenkov and Liquid Argon)
  - Possible muon g-2 experiment at Fermilab
  - Possible Future Project X experiments: (some of these experiments will be data intensive.)
    - Next generation mu2e, g-2
    - K→ $\pi \nu \nu$ experiments
    - Other neutrino experiments
- Energy Frontier
  - Lepton collider physics studies including R&D for Muon collider
- Accelerator component and beam dynamics modeling
  - Lepton colliders, neutrino factory, Project X

Fermilab – The Cosmic Frontier

DM: Dark Matter DE: Dark Energy








DM: ~10 kg COUPP, CDMS DE: SDSS P. Auger	DM: ~100 kg DE: DES P. Auger Holometer?	DM: ~1 ton DE: LSST BigBOSS	DE: LSST ??

Now 2013 2016 2019 2022

Astrophysics at Fermilab

Fermilab was a member of SDSS I and SDSS II

- Fermilab hosts 91 TB primary archive; 32,000 queries + 1 TB distributed per day to the general public.
- Through 2010, 3400 refereed papers with SDSS in title or abstract, >35,000 citations. Over half of publications are from outside the collaboration.
- Highest impact observatory 2002, 2003, 2006

DES follows this; Fermilab will host the secondary archive

Fermilab recently joined LSST

Computing for Astrophysics

- Sloan Digital Sky Survey
  - Data taking completed 2008
  - Commitment: Operate data archive through 2013
  - Archive: 22 nodes + 65 TB disk + Enstore tape
- Dark Energy Survey
  - Operations 2011-2013+
  - Commitment:
    - Simulations – 200 slots + 3.5 TB
    - Copy of DES archive - 420 TB
    - Analysis Computing (~200 cores needed)
    - Local Data Reprocessing (800 slots)
- Pierre Auger Observatory
  - Operations: 2011-2013+
  - Commitment:
    - Data Mirror (5 TB/yr)
    - Host Calibration database
    - Simulations – 100 slots
- Cold Dark Matter Search (SuperCDMS)
  - Operations through 2013+
  - Commitment:
    - Data processing and storage (800 slots, 64 TB/yr)
- Chicagoland Observatory … (COUPP)
  - Operations through 2013+
  - Commitment:
    - Data processing and storage (100 slots, 2 TB/yr)

Experimental Astrophysics -Future

- JDEM/WFIRST:
  - Steve Kent - Deputy Project Scientist
  - Erik Gottschalk - Science Operations Center Lead
  - JDEM effort is transitioning to other programs
- LSST:
  - Fermilab has just recently joined
  - No commitments yet. Ideas:
    - a) Data Access Center (DAC) - Science Analysis
    - b) Dark Energy Science Center
  - The Grid & Cloud Computing Department ran a pilot project to demonstrate the automated production of LSST detector simulation data using Open Science Grid resources.

DES Analysis Computing at Fermilab

- Fermilab plans to host a copy of the DES Science Archive. This consists of two pieces
  - A copy of the Science database
  - A copy of the relevant image data on disk and tape
- This copy serves a number of different roles
  - Acts as a backup for the primary NCSA archive, enabling collaboration access to the data when the primary is unavailable
  - Handles queries by the collaboration, thus supplementing the resources at NCSA
  - Enables the Fermilab scientists to effectively exploit the DES data for science analysis
- To support the science analysis of the Fermilab Scientists, DES will need a modest amount of computing (of order 24 nodes). This is similar to what was supported for the SDSS project.

CDMS/Minos/Minerva

- A fire in the Soudan Mine access shaft was detected on Thursday 17-Mar-2011.
- Specially trained first responders dealt with the fire and it was declared "99% out" on Sunday 20-Mar-2011.
  - The fire appears to have started in the timbers that line the access shaft between the 23rd and 25th levels of the mine.
- Water pumping stations on mine levels 12 and 22 were recommissioned on Saturday 26-Mar-2011
- First inspections of the underground detectors on level 27 were on Wednesday 30-Mar-2011.
- Restoration of the infrastructure is underway.
  - Limited power has been restored to the experimental halls.
  - Limited networking has been restored (likely will have to replace the fiber optic cable from the surface to the detector level).
  - Cleanup of the fire fighting foam residue and drying out is underway.
- Unclear when data taking will resume.

