Advanced Computing
FRA Physics Visiting Committee
April 24-25, 2008

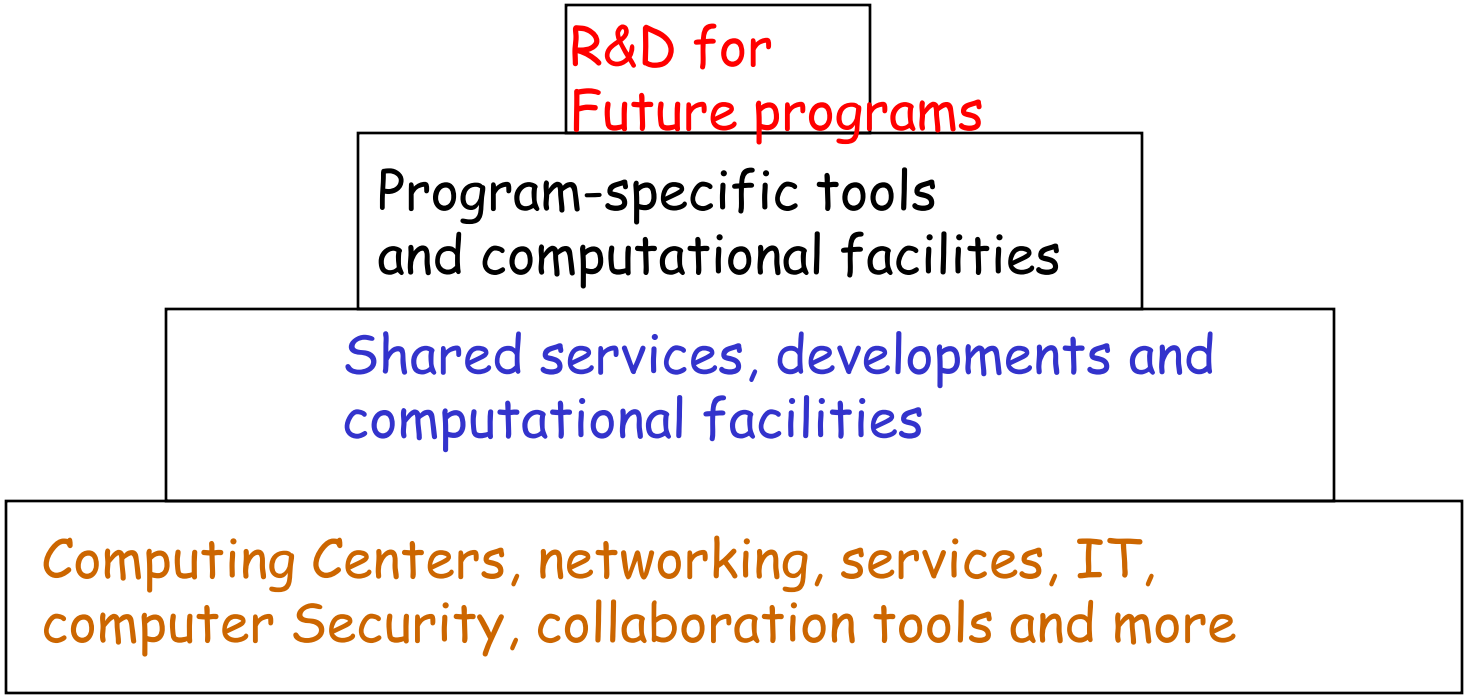
Victoria White
Head, Computing Division

Experiment + Theory + Computing

- The success of our current and future scientific programs relies not only on an adequate amount of computation, storage, networking, Trigger, DAQ, Control systems, databases and reliable IT services (for the program and the lab)
- But on a continuous and evolving Strategy for Advanced Computing because of
  - The scale of computing required for experiment data and theoretical science
  - The unprecedented volumes of data we are dealing with (~9 Petabytes of scientific data in Fermilab storage systems)
  - Global collaborations and global involvement in computing
  - The requirement to keep up with technology and to find innovative, affordable solutions which are sensitive to
    - Space, Power and cooling costs
    - Operational loads on people
    - Performance and Reliability which scales

Advanced Computing Strategy

- Invest in ALL layers of computing judiciously
  - In the places where our field demands different solutions to those common in industry or other labs
  - Share as much as possible - leverage, consolidate, focus on robust solutions

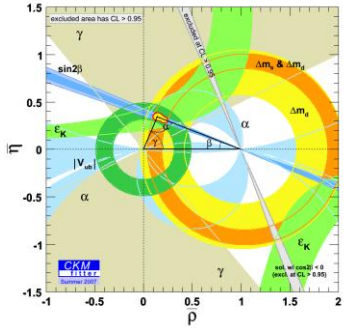


Program Specific Tools and Facilities

- Lattice QCD
- Computational Cosmology
- Computational physics program
  - Accelerator modeling tools
  - Detector modeling and applications
  - Tools for studying and optimizing physics algorithms and computational applications
- CMS Software
  - Fermilab led the redesign of the CMS framework – for efficient access to data
  - Strong contribution to architecture and development of the workflow tools for production processing of data
    - Processed 2B events since beginning of the year using this "ProdAgent" on Tier 0 at CERN and throughout the Worldwide LHC Computing Grid (WLCG)

Lattice QCD computing

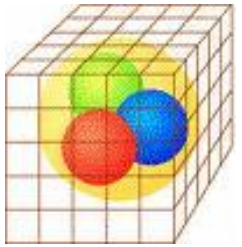
Physics goals:



- Quark flavor physics
  - Essential for B factories, CDF/D0, LHC-b
  - Essential for high precision K and charm physics. Project X
- QCD thermodynamics
  - Relevant to RHIC and early universe
- Hadron spectroscopy and structure
  - Relevant to JLab and pdfs needed at Tevatron and LHC

Lattice QCD physics program is covered in detail by A. Kronfeld.

Simulations:



- 2 TF-yrs to simulate configurations of the gluon field.
- 2 TF-yrs to simulate hadrons in the gluon fields, extract physics results.
- Repeat steps 1 and 2 multiple times to understand the systematic errors.
- Requires tightly coupled highly parallel hardware computing with scalable performance.
- Requires optimized portable codes that can be efficiently run on a wide variety of hardware.

Fermilab QCD and USQCD

- Fermilab builds and operates Lattice QCD facilities as part of the DOE Office of Science's national plan for LQCD.
- USQCD scientific collaboration represents lattice theorists from BNL, Fermilab, JLab and universities.
- Five year (FY05-09) OMB 300 project with a $9.2M budget of which $5.4M is designated for hardware and operations at Fermilab.
  - LQCD project has received a "green" rating (best) in all of its DOE project reviews.
  - Recent USQCD user survey found many favorable comments regarding the operations of the Fermilab LQCD facility.
- Follow-on project to start FY10 has been proposed to DOE.
- DOE SciDAC grants fund software research and development.
- SciDAC investigations at Fermilab:
  - multicore optimizations
  - scientific workflows for LQCD
  - an automated fault monitoring and mitigation system

Office of Science SciDAC Scientific Discovery through Advanced Computing USQCD US Lattice Quantum Chromodynamics

LQCD Clusters

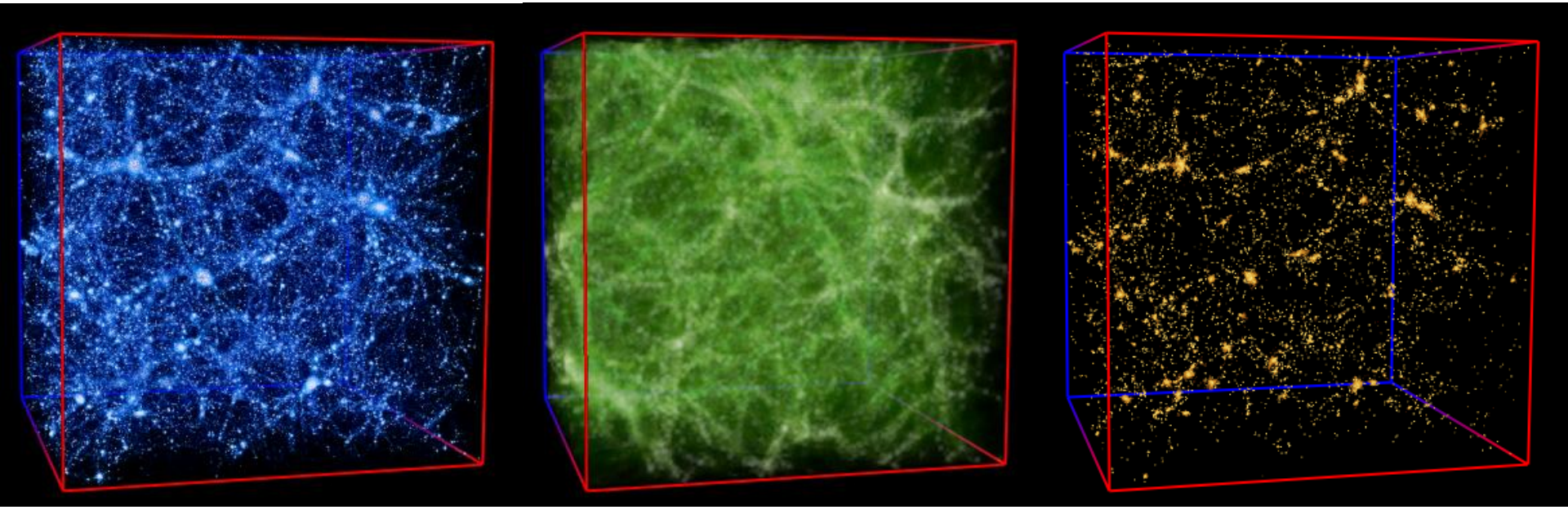
Lattice Computing Center
(at the New Muon Lab)

