Fermilab, the Grid & Cloud Computing Department and the KISTI Collaboration

Overview

- Fermi National Accelerator Laboratory
- The Computing Sector
- The Grid & Cloud Computing department
- Collaboration with KISTI

GSDC - KISTI Workshop
Jangsan-ri, Nov 4, 2011

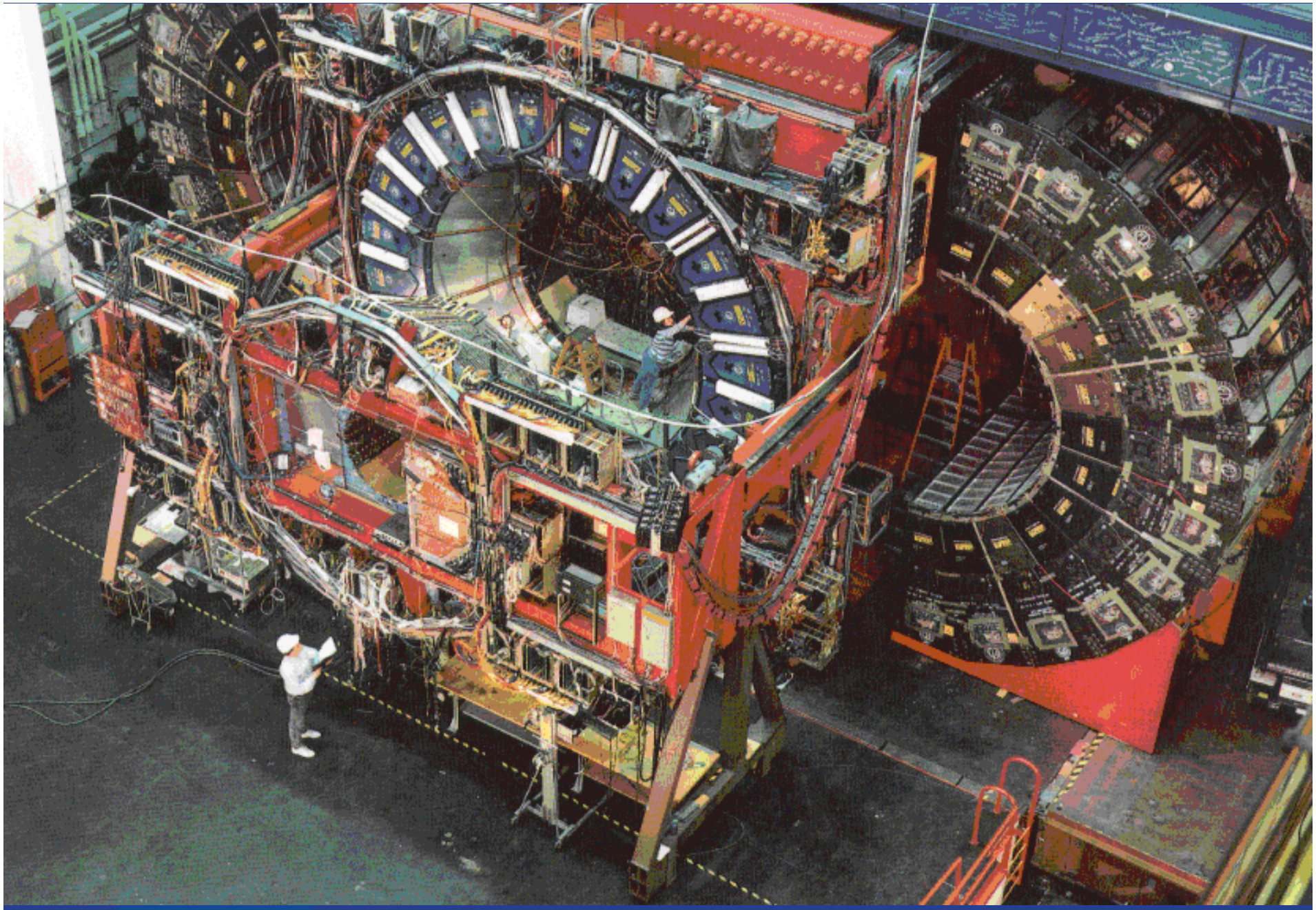
Gabriele Garzoglio
Grid & Cloud Computing Department, Associate Head
Computing Sector, Fermilab

U.S. DEPARTMENT OF ENERGY

Fermilab

Fermi National Accelerator Laboratory

USGS
0
0.5mi

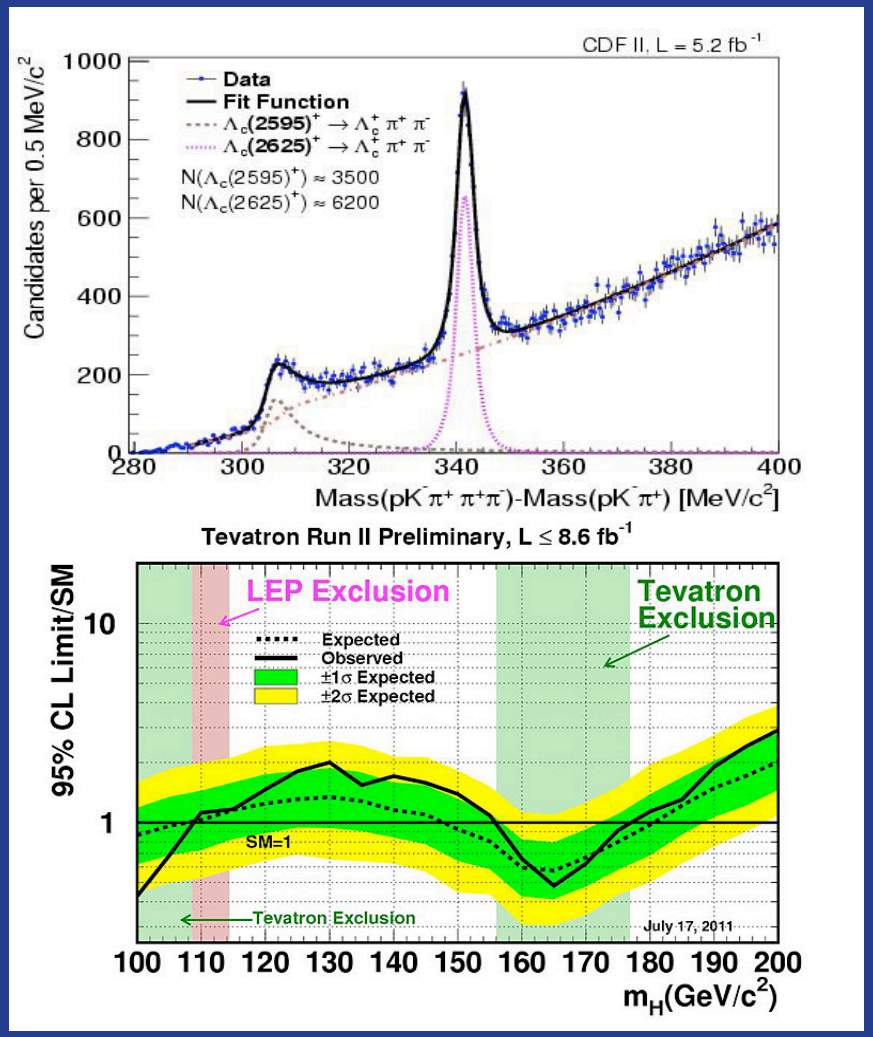


TeVatron Shutdown

- On Friday 30-Sep-2011, the Fermilab TeVatron was shut down following 28 years of operation,
- The collider reached peak luminosities of 4×10^{32} per centimeter squared per second,
- The CDF and Dzero detectors recorded 8.63 PB and 7.54 PB of data respectively, corresponding to nearly 12 inverse femtobarns of data,
- CDF and Dzero data analysis continues,
- Fermilab has committed to 5+ years of support for analysis and 10+ years for access to data.