http://www.fnal.gov/pub/today/archive_2011/soudan_mine_fire.html

Fermilab Computing Infrastructure

FermiGrid - CPU resources for the user community

User Login & Job Submission

Open Science Grid

TeraGrid
WLCG
NDGF

FermiGrid Authentication /Authorization Services

FermiGrid Site Gateway

FermiGrid Infrastructure Services

FermiGrid Monitoring/ Accounting Services

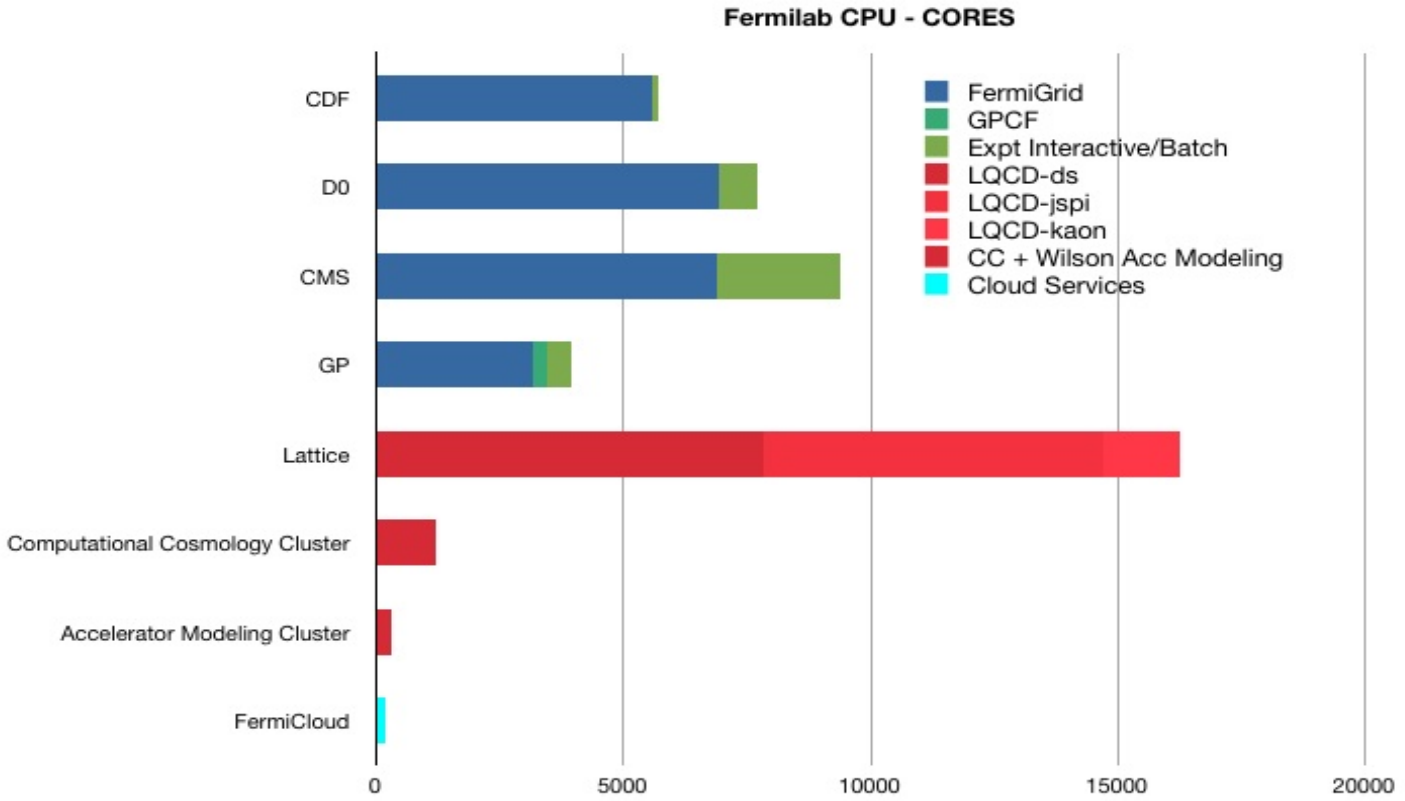
GP	D0	CDF	CMS
3284 slots	6916 slots	5600 slots	7485 slots

Measured FermiGrid Availabilty(*)

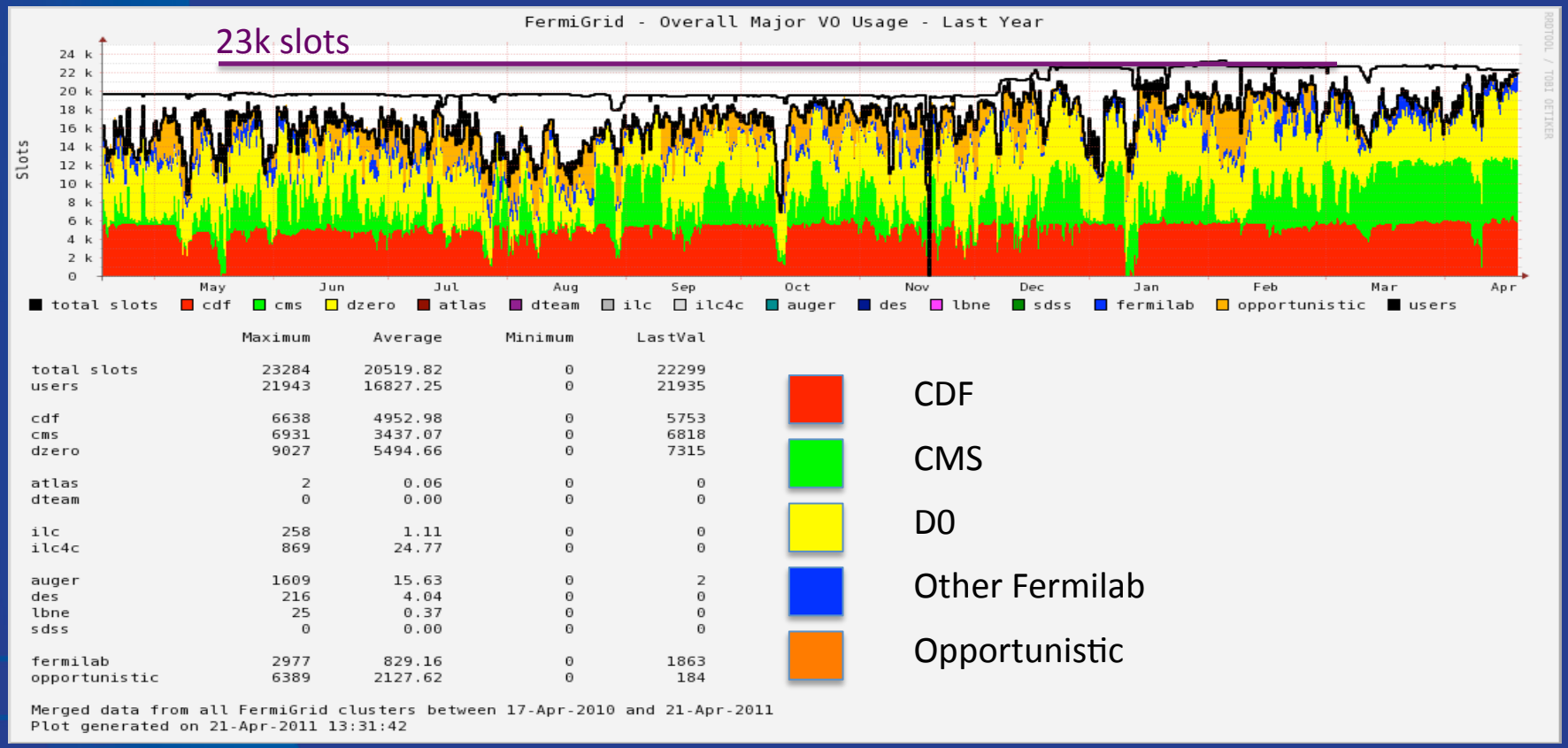
Service	Availability	Downtime
VOMS-HA	100%	0m
GUMS-HA	100%	0m
SAZ-HA (gatekeeper)	100%	0m
Squid-HA	100%	0m
MyProxy-HA	99.943%	299.0m
ReSS-HA	99.959%	215.4m
Gratia-HP	99.955%	233.3m
Database-HA	99.963%	192.6m