Pion Cluster

Name	CPU	Cores	Network	Online
QCD	Single 2.8 GHz Pentium 4	127	Myrinet 2000	June 2004 0.15 TFlops
Pion	Single 3.2 GHz Pentium 640	518	Infiniband Single Data Rate	June 2005 / Dec 2005 0.86 TFlops
Kaon	Dual 2.0 GHz Opteron 240	2400	Infiniband Double Data Rate	Oct 2006 2.56 TFlops
J/psi	TBD	TBD	TBD	Jan / Mar 2009 > 6.2 TFlops

>9.6 Tflops/s sustained QCD performance by 2009

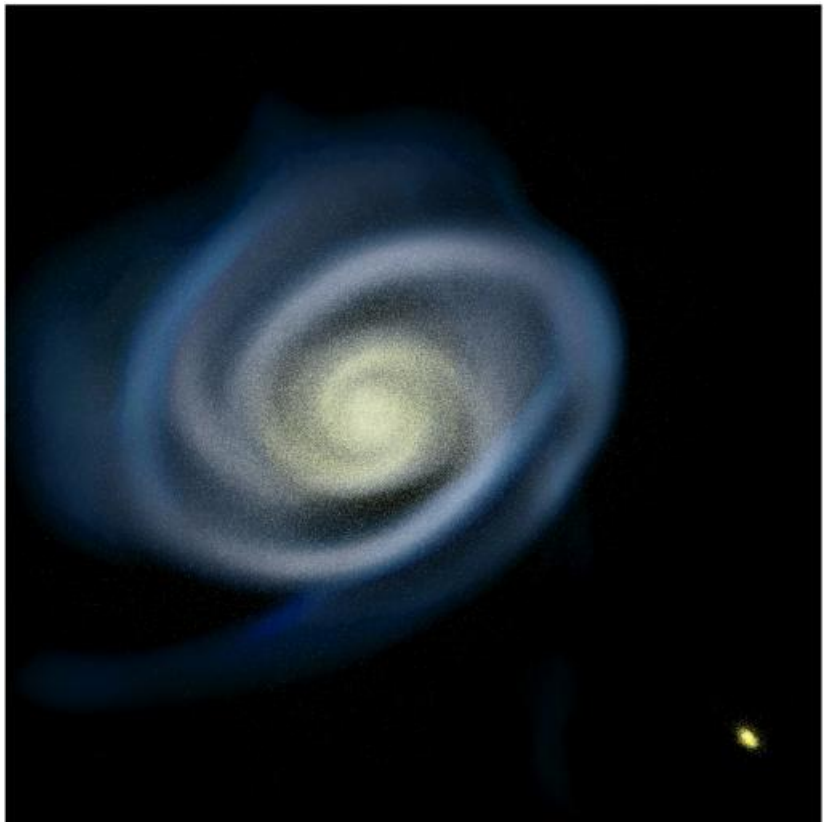
Computational Cosmology: Large Scale Simulations

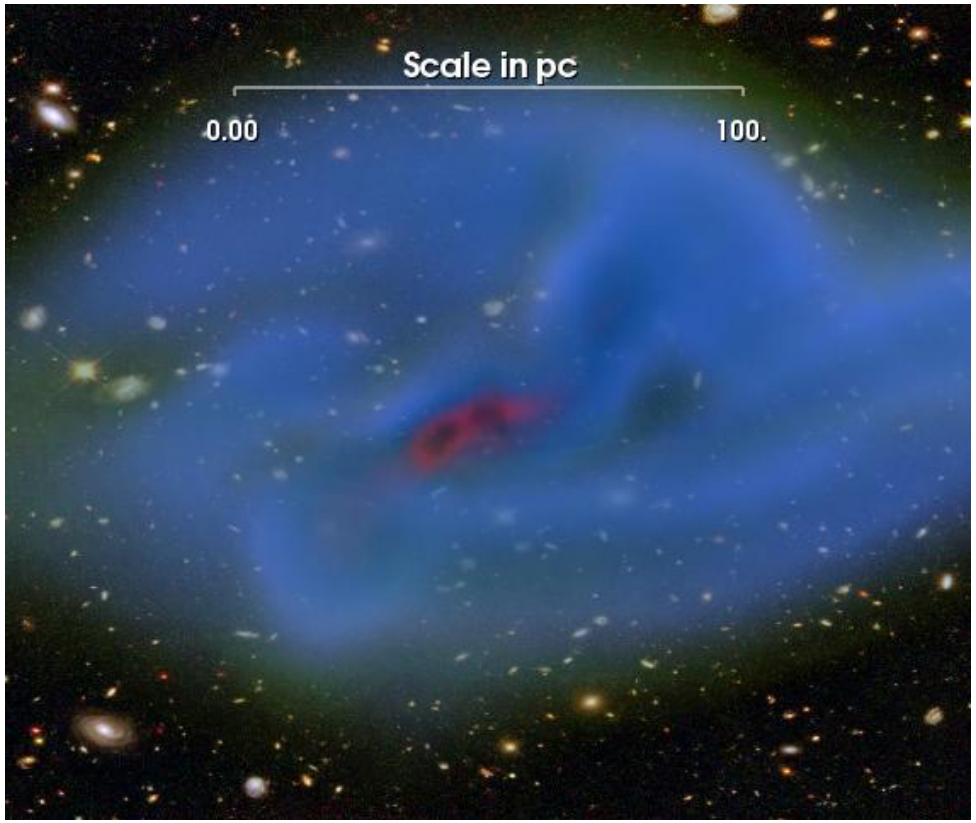


Dark matter Gas Galaxies

Modeling baryonic effects on large scales is important for unbiased calibration of weak lensing signal from future Dark Energy experiments.

Computational Cosmology: Galaxy and black hole formation





Understanding galaxy and black hole formation is the primary goal of modern Astrophysics, and is also important for calibrating Dark Energy experiments.

New Cluster for Computational Cosmology

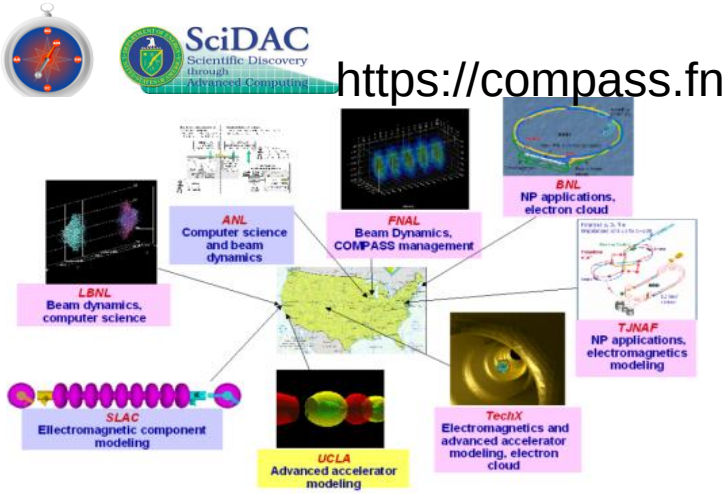
- Leveraged from Lattice QCD cluster work
- Operated by Lattice QCD department
- $$s from FRA, U of C (KICP), Fermilab
- Current Status
  - 69 nodes/560 cores/16 GB per node.
  - About 70% online at any given time.
  - ~25 users (together with KICP).
  - On average running at 70% capacity.
- Goals for this year
  - 125 nodes/1,000 cores.
    - resources would then be on a par with the Institute for Theory and Computation at Harvard

Computational Physics mission

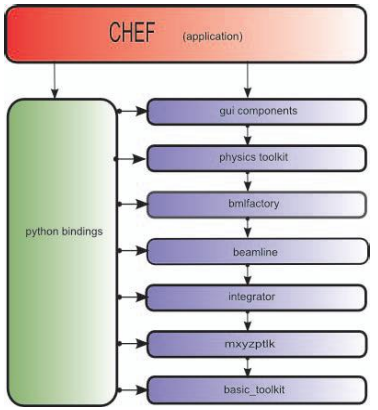
- Develop computational tools
  - Physical system modeling
    - Detector, accelerator, cosmology, HEP theory
  - Data analysis
  - Algorithm and application optimization
- Provide expertise on deployment and utilization of computational tools
  - Both internally and externally developed
  - On local, distributed (Grid), and DOE & NSF "supercomputing" facilities
- Develop or contribute to the development of applications of such tools