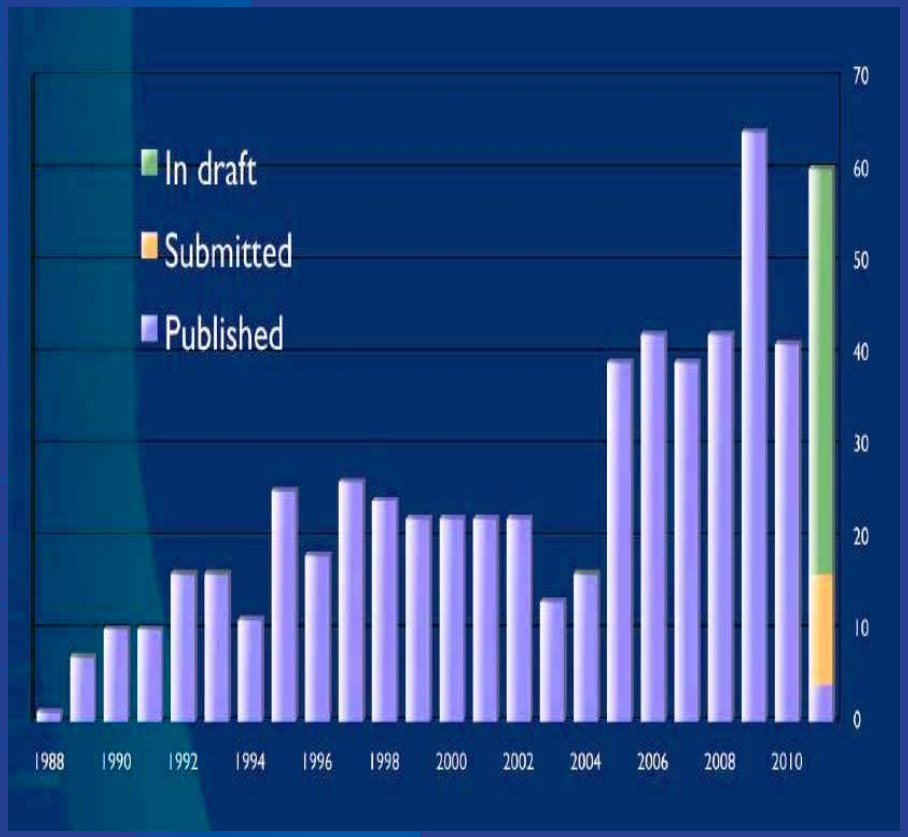
Physics!

- CDF & D0 - Top Quark Asymmetry Results.
- CDF - Discovery of Ξ_b^0.
- CDF - $\Lambda_c(2595)$ baryon.
- Combined CDF & D0 Limits on standard model higgs mass.