* = Excluding building or network failures

FNAL CPU - core count for science



FermiGrid – Past Year Slot Usage



	Maximum	Average	Minimum	LastVal
total slots	23284	20519.82	0	22299
users	21943	16827.25	0	21935
cdf	6638	4952.98	0	5753
cms	6931	3437.07	0	6818
dzero	9027	5494.66	0	7315
atlas	2	0.06	0	0
dteam	0	0.00	0	0
ilc	258	1.11	0	0
ilc4c	869	24.77	0	0
auger	1609	15.63	0	2
des	216	4.04	0	0
lbne	25	0.37	0	0
sdss	0	0.00	0	0
fermilab	2977	829.16	0	1863
opportunistic	6389	2127.62	0	184

Merged data from all FermiGrid clusters between 17-Apr-2010 and 21-Apr-2011
Plot generated on 21-Apr-2011 13:31:42

- CDF
- CMS
- D0
- Other Fermilab
- Opportunistic

http://fermigrid.fnal.gov

HEP – a Data Intensive Science

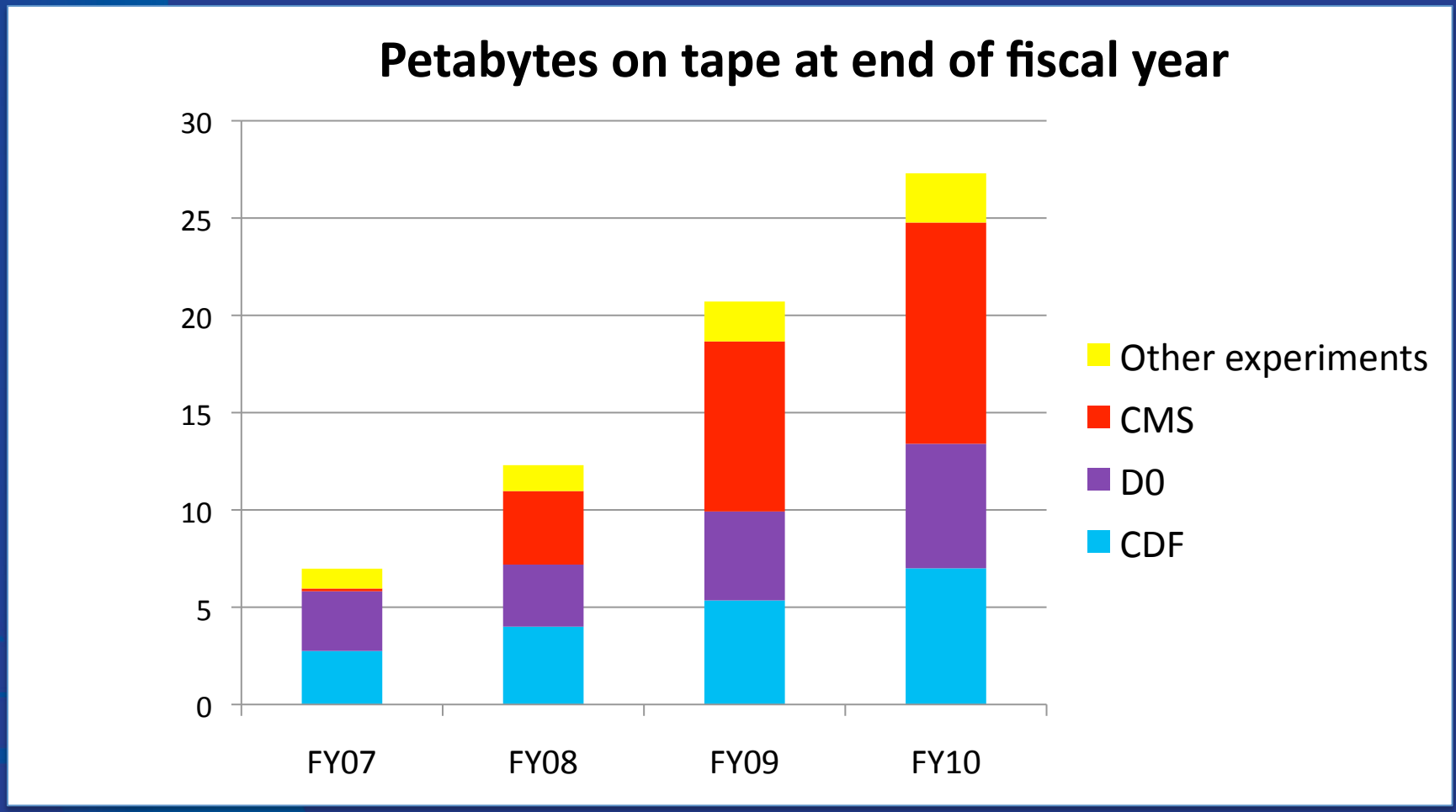
- Large datasets are now a hallmark of nearly all our scientific programs – including those involving theory and simulations.