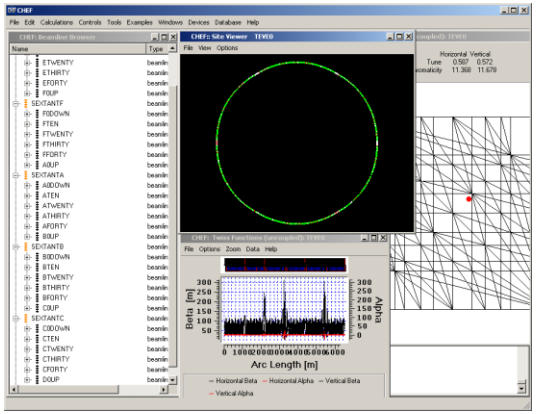
Accelerator modeling tools and applications

- Development of HPC accelerator modeling tools and applications
  - Multi-particle dynamics
  - Management of ComPASS collaboration (a DOE SciDAC funded program)
- Development and support of CHEF
  - general framework for single particle dynamics
- Application focus
  - Run II accelerators, ILC (-> Project-X) design



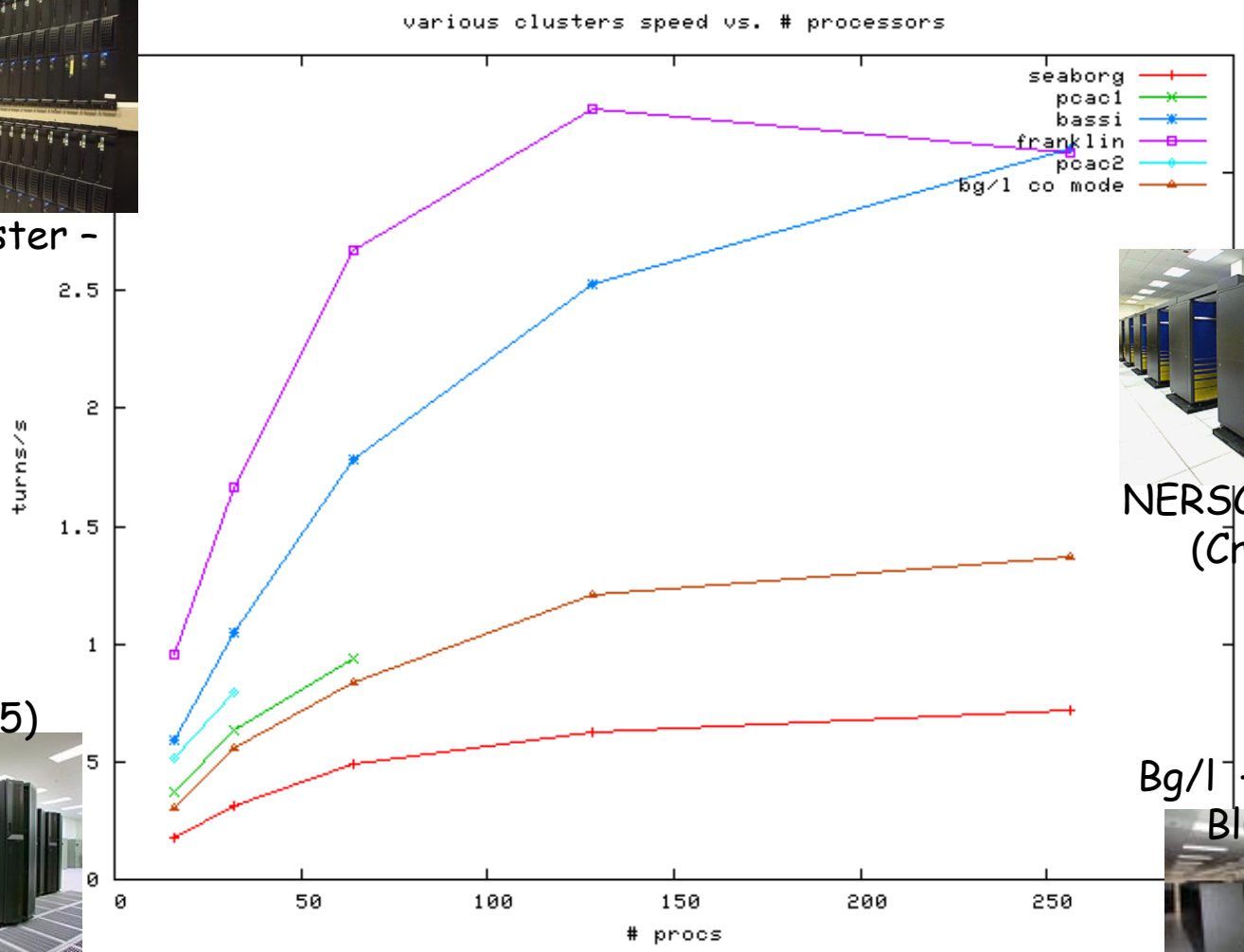
Collaborative Hierarchical Expandable Framework





Performance of accelerator modeling tools on HPC platforms

Fermilab Cluster – pcac2



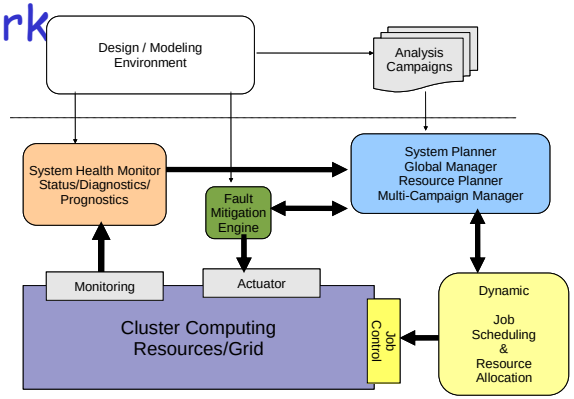
NERSC – franklin (Cray XT4)

NERSC – bassi (ibm power 5)

Bg/l – ANL (IBM Blue Gene L)

Detector modeling tools and performance tools

- Members of the Geant4 collaboration, responsibilities
  - hadronic physics model development and validation, numerical stability, time performance
- Support for USCMS LHC Physics Center (LPC) activities:
  - MC/data overlap tool, GFlash parameterizations, generators/simulation common tasks
- Develop and deploy tools to study and optimize performance of physics algorithms and applications
  - Profiling tools
    - Utilized to study and improve the performance of G4 algorithms in the CMS environment
    - Continue with G4, extend utilization in lQCD framework, potential use in accelerator and cosmology modeling frameworks
  - Computational services reliability framework
    - Used in monitoring/optimizing lQCD cluster
      - Study and analyze failures
      - Predict potential failures
      - Provide mitigation actions



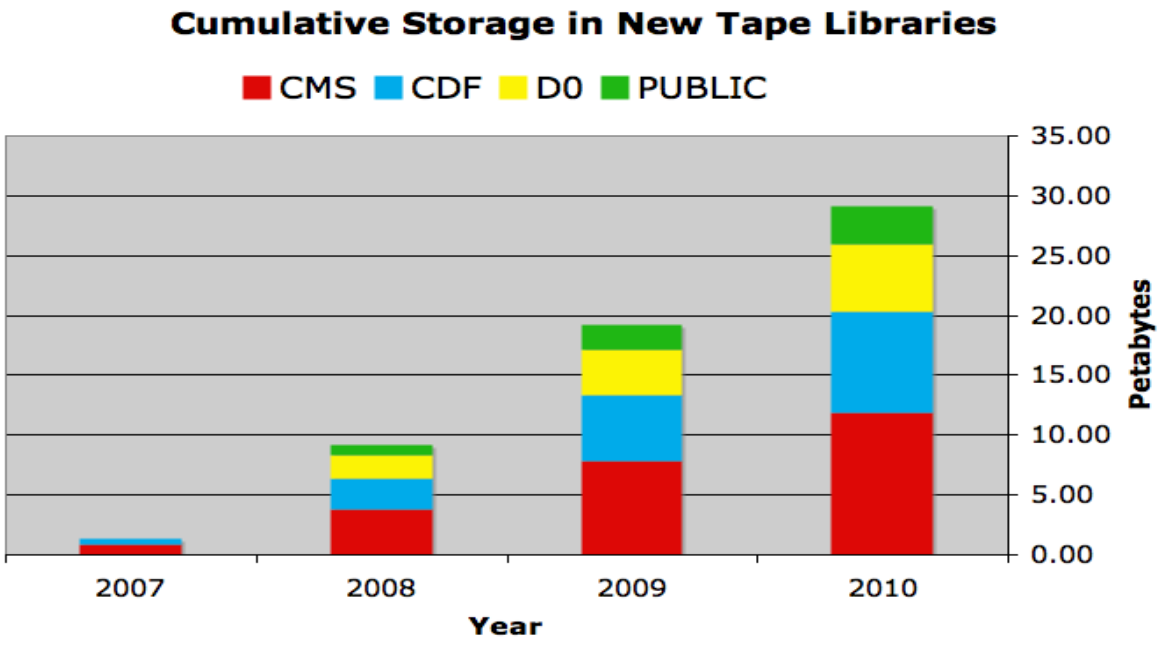
GRID and shared facilities and services

- Storage and data caching systems
- CMS Tier 1 facility and LPC analysis center
- FermiGrid
- Open Science Grid
- Wide area networking