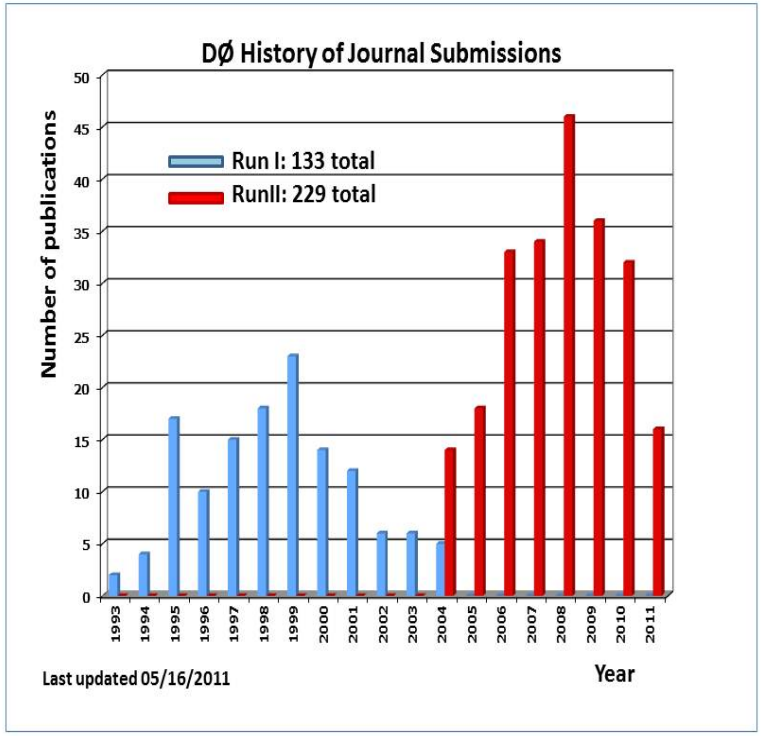


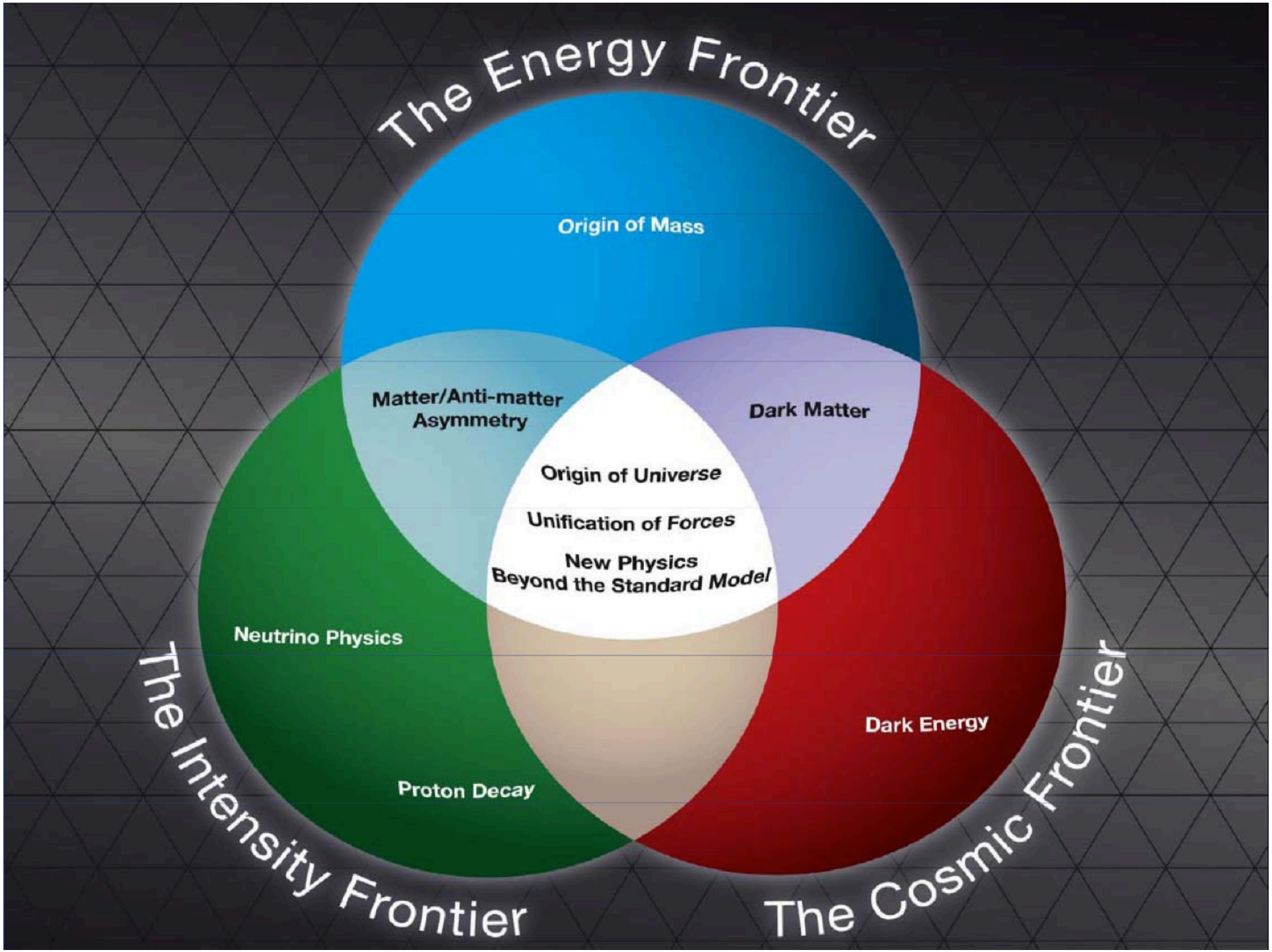
CDF & D0 Publications

CDF

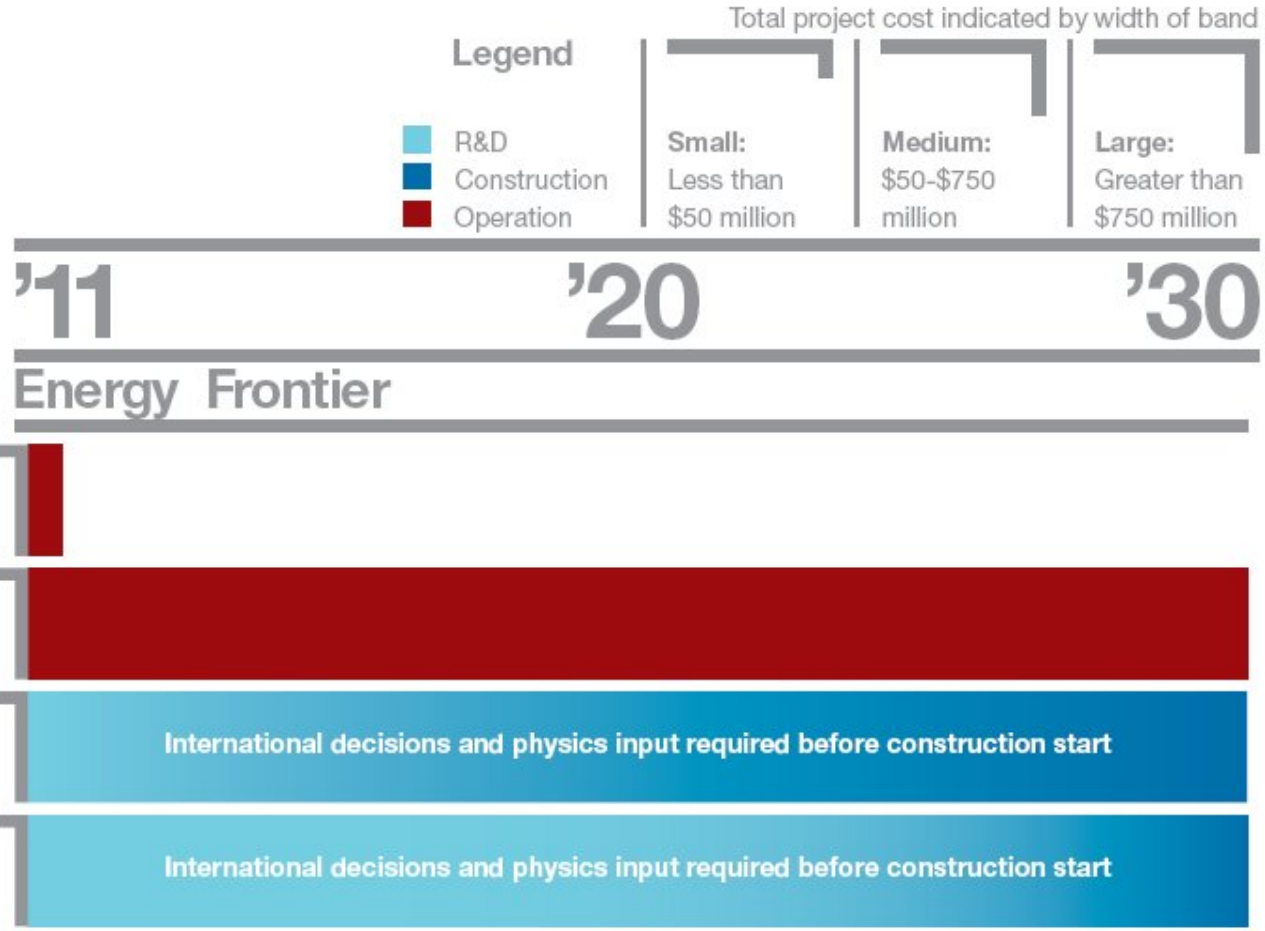


D0

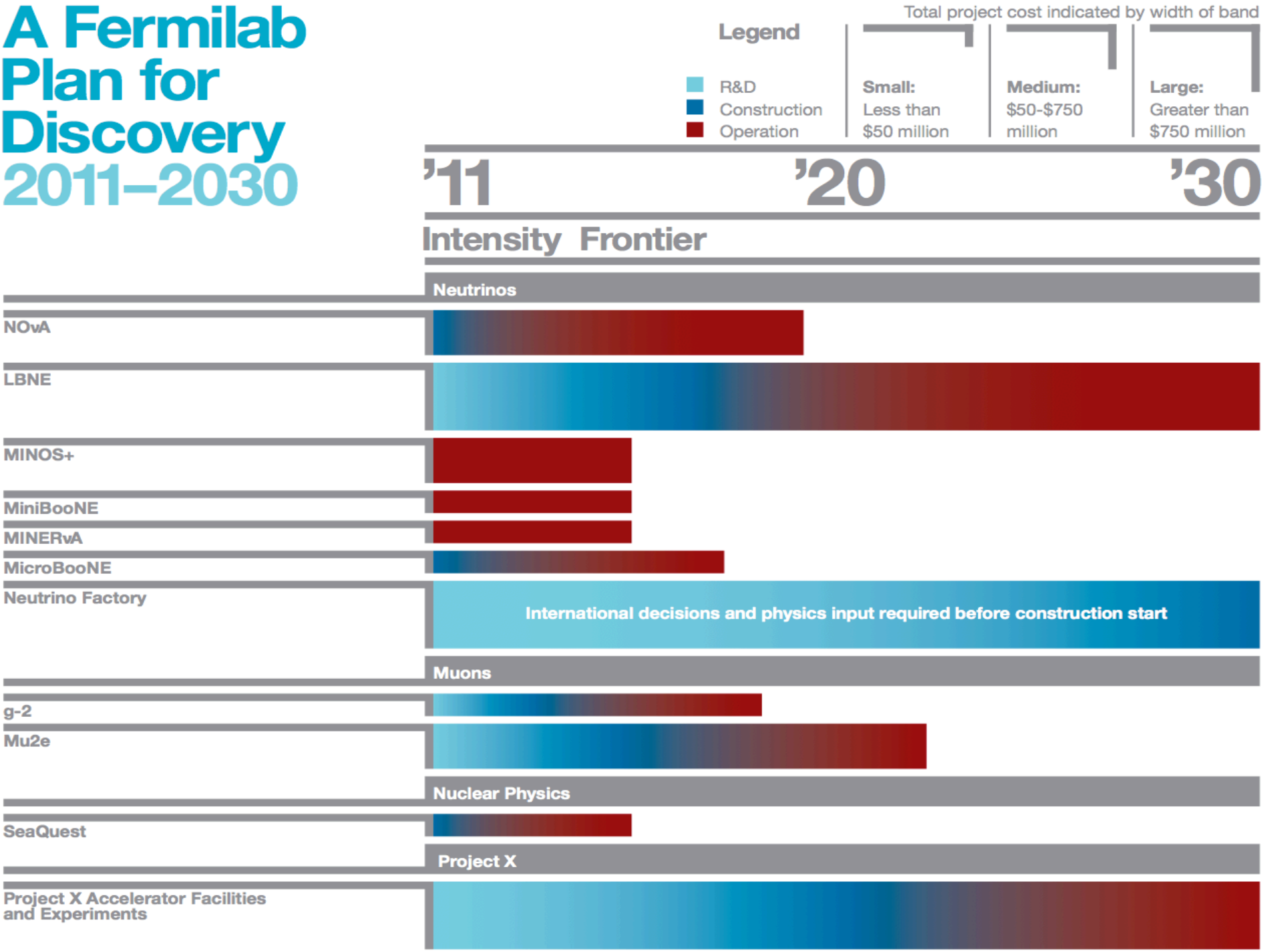



The Energy Frontier
Origin of Mass
Matter/Anti-matter Asymmetry
Dark Matter
Origin of Universe
Unification of Forces
New Physics Beyond the Standard Model
Neutrino Physics
Proton Decay
Dark Energy
The Intensity Frontier
The Cosmic Frontier