- Fermilab Computing Division provides central, shared data movement and storage services for ALL lab scientific program
  - Some resources are dedicated (e.g. disks)
  - Some resources are paid for by specific lab scientific programs and they have "guaranteed use" of (e.g. number of tape drives of a specific type)
  - Some resources are shared across a broad set of scientific programs (e.g. public disk cache)

Data Movement and Storage Services

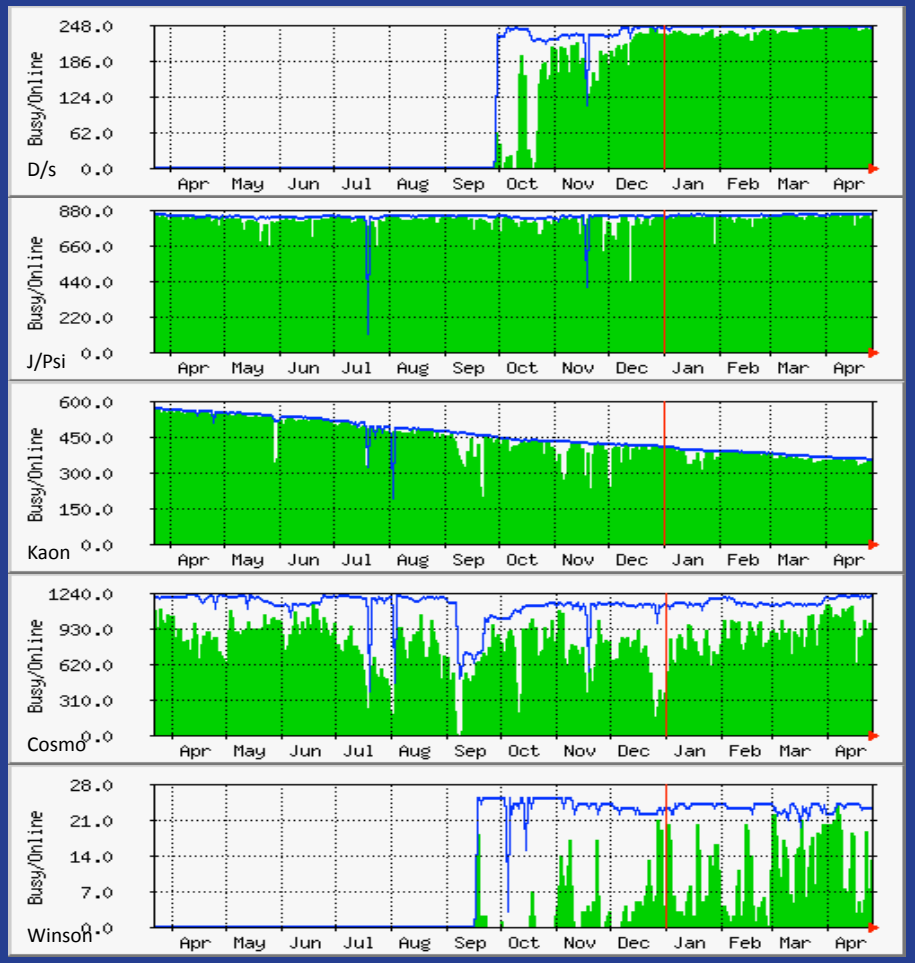
- Enstore: tape Libraries and storage system – for primary and archival data storage
  - Developed at Fermilab (+ part in common with dcache)
  - Also used at PIC (Spain)
  - Continuous migration to new tape technologies
  - Must scale to requests for many thousands of files simultaneously
  - 26 Petabytes on Tape
- dCache: tape-backed and non tape-backed disk cache for data processing (Fermilab-DESY-NGDF-others collaboration)
  - Scales for current usage – concerns for future demands
  - Investigating Lustre file system as future dCache replacement
    - broader-than-HEP open source product
- Space has been identified for two additional SL8500 Robots in FCC.
  - One of these robots has been ordered and delivery is expected in May 2011.