Managing Petabytes of Storage

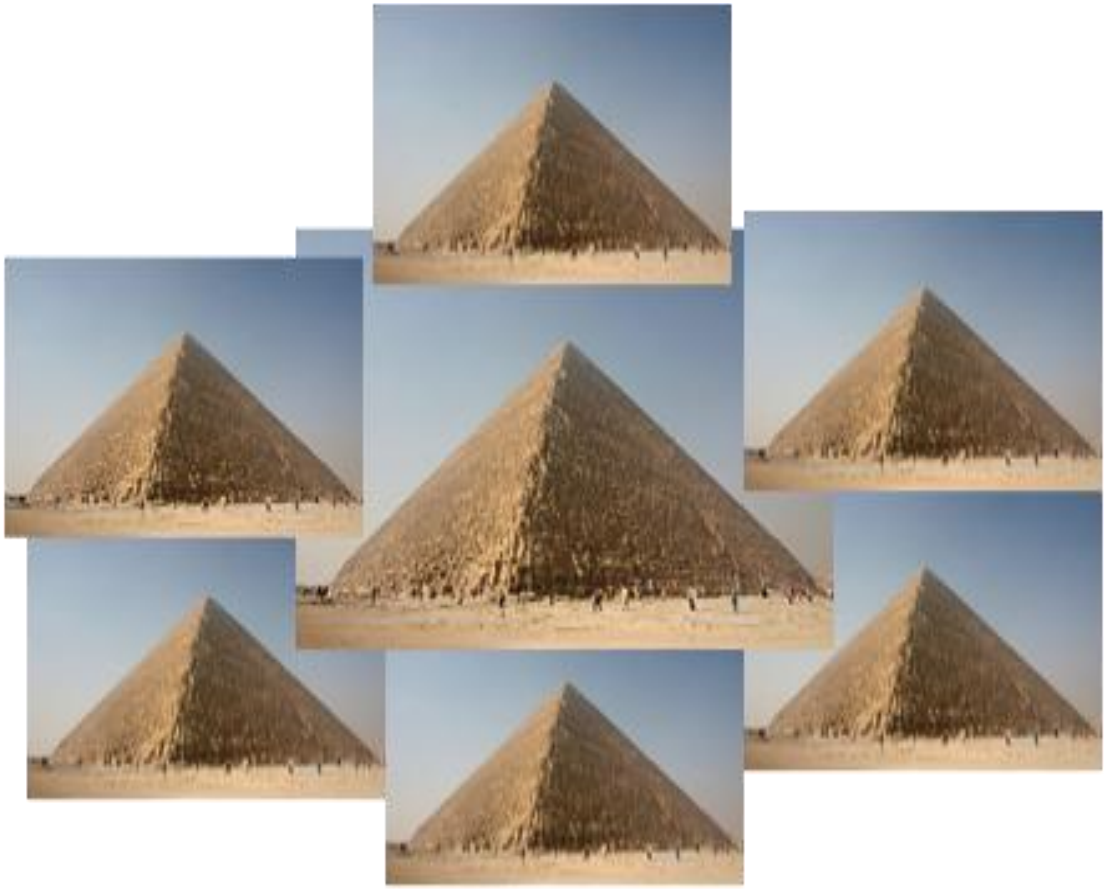
- Fermilab currently has nearly 9 Petabytes of near-line (actively used) long lived tape storage
  - Projecting > 25 PB by the end of FY 2010
- Vigilence, ongoing developments are required to assure:-
  - Integrity and Protection of data
  - Access to data – locally and globally, with enormous peak loads
  - Standard Grid access protocols (SRM) – global collaboration
  - Ongoing collaboration with DESY on SRM-dcache
  - Need to look at next generation solutions while scaling up and operating

Migrating all data from end-of-life tape libraries by 2009



Visualizing the data volume

If each byte were a grain of sand, 8.79×10^{15} bytes of data now on tape could build the Great Pyramid of Khufu seven times over.



We move more of this data and take better care of it than any other institution we know

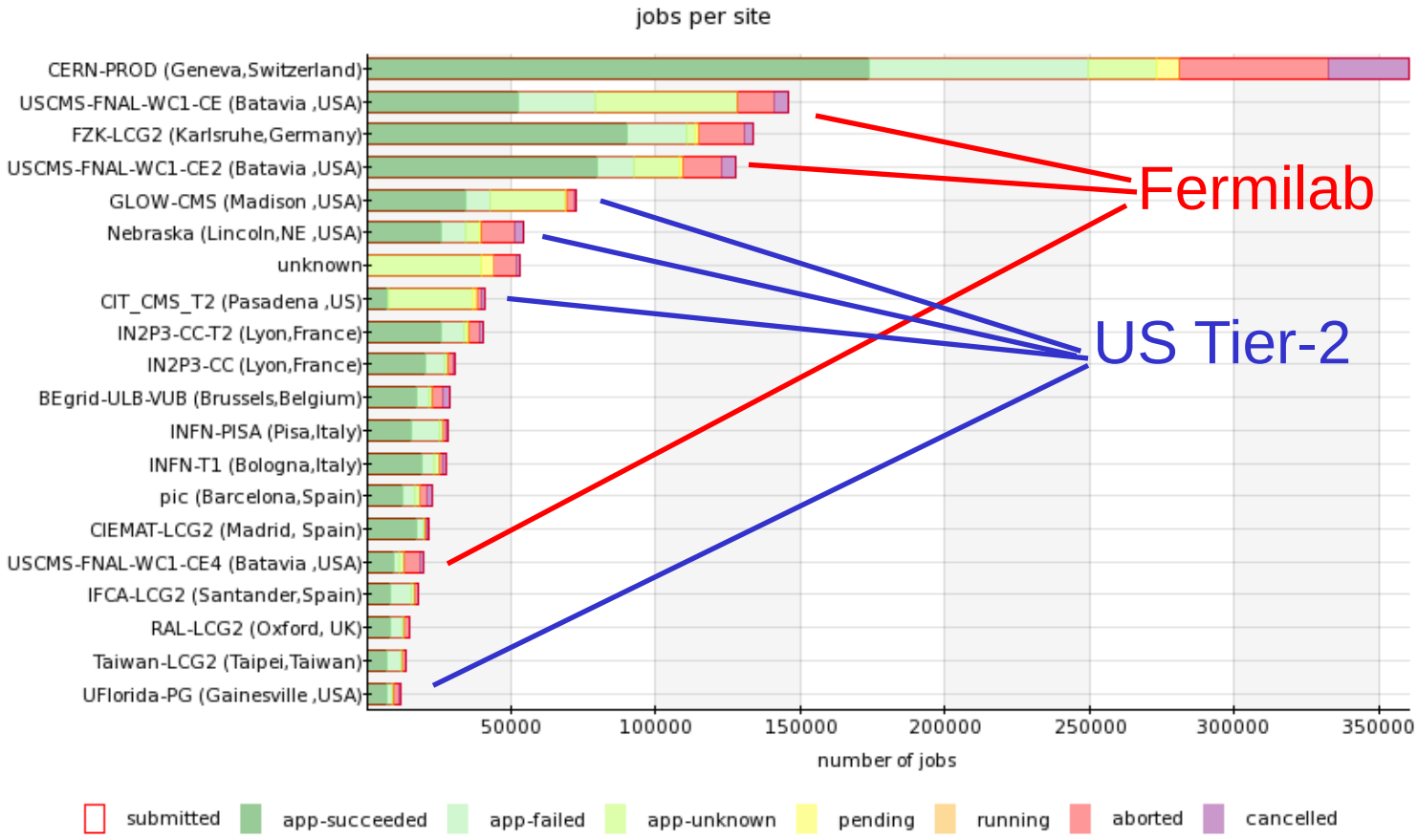
The US-CMS Tier-1 Computing Center

- The US-CMS Tier-1 at FNAL will complete its four year procurement ramp-up in the summer of 2008
  - The US-CMS Tier-1 is comparable in size to the CMS computing resources at CERN
  - The largest single Tier-1 computing center for CMS and the only CMS Tier-1 in the Americas