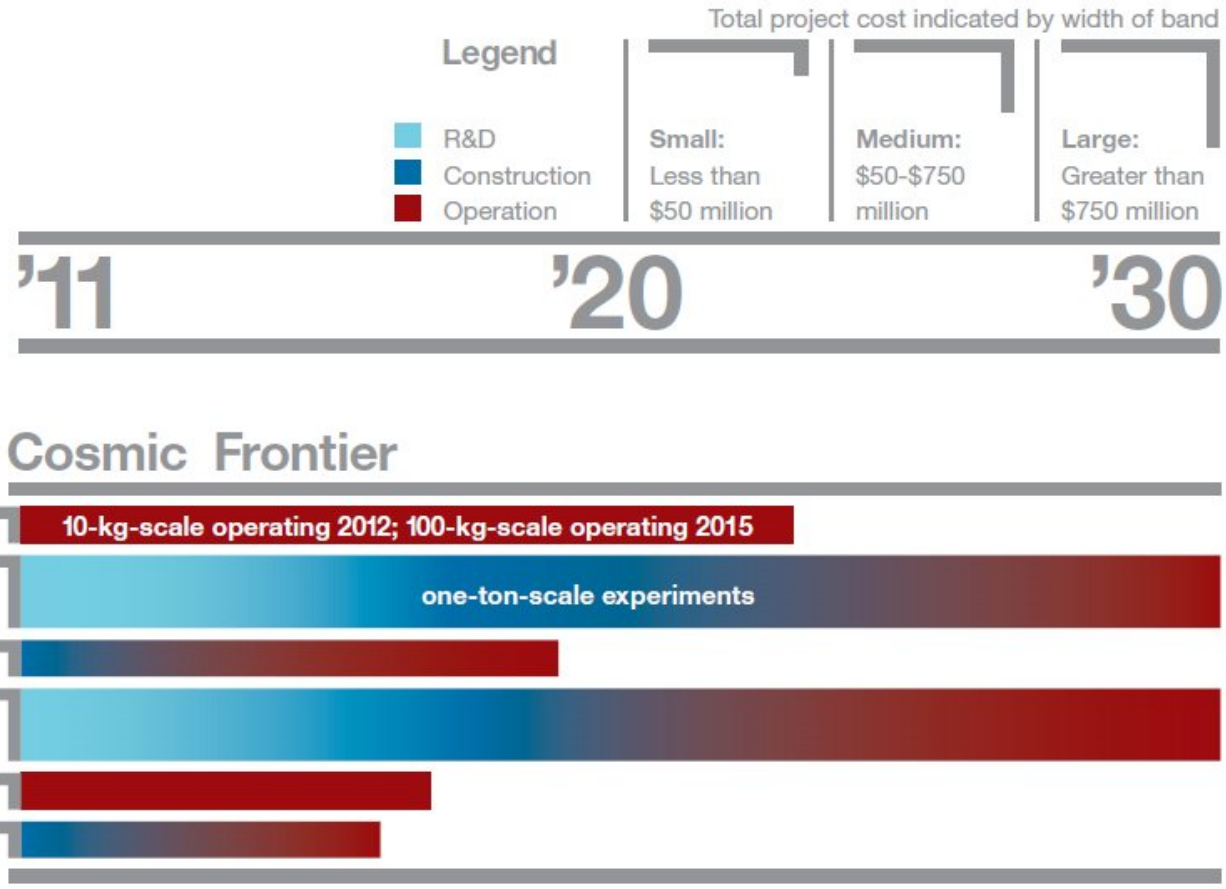
A Fermilab Plan for Discovery 2011–2030



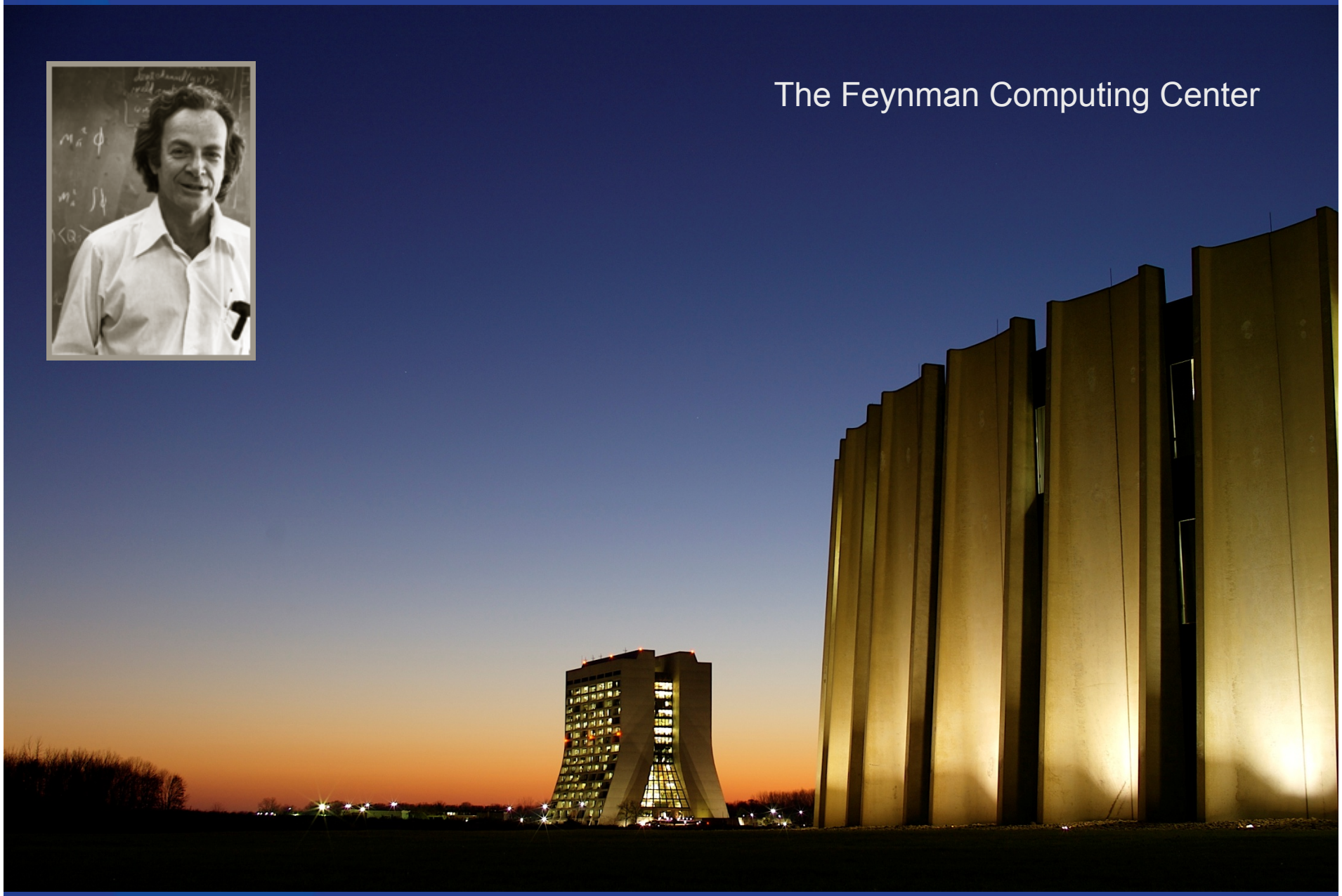
A Fermilab Plan for Discovery 2011–2030



A Fermilab Plan for Discovery 2011–2030



The Feynman Computing Center

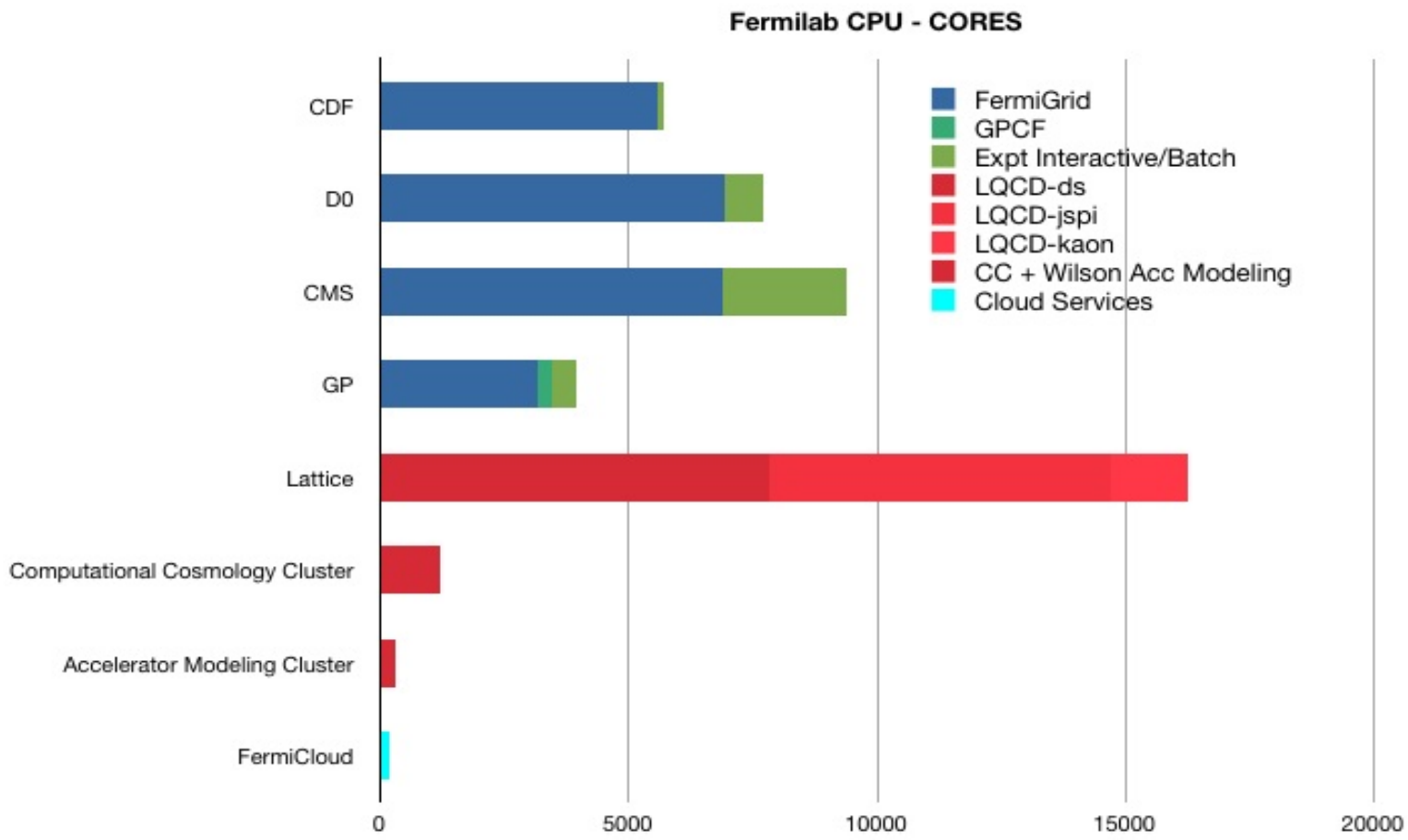


Fermilab Computing Facilities

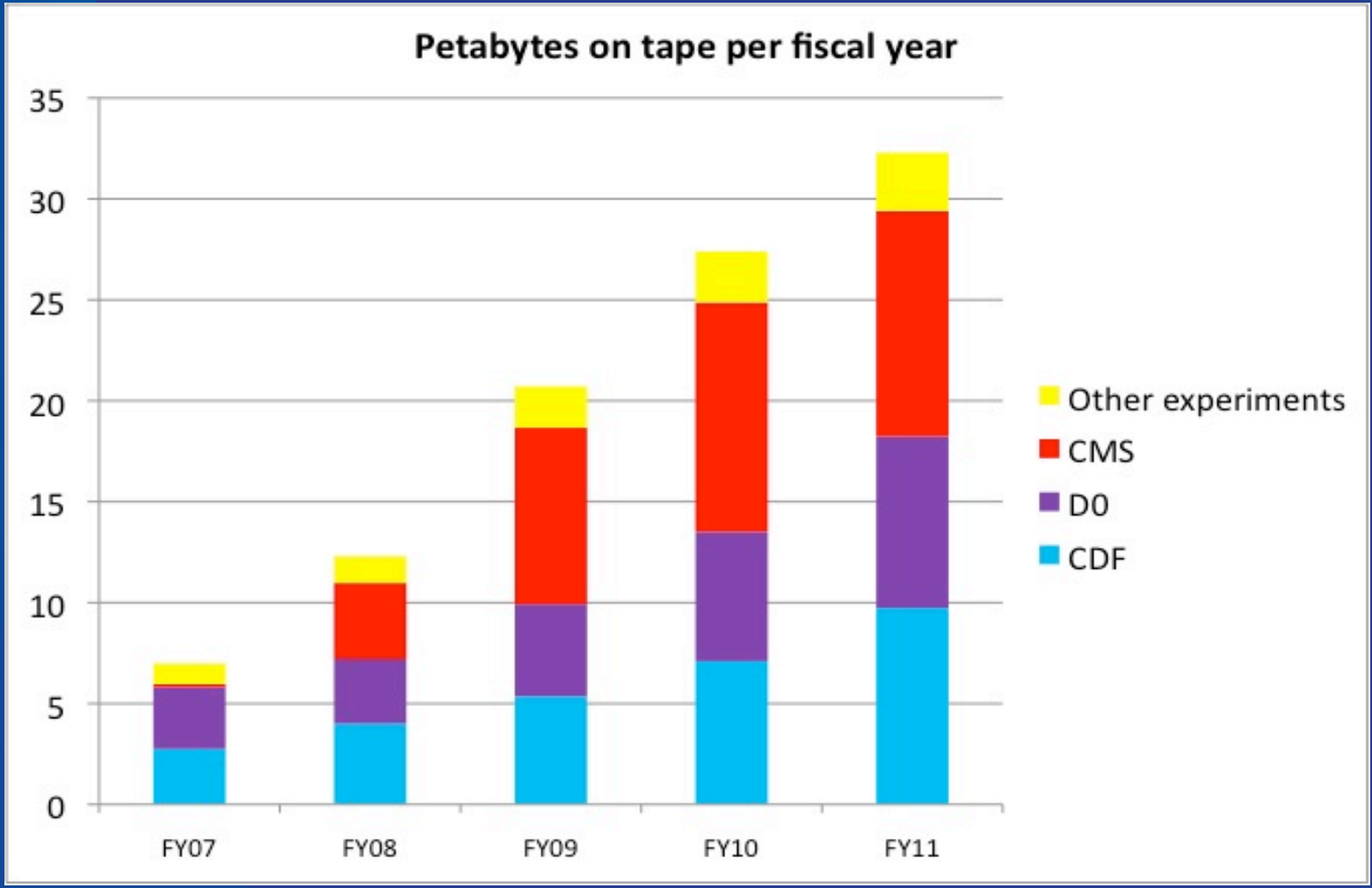
- Feynman Computing Center:
  - FCC2
  - FCC3
  - High availability services – e.g. core network, email, etc.
  - Tape Robotic Storage (3 10000 slot libraries)
  - UPS & Standby Power Generation
  - ARRA project: upgrade cooling and add HA computing room - completed
- Grid Computing Center:
  - 3 Computer Rooms – GCC-[A,B,C]
  - Tape Robot Room – GCC-TRR
  - High Density Computational Computing
  - CMS, RUNII, Grid Farm batch worker nodes
  - Lattice HPC nodes
  - Tape Robotic Storage (4 10000 slot libraries)
  - UPS & taps for portable generators
- Lattice Computing Center:
  - High Performance Computing (HPC)
  - Accelerator Simulation, Cosmology nodes
  - No UPS