Data Storage at Fermilab - Tape



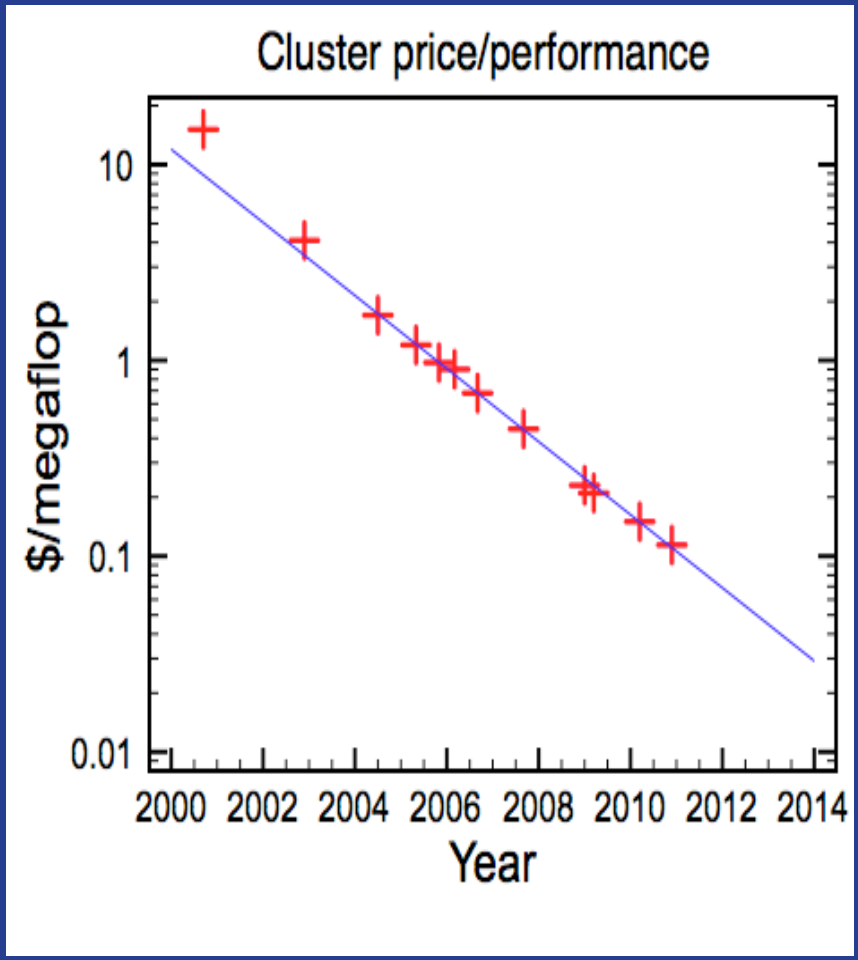
HPC

- Fermilab operates 5 HPC clusters:
  - The "D/s", "J/Psi" and "Kaon" clusters for Lattice QCD calculations,
  - The "Cosmology" cluster for cosmological calculations,
  - The "Wilson" cluster for accelerator modeling.
- In addition to the above production clusters we also have a small GPU cluster (not shown).



HPC – Price/Performance

- The price/performance of clusters installed at Fermilab for lattice QCD has fallen steadily over the last six years, with a halving time of around 1.5 years, as shown by the solid line in the graph at right.
- Product roadmaps provided by vendors of system components make clear that this trend is likely to continue for the next several years.

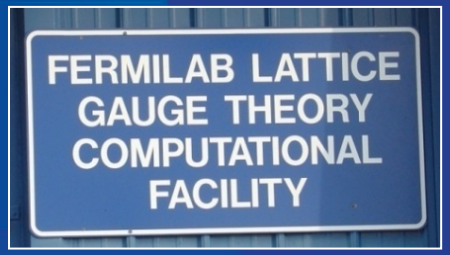


COMPUTING FACILITIES

Fermilab Computing Facilities

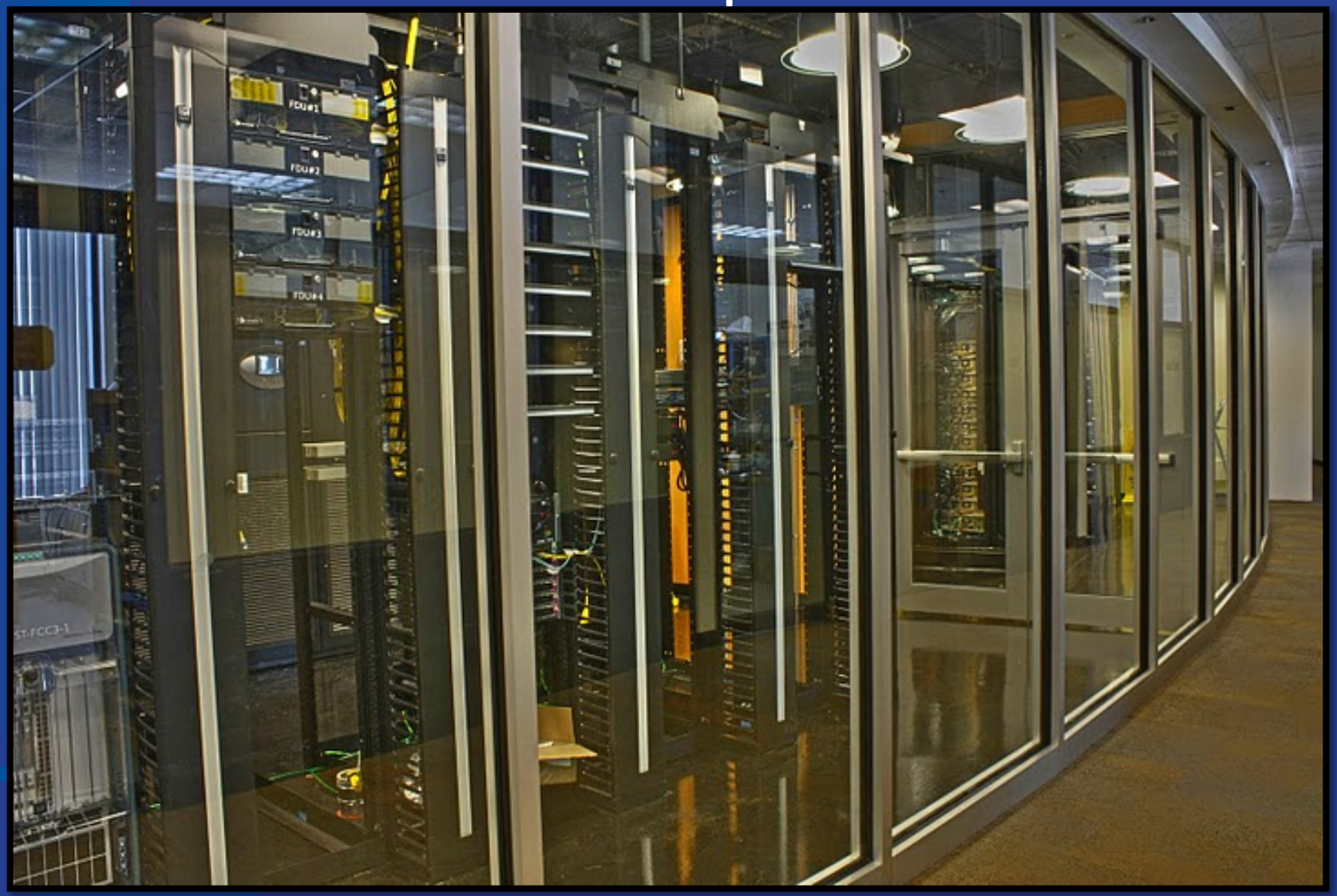
- Feynman Computing Center (FCC)
  - •High Availability Services
  - •Networking, Computer Security, BSS, Email, Web, Databases, Core Site Services
  - •Tape Robotic Storage
  - •UPS & Standby Power Generation
  - •ARRA project: upgrade cooling and add HA computing room - completed