FNAL Tier-1 and LPC Summer 2008	CPU T1	5.5MSI2k	Tier-1 Processing Nodes
	CPU LPC	2.5MSI2k	Dedicated to Local Analysis
	Disk T1	2.2PB	dCache (1600MB/s IO)
	Disk LPC	0.5PB	Dedicated to Local Analysis
	Network	20Gb/s	CERN to FNAL
	People	30FTE	Includes Developers and Ops

CMS: Getting Ready for Physics

- Analysis jobs since the first of the year



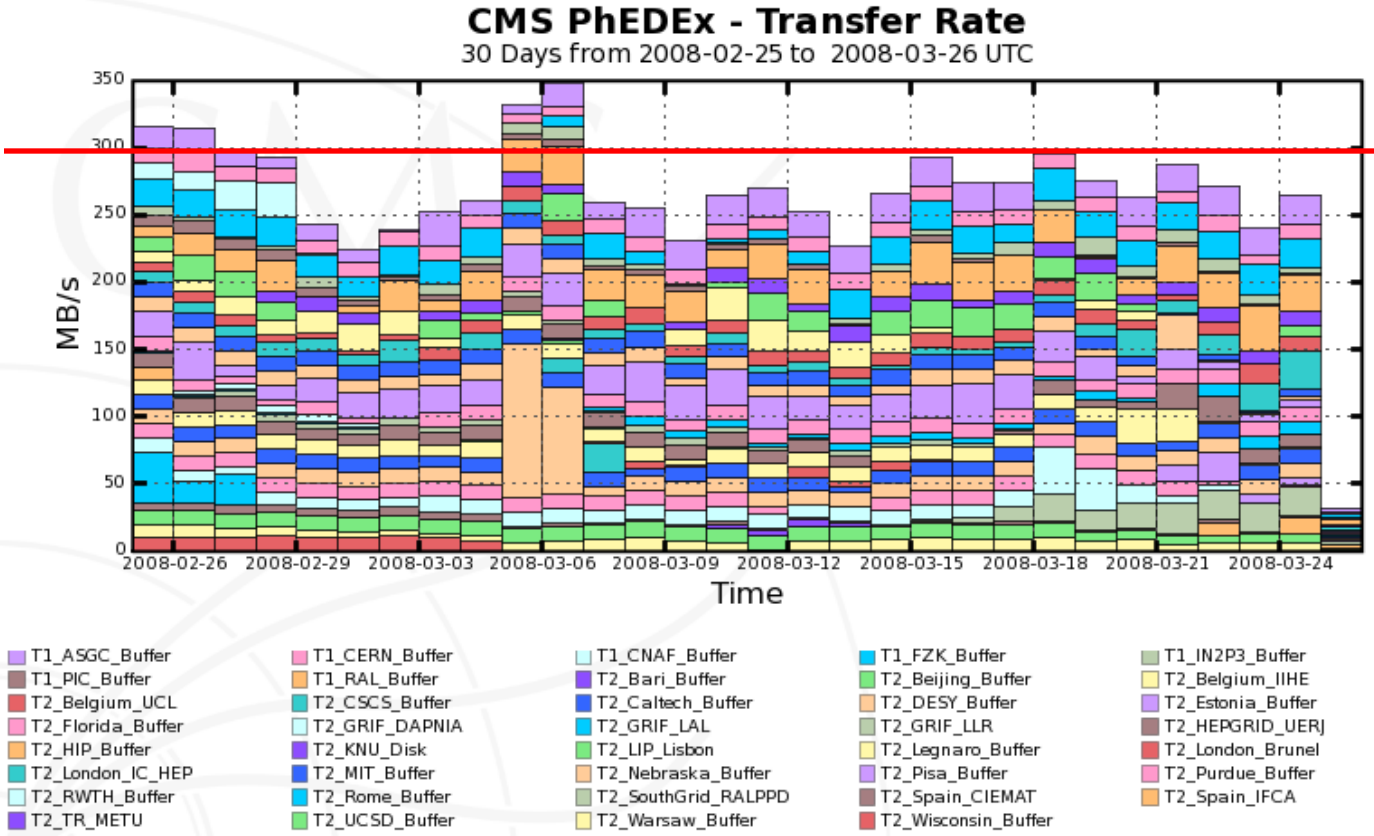
Supported by DOE and NSF jointly and running jobs on U.S and worldwide Grid resources

CMS: Data Transfers

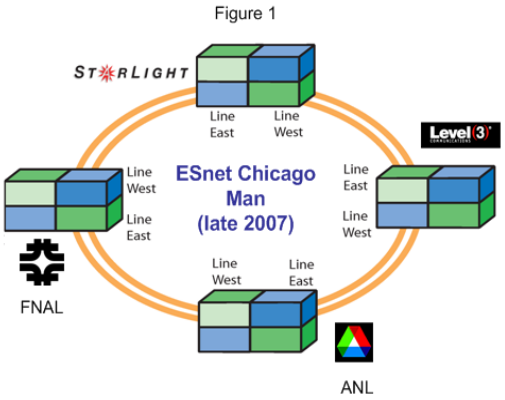
- Fermilab is able to routinely export data as multi-terabytes per day to US and external Tier-2 sites
  - In the last month data has been successfully transferred to 40 CMS sites

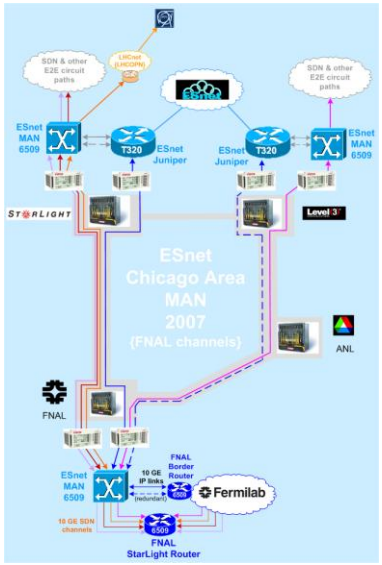
300MB/s

Export from Fermilab over the last 30 days



Chicago metropolitan area network (MAN)



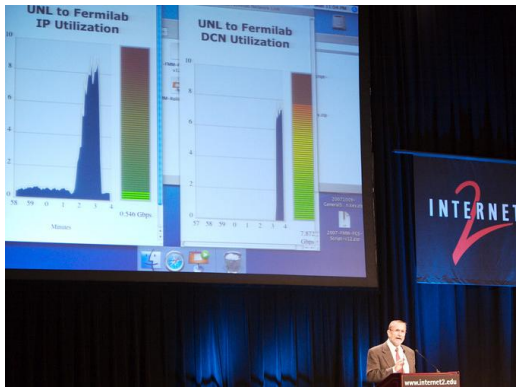


- Collective effort by Fermilab, Argonne N.L., and ESnet
  - Fermilab/Argonne jointly provide operational support
- Completed in December, 2007:
  - Redundant optical fiber cable paths for internet connectivity
  - Provides Fermilab with six 10Gb/s channels for offsite traffic
    - Exceeds ESnet's backbone

	StarLight	Level3 PoP	Total
Argonne	1	2	3
Fermilab	4	2	6
		MAN total:	9

Lambda Station: - A Fermilab R&D Networking Project

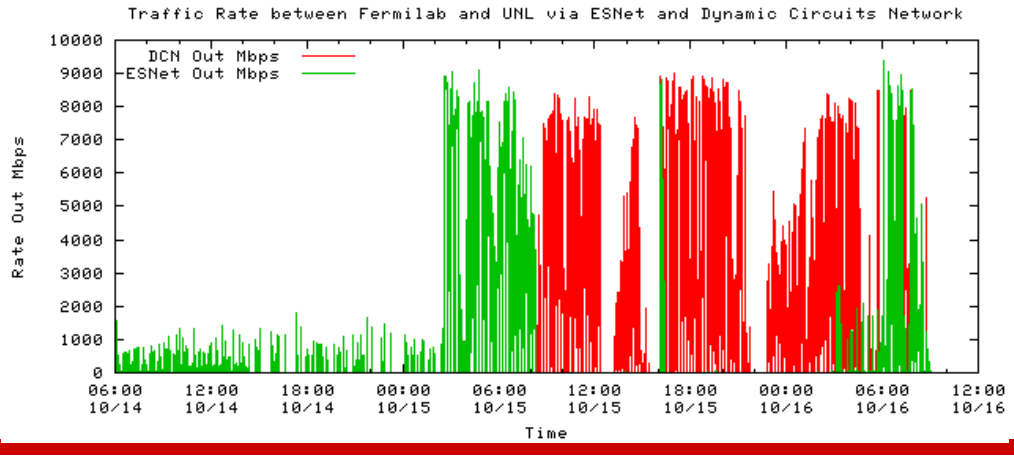
- Mission - enable local computing facilities to use dynamic wide area end-to-end data circuits
- Strategy - provide a network path selection service:
  - Initiates setup & tear down of dynamic circuits
  - Reroutes specific application traffic over circuit path
  - Graceful cutover and fallback on path changes
- Deployed between Fermilab & Univ. of Nebraska (a CMS Tier 2 site)
  - Successful demos at SC07 & Internet-2 Members Meeting (right)
  - A major component to emerging dynamic circuit services