EPA Energy Star award for 2010

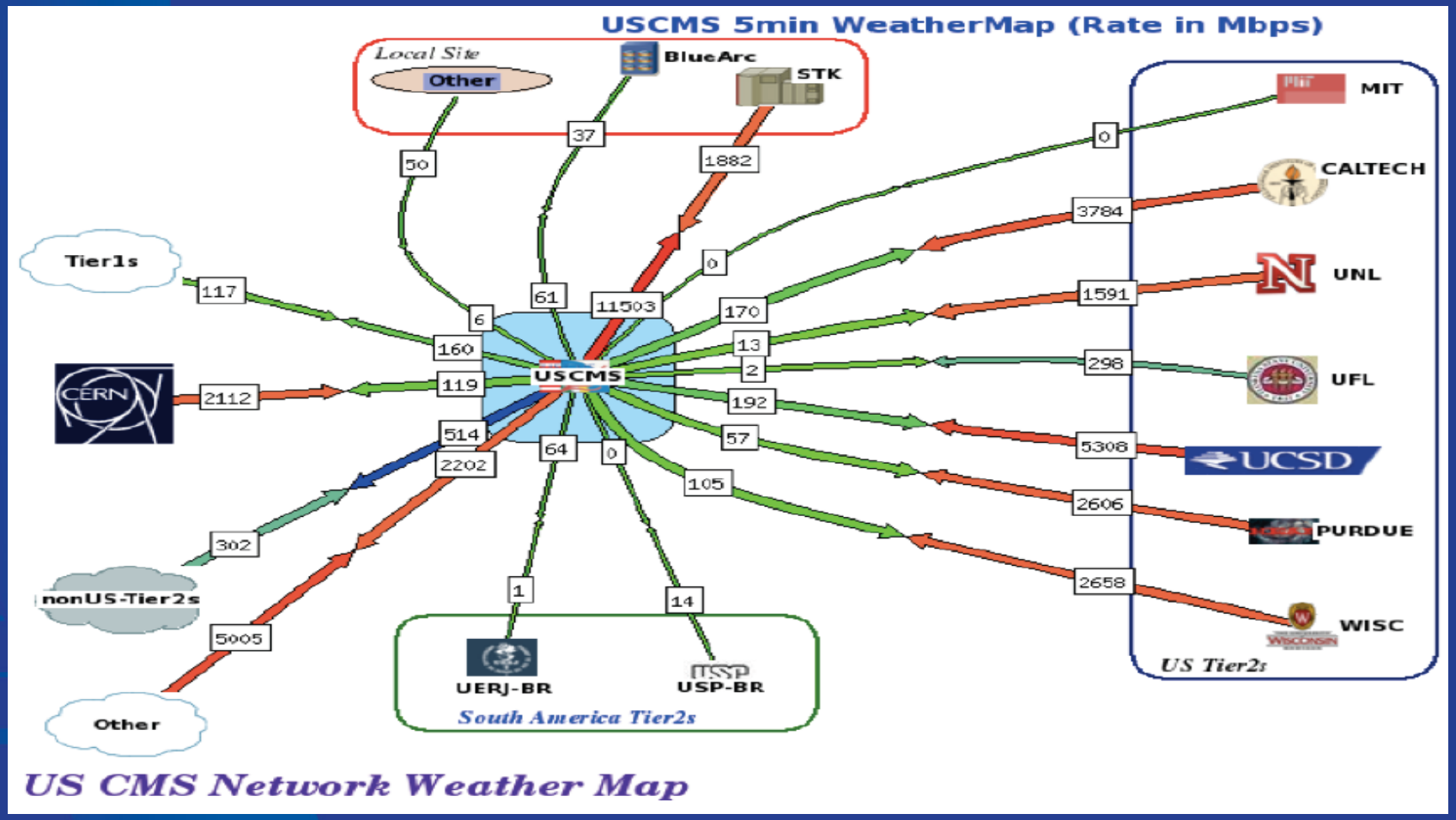
Fermilab CPU Core Count

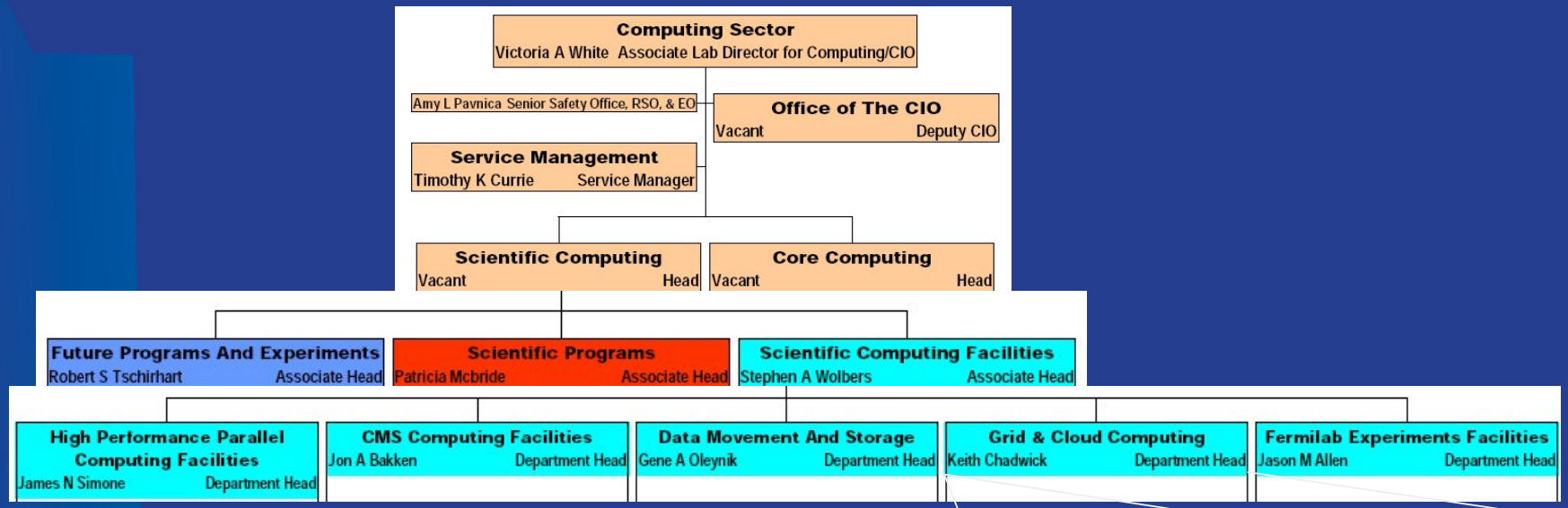


Data Storage at Fermilab



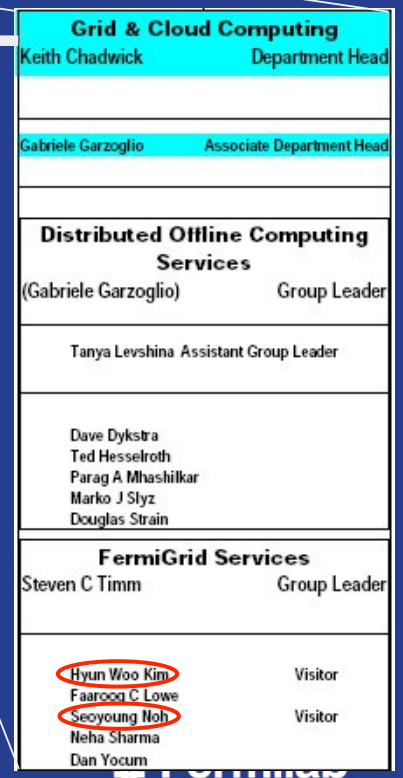
High Speed Networking





To establish and maintain Fermilab as a high performance, robust, highly available, expertly supported and documented Premier Grid Facility which supports the scientific program based on Computing Division and Laboratory priorities.

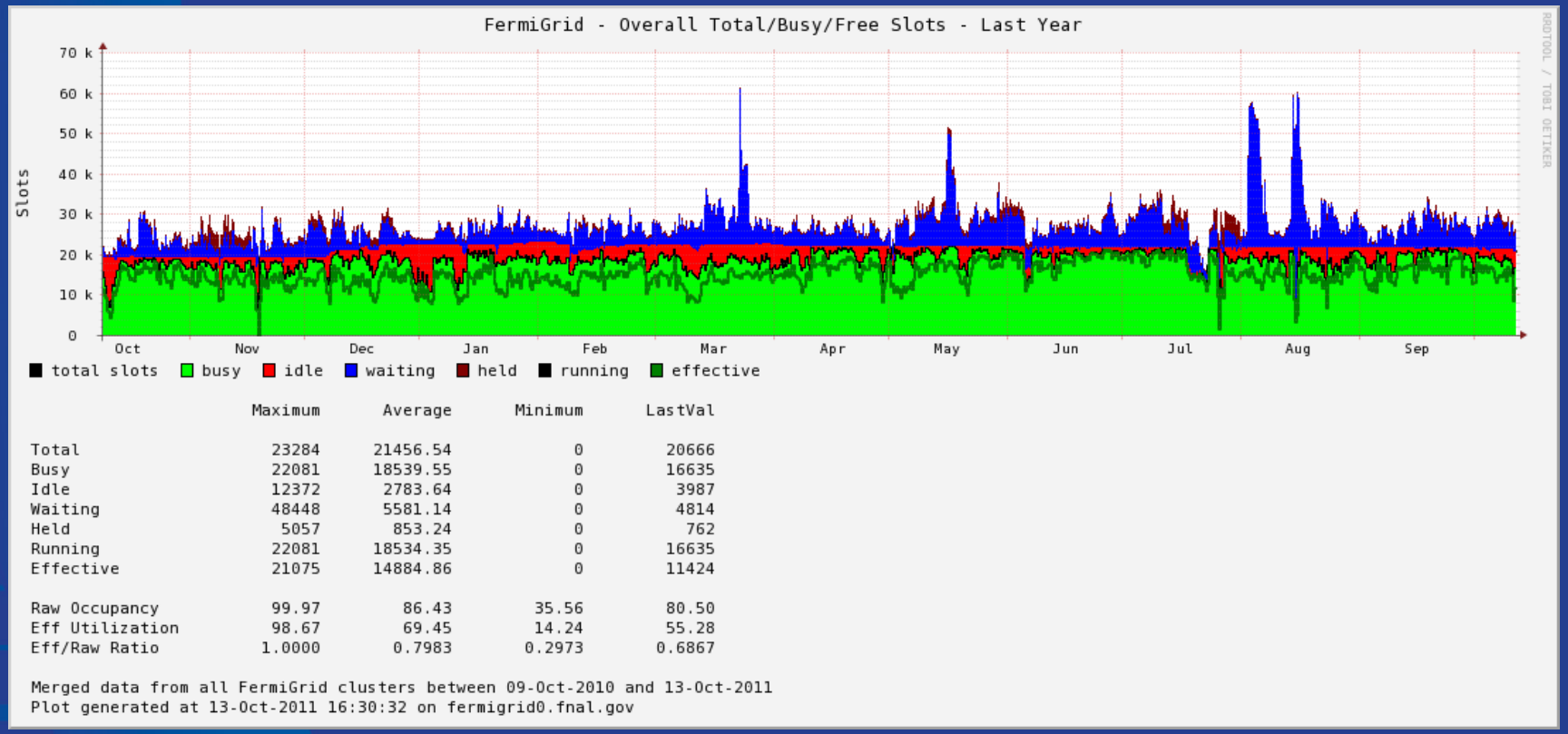
The department provides production infrastructure and software, together with leading edge and innovative computing solutions, to meet the needs of the Fermilab Grid community and takes a strong leadership role in the development and operation of the global computing infrastructure of the Open Science Grid.