- •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

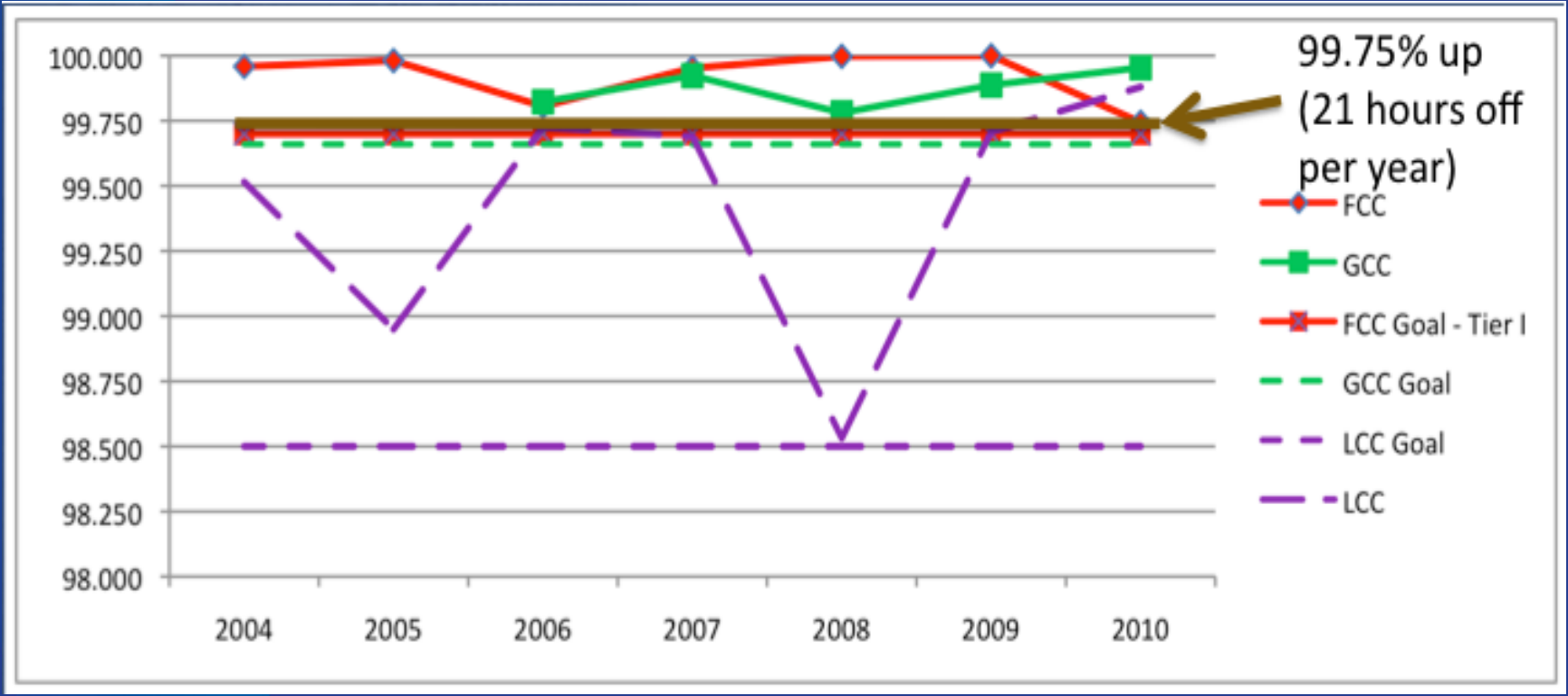


- •Lattice Computing Center (LCC)
  - •High Performance Computing (HPC)
  - •Accelerator Simulation, Cosmology nodes
  - •No UPS