Large-scale data recovery via dynamic circuit

- Shortly after deploying their dynamic circuit, Nebraska lost their entire Tier-2 data cache

- 50TB Data recovered by data transfer from FNAL Tier-1
  - Completed in 32 hours
  - Lambda Station switching over to dynamic Internet2/ESnet circuit
  - Graceful cutover to & fall back from the dynamic circuit

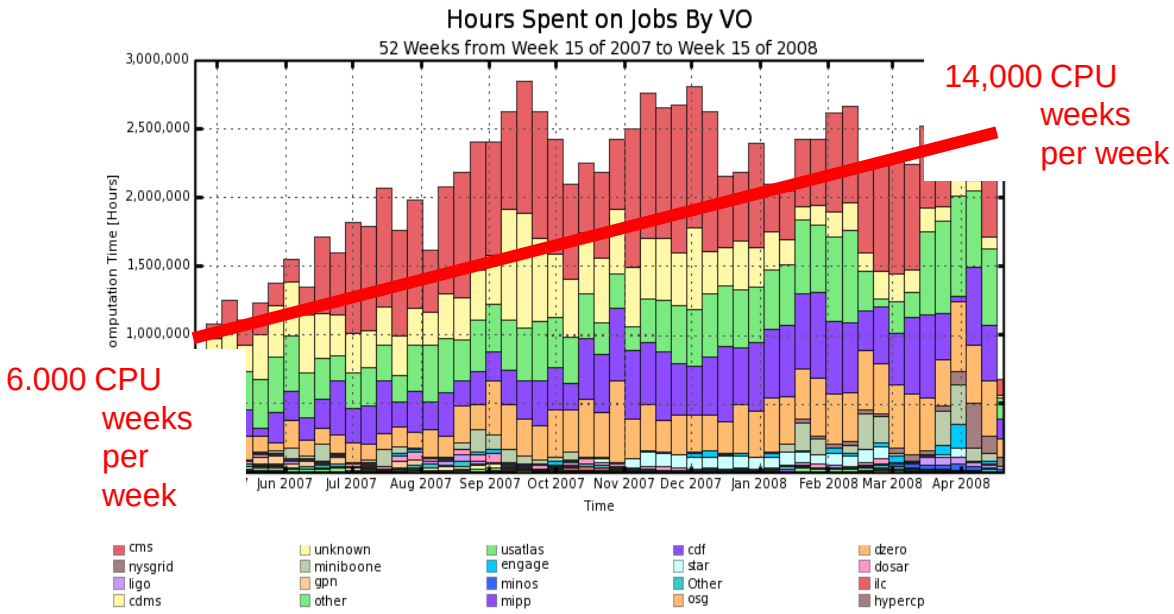


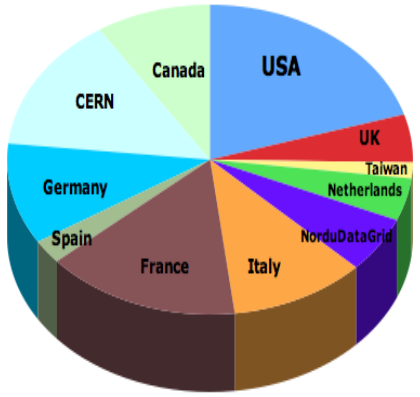
FermiGrid – a campus Grid

- Fermilab has adopted a strategy of placing all of its production computing resources in a Grid "meta-facility" called FermiGrid.
- Optimizes use of resources and allows Fermilab to:
  - Insure that the large experiments have first priority usage of dedicated resources purchased on their behalf.
  - Allow opportunistic use of dedicated and shared resources by Virtual Organizations (VO's) that participate in the Fermilab experimental program and by certain VO's that use the Open Science Grid (OSG)
  - Fully support Open Science Grid and the LHC Computing Grid and gain positive benefit from this infrastructure in the US and Europe
  - Provide shared, high availability, robust grid services for all stakeholders at Fermilab (including CMS)
- It has been highly successful and almost all Fermilab scientific programs and experiments are now able to benefit from use of the broader Grid (if they need to) - e.g. CDF, D0, MiniBoone, MINOS, legacy experiments, DES, SDSS,….
  - We have run 3 FermiGrid "schools" now – very well attended

Open Science Grid (OSG) – a DOE and NSF supported National Grid Infrastructure

- Recorded use of OSG more than doubled in past year. Over 100,000 jobs per day from 70 production sites.
- Use by more than 12 non-physics users reaching 10% of the total, harvesting unused cycles
- Contributions to Worldwide LHC Computing Grig (WLCG) from US more than 25% of overall processing throughput.
  - All US LHC (as well as STAR and LIGO) Tier-2s and more than 10 Tier-3s accessible through OSG.



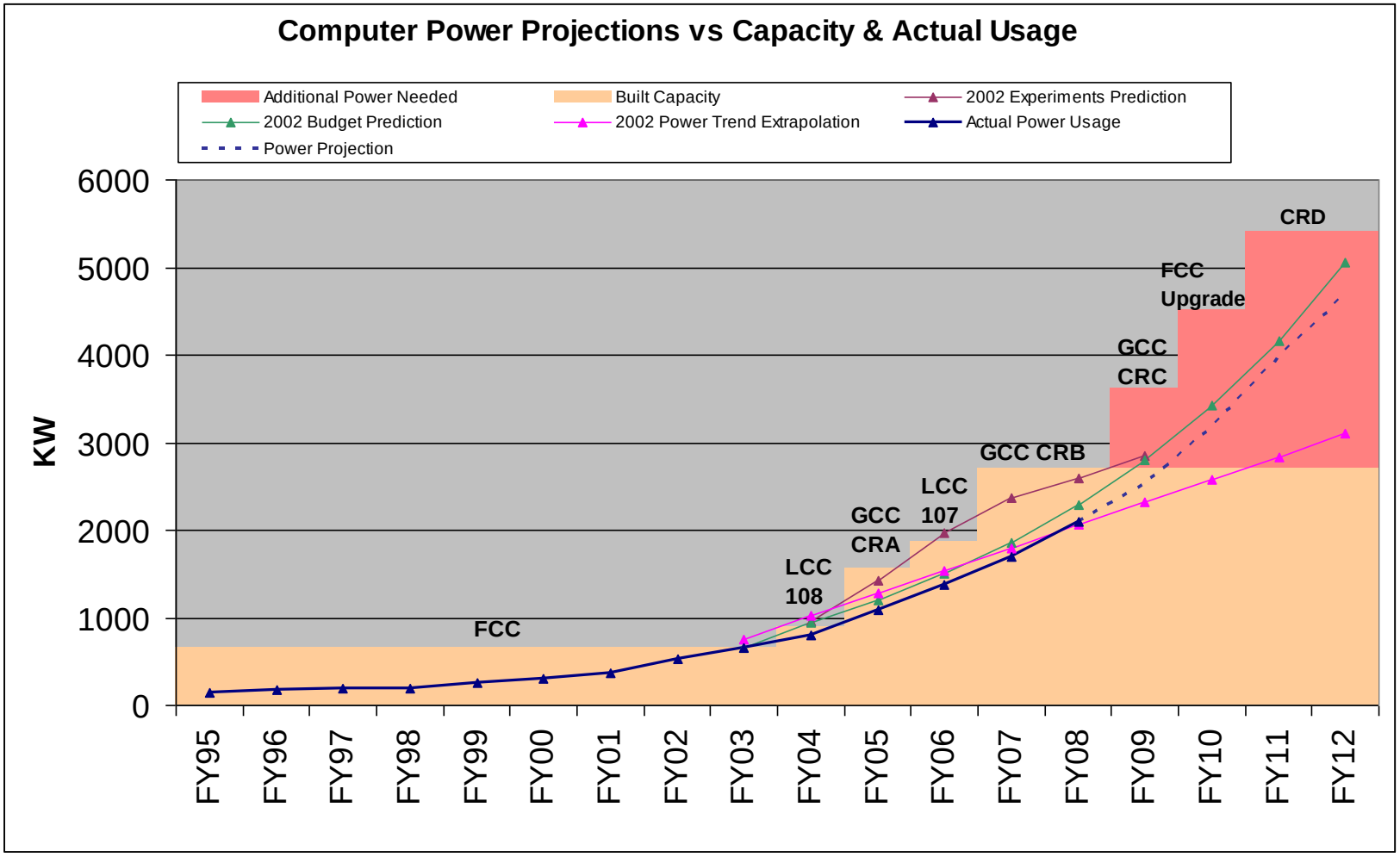


WLCG ATLAS and CMS processing usage for the past year (WLCG Apel reports).