FermiGrid Occupancy & Utilization

Cluster(s)	Current Size (Slots)	Average Size (Slots)	Average Occupancy	Average Utilization
CDF	5630	5477	93%	67%
CMS	7132	6772	94%	87%
D0	6540	6335	79%	53%
GP	3042	2890	78%	68%
Total	21927	21463	87%	70%

Overall Occupancy & Utilization



	Maximum	Average	Minimum	LastVal
Total	23284	21456.54	0	20666
Busy	22081	18539.55	0	16635
Idle	12372	2783.64	0	3987
Waiting	48448	5581.14	0	4814
Held	5057	853.24	0	762
Running	22081	18534.35	0	16635
Effective	21075	14884.86	0	11424
Raw Occupancy	99.97	86.43	35.56	80.50
Eff Utilization	98.67	69.45	14.24	55.28
Eff/Raw Ratio	1.0000	0.7983	0.2973	0.6867

Merged data from all FermiGrid clusters between 09-Oct-2010 and 13-Oct-2011
Plot generated at 13-Oct-2011 16:30:32 on fermigrid0.fnal.gov

FermiGrid-HA2 Service Availability

Service	Raw Availability	HA Configuration	Measured HA Availability	Minutes of Downtime
VOMS – VO Management Service	99. 657%	Active-Active	100.000%	0
GUMS – Grid User Mapping Service	99.652%	Active-Active	100.000%	0
SAZ – Site AuthoriZation Service	99.657%	Active-Active	100.000%	0
Squid – Web Cache	99.640%	Active-Active	100.000%	0
MyProxy – Grid Proxy Server	99.954%	Active-Standby	99.954%	240
ReSS – Resource Selection Service	99.635%	Active-Active	100.000%	0
Gratia – Fermilab and OSH Accounting	99.365%	Active-Standby	99.997%	120
Databases	99.765%	Active-Active	99.988%	60

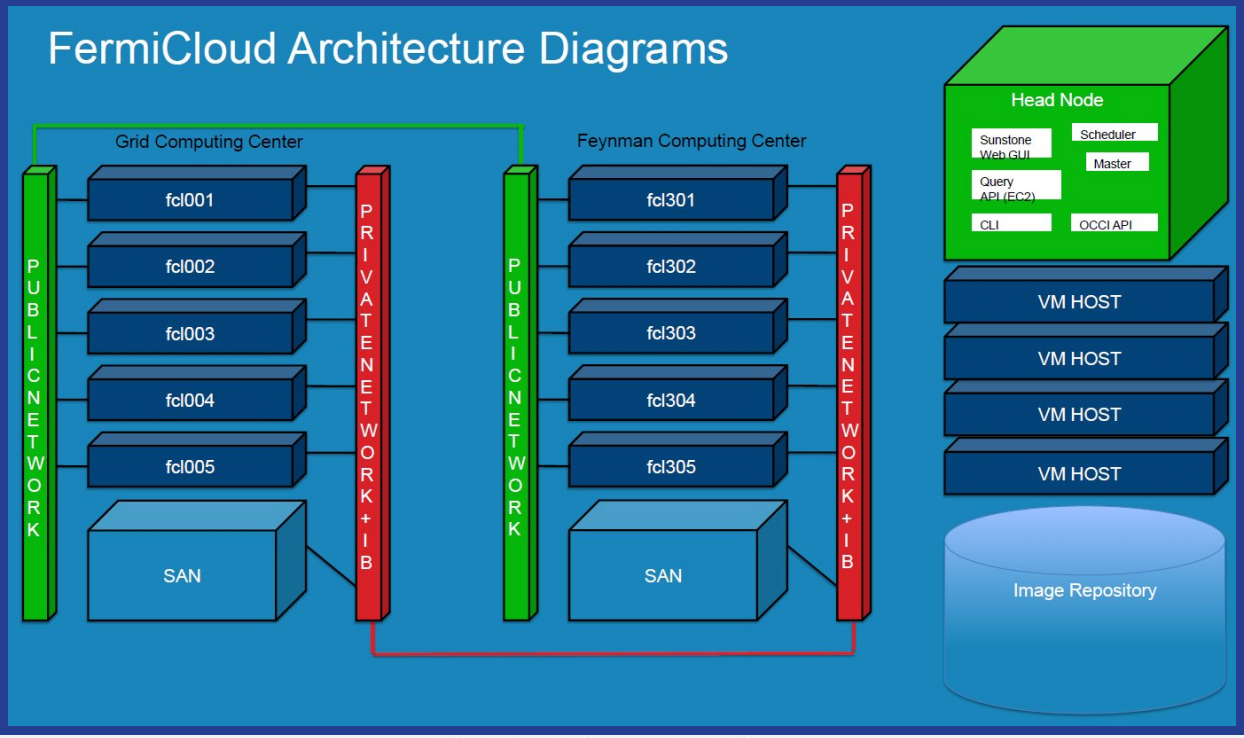
FY2011 Worker Node Acquisition

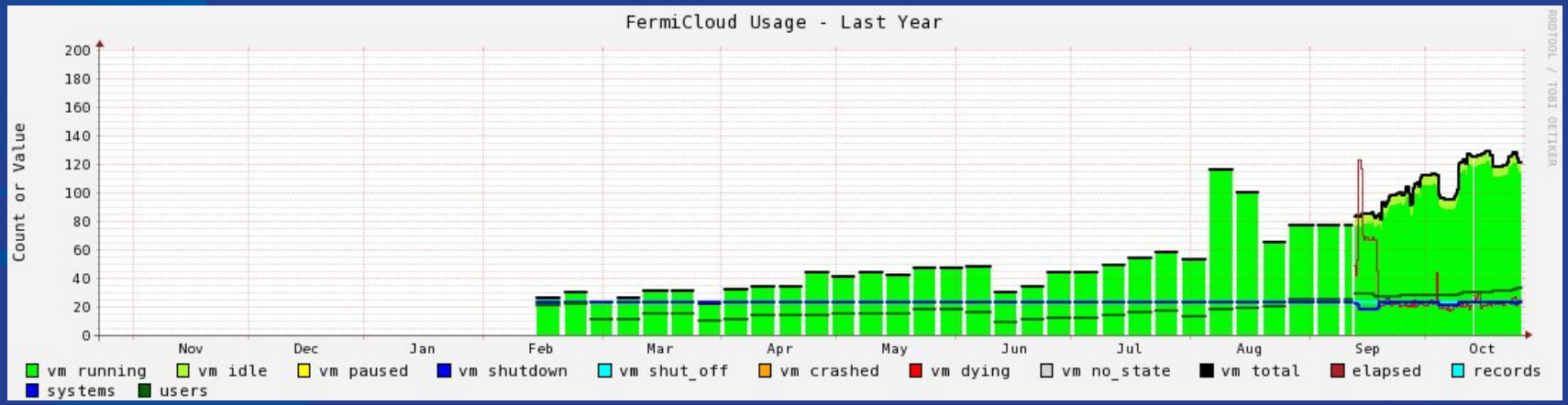
Specifications:

- Quad processor,
- 8 core,
- AMD 6128/6128HE, 2.0 GHz,
- 64 Gbytes DDR3 memory,
- 3x2 Tbytes disk,
- 4 year warranty,
- $3,654 each.