New FCC3 Computer Room

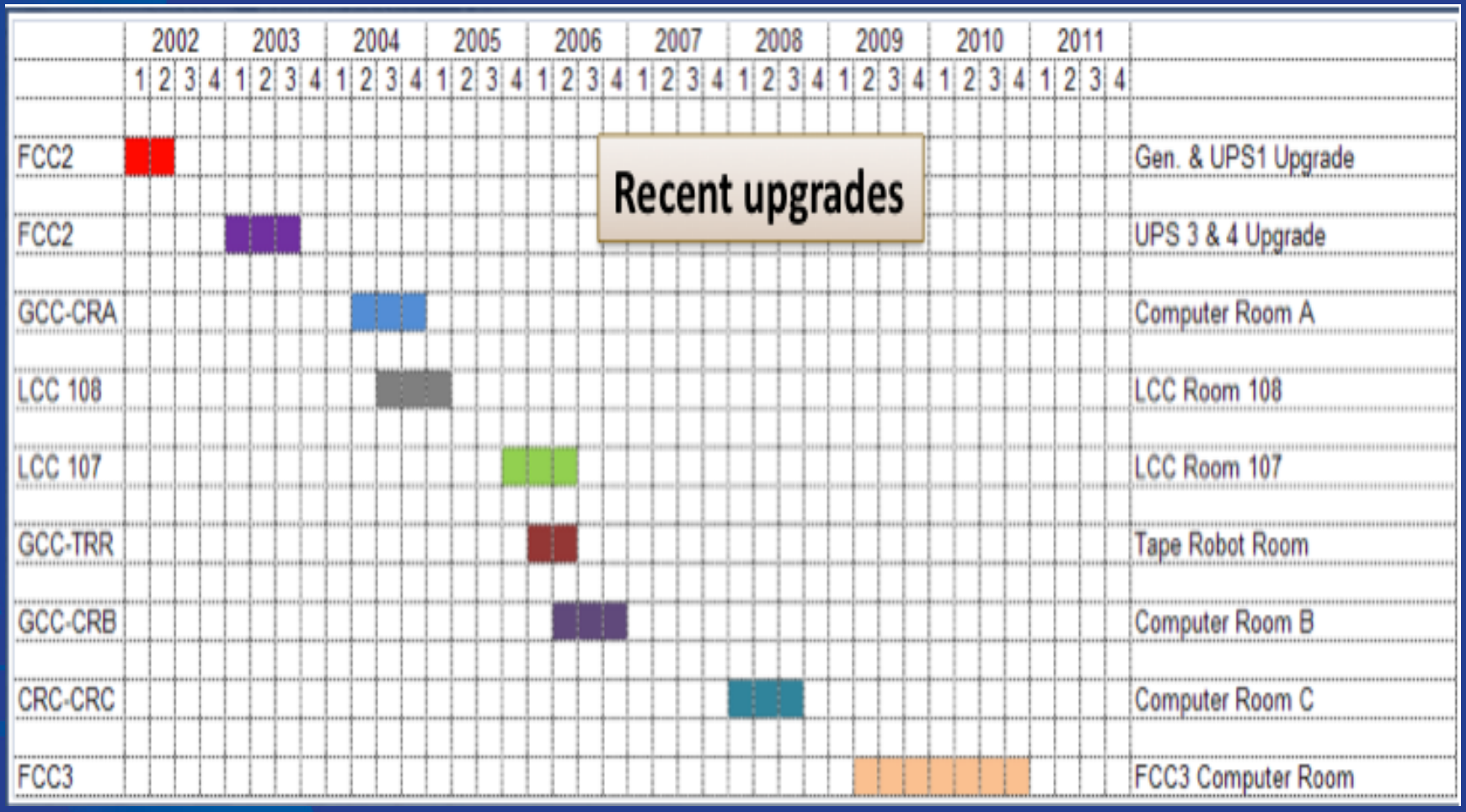


Computer Room Availability 2004-2010



- >99.75% in FCC and GCC (server and CPU rooms) over many years
- Better than the expectation based on the building infrastructure
- Excellent performance

Facility Upgrades

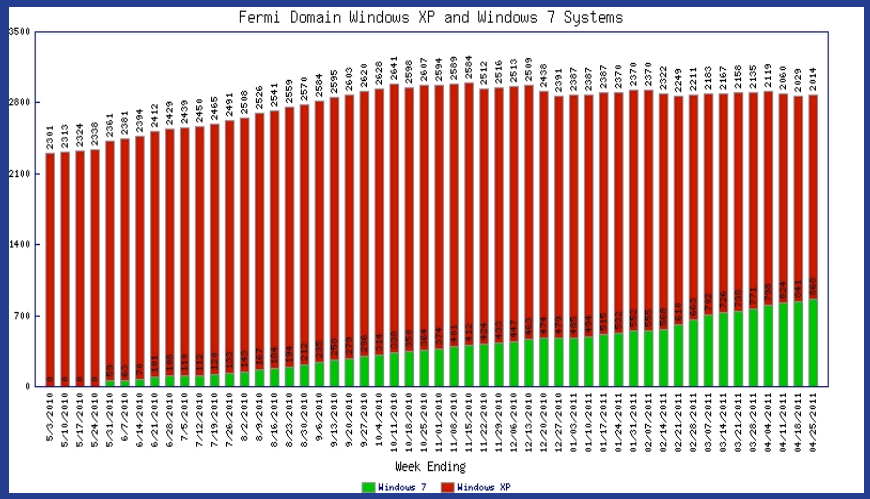


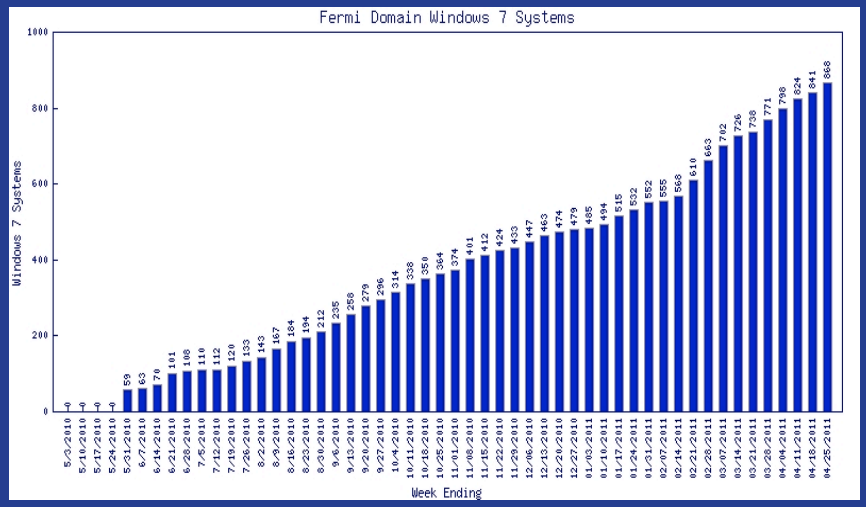
Significant Infrastructure Events

- ☹ November 2010:
  - Overheating distribution breaker found during infrared survey of electrical distribution panels required ~8 hour shutdown of FCC computer rooms to replace.
  - During the startup of the systems following the distribution breaker replacement, the upstream 1400A breaker tripped.
  - This is the same 1400A breaker that tripped twice in February 2010.
  - The decision was made to replace the 1400A breaker - ~16 hour shutdown.
- ☺ January 2011:
  - FCC3 Computer Room commissioned.
- ☹ February 2011:
  - Spanning tree issues caused at least two site wide network outages (each lasting 15-30 minutes).
  - Failing supervisor module in 6509 repeatedly caused loss of connectivity for several production services to the remainder of the site network over a five day period.
  - Various mitigations were attempted, but eventually a "whole switch" replacement was performed over an ~8 hour period.