Computer Centers, "IT", Computer Security

- Space, power and cooling
- Open Science security
- ITIL –relevant to both scientific computing and classic "IT"
  - service desk
  - service level agreements, metrics
  - incident management, configuration management
  - an emerging standard for running "IT" shops. From Wikipedia:
    - The **Information Technology Infrastructure Library** (**ITIL**) is a set of concepts and techniques for managing information technology (IT) infrastructure, development, and operations.
    - ITIL is published in a series of books, each of which cover an IT management topic. The names *ITIL* and *IT Infrastructure Library* are registered trademarks of the United Kingdom's Office of Government Commerce (OGC). ITIL gives a detailed description of a number of important IT practices with comprehensive checklists, tasks and procedures that can be tailored to any IT organization.

Computer Centers: Space, power, cooling



Computer Security and Open Science

- Surprisingly this is "Advanced Computing"
  - We must pay attention to this to assure we can keep our facilities "open" – yet safe
- Hope for opportunities for R&D and broader collaboration
  - DOE and Office of Science initiatives – Security R&D working group
  - PhD Computer Science researcher is OSG Security Officer
  - Can we do what Deming did to quality management to information security?
  - Submitted one proposal with LANL and PNNL – on the risks associated with huge set of resources (such as OSG, WLCG)
  - Looking at collaboration with social scientists on how to integrate security in users thinking

Extra Slides

- Some achivements of the past year including some R&D projects that are changing/have changed as a result of changes in the HEP program

ILC Detector Simulation was established in the Computing Division during 2007 (on hold now)

ILC Detectors are precision microscopes that push the frontier of simulation & software.

R. Kutschke

Development of reconstruction software for high precision tracking systems ramped up in Winter 2007.

H. Wenzel.

SNAP DAQ until recently

80 Detectors

*2Gb/s

80 Electronics Modules

*2Gb/s

2 Instrument Control Unit

*50Mb/s

Data compression
Data management
Data formatting
Interface to electronics modules

2 Slice Box

*300Mb/s

2 Telemetry

*300Mb/s

*8Gb/s

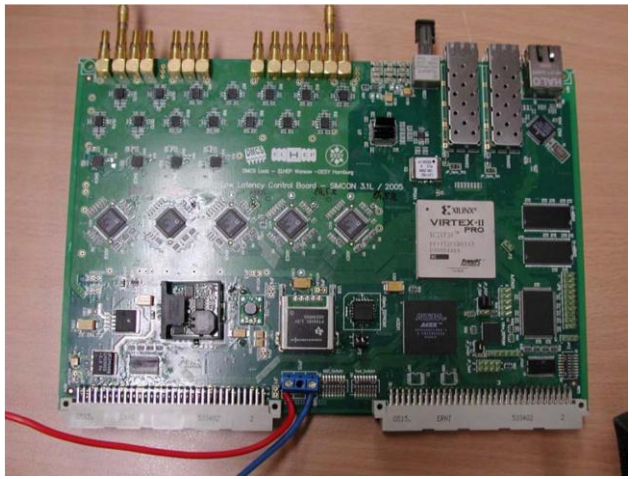
80 Memory Modules

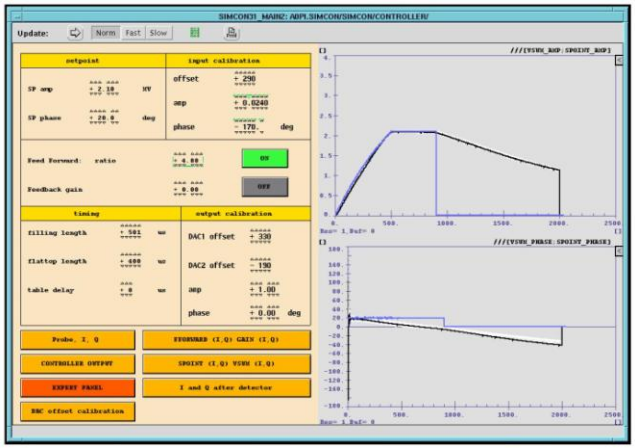
Participation:
- **Detector and readout electronics R&D: collaborating with LBL, UofM and FNAL/PPD**
- **Electronics modules: design of BeO port card**
- **Slice box: design of slice card and ICU interface**
- **Instrument control unit: collaborating with SLAC and LBL**
- **Memory modules: collaborating with FNAL/PPD**
- **Telemetry: off-the-shelf solution**

*** Peak data rates displayed. Data comes in bursts with a 10% duty cycle.**

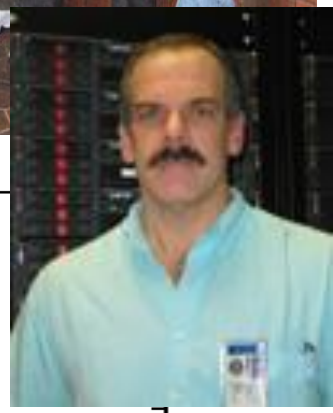
ILC Controls R&D

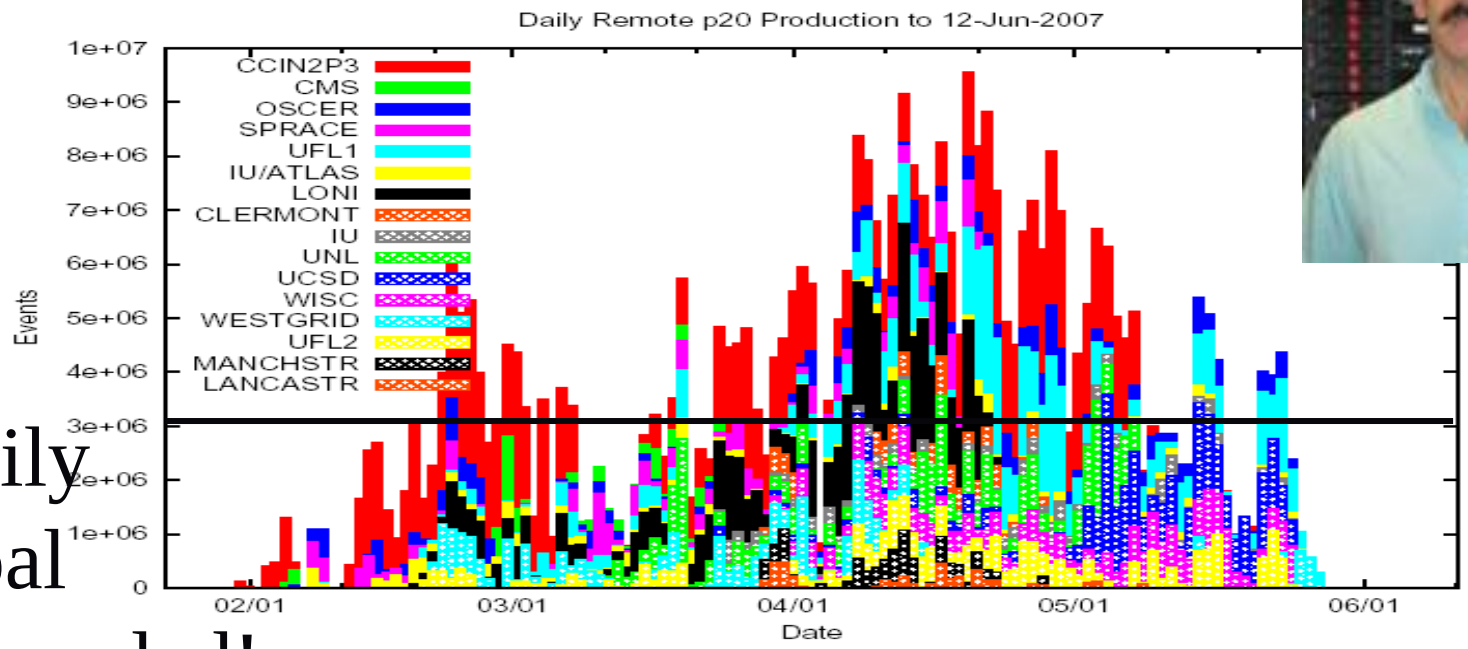
- CD collaborated on design of the ILC control system
  - Global effort with DESY, KEK, US institutions
  - Prototyping and testing at ILC test facilities
- R&D in close collaboration with other US groups (FNAL/ANL/SLAC/LBNL/ORNL/UPenn)
  - Timing and RF distribution
  - High availability and reliable hardware and software
  - Frameworks for controls software
  - Low Level RF simulations and controls