Microsoft Windows

Windows Version	Inventory
Windows NT 4.0	3
Windows 2000 Professional	9
Windows 2000 Server	2
Windows Server 2003	195
Windows Server 2008	94
Windows XP Professional	2046
Windows XP Tablet PC Edition	7
Windows Vista Business	22
Windows Vista Enterprise	18
Windows Vista Ultimate	1
Windows 7 Enterprise	897
Windows 7 Professional	1
Total	3295





Scientific Linux

Scientific Linux 3 was deprecated on 10-Oct-2010!

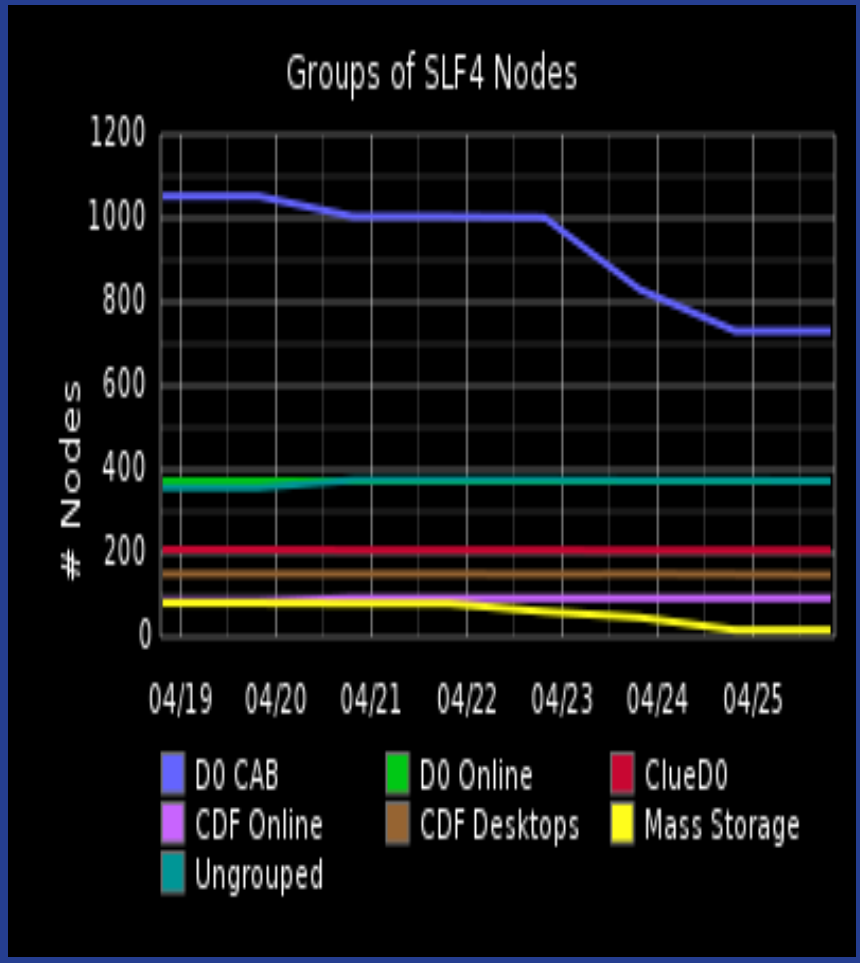
- We are down to four SLF3 systems to migrate.

Scientific Linux 4 will be deprecated on 12-Feb-2012!

- Migration off SL(F)4 ➔

Scientific Linux 5 + 6

- See Troy's presentation later this week.



Apple / Android

- The Computing Division is extending it's support for Apple desktops and laptops to include support for Apple iOS based devices (iPhone/iPod/iPad).
- We will be also be deploying support for Android devices also...

Progress on FermiCloud

- FermiCloud has largely completed the Phase 1 Project Goals:
  - The selected Cloud Computing management framework was OpenNebula.
  - The storage performance investigation work is still ongoing.
- We are currently working on Phase 2 of the FermiCloud project:
  - We have developed and deployed an X509 AuthZ/AuthN framework for OpenNebula.
  - We have deployed a small number of "low impact" production services.
  - We are working on extending FermiCloud to host other production services.
  - Other items in Phase2 include Accounting, Infiniband/HPC, Idle VM detection and backfill.
- We are beginning the planning for Phase 3 of the FermiCloud project:
  - This will likely include work to support FermiCloud-HA (High Availability).
  - We will incorporate the "lessons learned" from the FermiGrid-HA2 deployment.

ITIL

- Fermilab has recently made the decision to replace remedy with Service-Now.
  - The target go-live date is 1-Aug-2011.
- Anurag Chadha joined the Office of Enterprise Architecture and Configuration Management in April 2011 (replacement for Don Petravick who left in June 2010).

Structure

- **In January 2011, the Fermilab Directorate announced that** the Computing Division will soon become the Computing Sector:
  - http://www.fnal.gov/pub/today/archive_2011/Organization-Fermilab-2011.pdf
- The sector will be lead by the Laboratory CIO (Vicky White).
- Within the sector, there will be two divisions instead of four quadrants.
- Current division management will begin interviewing internal candidates to head these two divisions in the near future.

Summary

- Fermilab has a strong, effective scientific computing program that serves a very broad range of scientific programs
  - We plan to continue to develop and use common solutions working with the broad community collaboratively
  - Lots of new opportunities for the future
- We have developed well-founded strategies for meeting the challenges of the current and future scientific program:
    - Tevatron Run II → CMS and Intensity Frontier
    - SDSS→ DES→ LSST
    - Scientific Computing