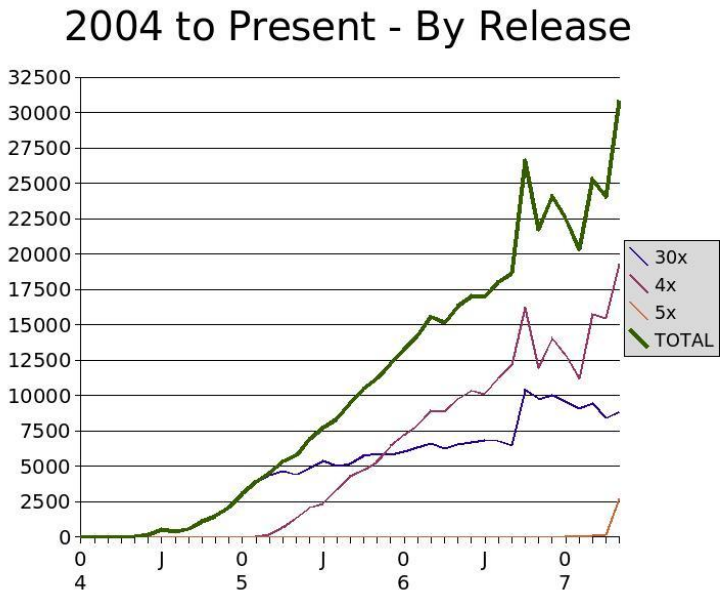


February ’07-April ’07
DO reprocesses
500M Run IIb
Collider events
Using OSG and other
Grid resources





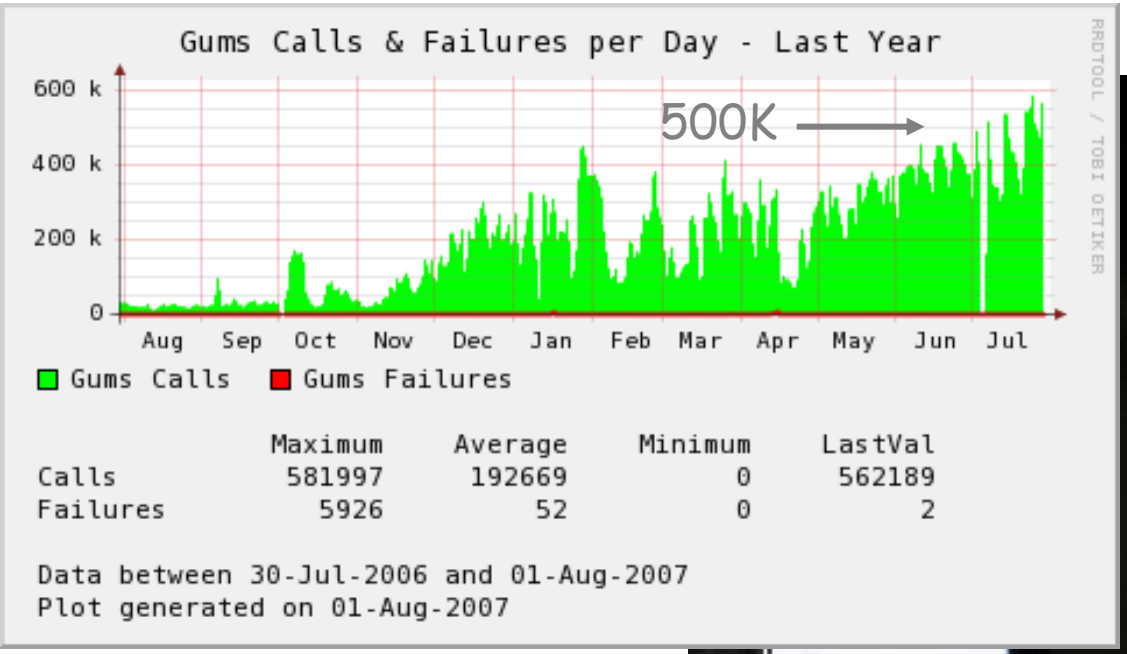
Daily Goal Exceeded!



April 2007: Systems running Fermilab Scientific Linux exceed 30,000 in less than 3 years.

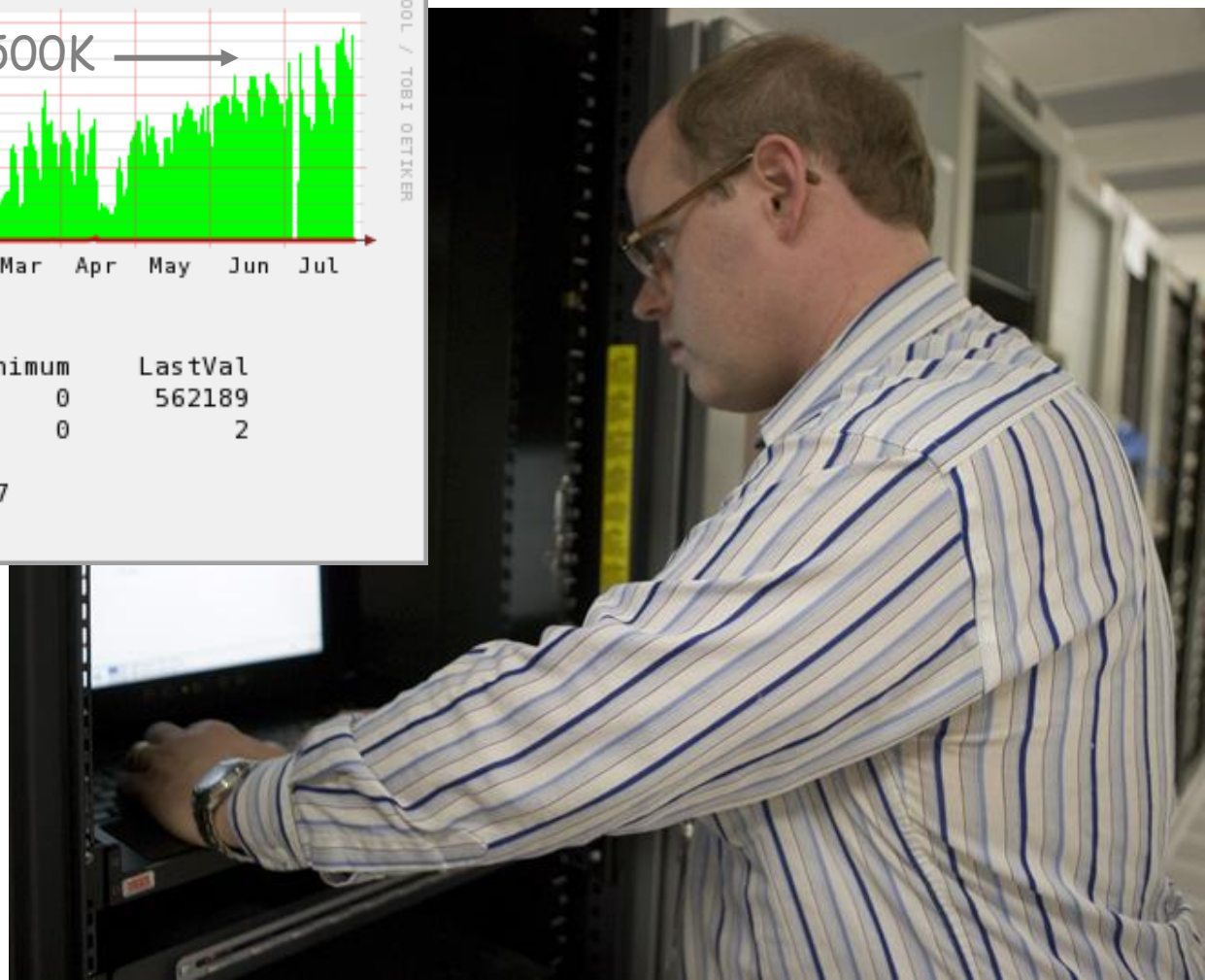
Scientific Linux

July 2007 : Surpass 500,000 calls per day to FermiGrid user mapping service



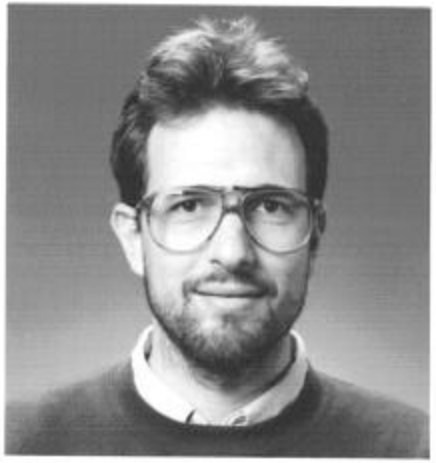
	Maximum	Average	Minimum	LastVal
Calls	581997	192669	0	562189
Failures	5926	52	0	2

Data between 30-Jul-2006 and 01-Aug-2007
Plot generated on 01-Aug-2007



July 2007: >125 web sites centrally hosted, over 100,000 visitors a month to www.fnal.gov…

…and over 3 million email messages delivered each week.



Daniel Elvira

Fermilab Joined the Geant4 Collaboration in June 2007 and found ways to speed up the simulation package by 4%

CMS Simulation at the LHC Physics Center