	Retire-ments	Base	Option	Purchase	Assign
CDF	200	0	16+36	52	36
CMS	--	40	4+20	64	64
D0	23	30	37	67	67
IF	--	0	40	40	0
GP	48	0	9	9	65
Total		99	111+61	271	271

FermiCloud Utilization



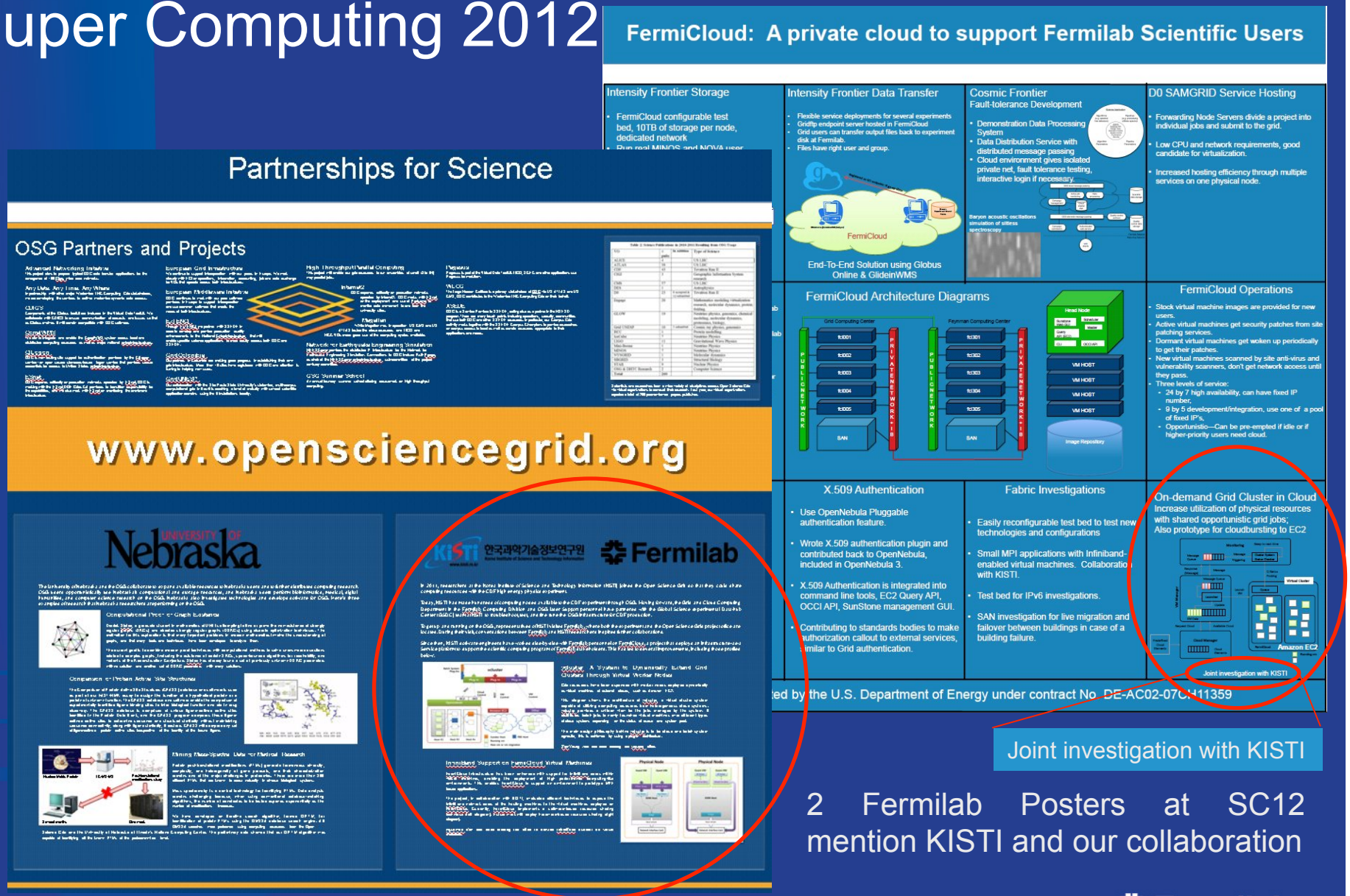


Fermilab

KISTI and GCC: working together…

- Maintaining frequent in-person visits at KISTI and FNAL
- Working side-by-side in the Grid and Cloud Computing department at FNAL
  - Seo-Young: 3 mo in the Summer 2011
  - Hyunwoo Kim: 6 months since Spring 2011
- Sharing information about Grid & Cloud computing and FNAL ITIL service management
- Consulting on operations for CDF data processing and OSG

Super Computing 2012



Joint investigation with KISTI

2 Fermilab Posters at SC12 mention KISTI and our collaboration

FermiCloud Technology Investigations

- Testing storage services with real neutrino experiment codes.
- Evaluate ceph as a FS for the image repository.
* Using Infiniband interface to create sandbox for MPI applications.
* Batch queue look-ahead to create worker node VM's on demand.
* Submission of multiple worker node VM's, grid cluster in the cloud.
* Bulk launching of VMs and interaction with private nets
- Idle VM detection and suspension, backfill with worker node VM's.
- Leverage site “network jail” for new virtual machines.
- IPv6 support.
- Testing dCache NFS4.1 support with multiple clients in the cloud.
- Interest in OpenID/SAML assertion-based authentication.
- Design a high-availability solution across buildings
- Interoperability: CERNVM, HEPiX, Glidein WMS, ExTENCI, Desktop

* In Collaboration with KISTI

- Seo-Young and Hyunwoo: Summer 2011
- Seo-Young / KISTI: Summer 2011; Proposed Fall / Winter 2011
- Hyunwoo: Proposed Fall / Winter 2011

vcluster and FermiCloud

A System to Dynamically Extend Grid Clusters Through Virtual Worker Nodes

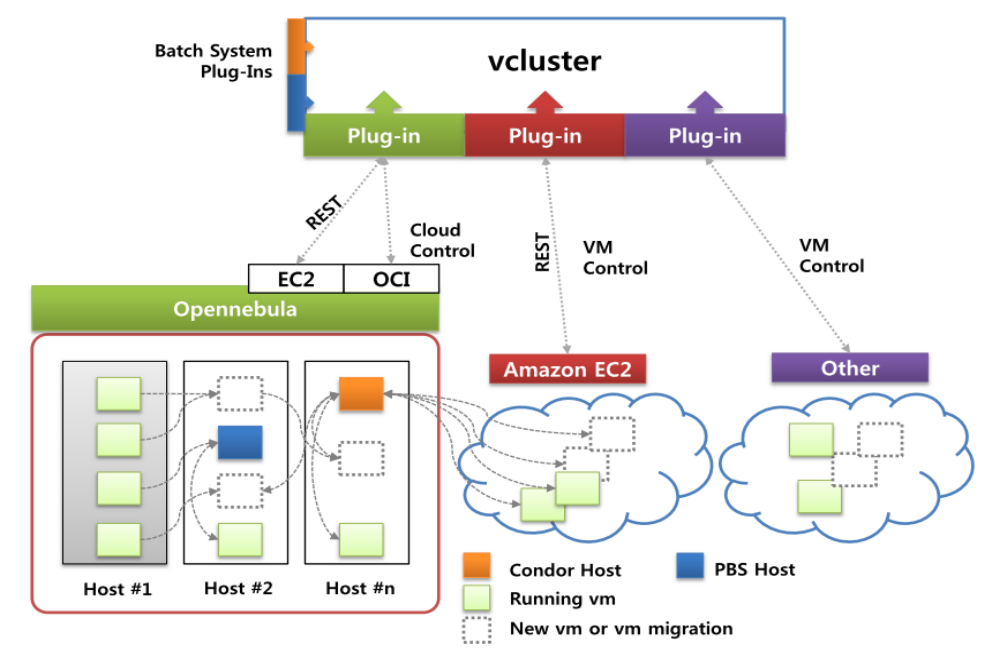
Expand Grid resources with worker nodes deployed dynamically as VM at clouds (e.g.EC2)

vcluster is a virtual cluster system that

- uses computing resources from heterogeneous cloud systems
- provides a uniform view for the jobs managed by the system
- distributes batch jobs to VM over clouds depending on the status of queue and system pool.

vcluster is cloud and batch system agnostic via a plug-in architecture.

Dr. Seo-Young Noh has been leading the vcluster effort.



InfiniBand and FermiCloud

InfiniBand Support on FermiCloud Virtual Machines

The project enabled FermiCloud support for InfiniBand cards within VM.

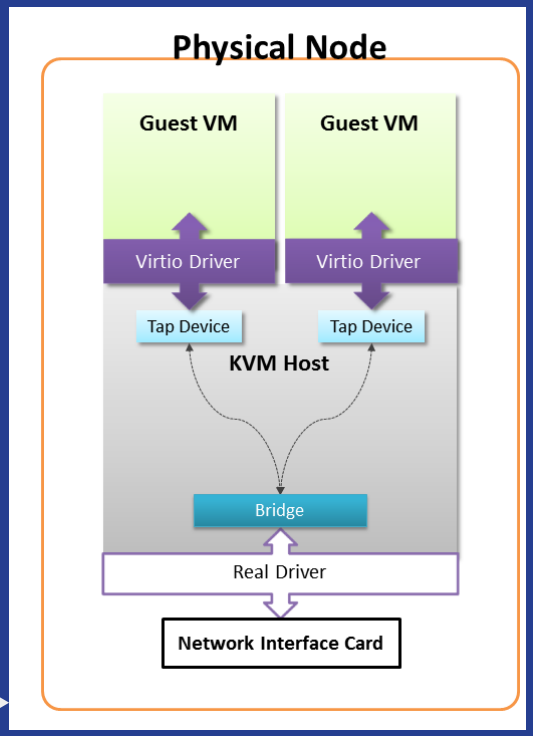
Goal: deploy high performance computing-like environments and prototype MPI-based applications.

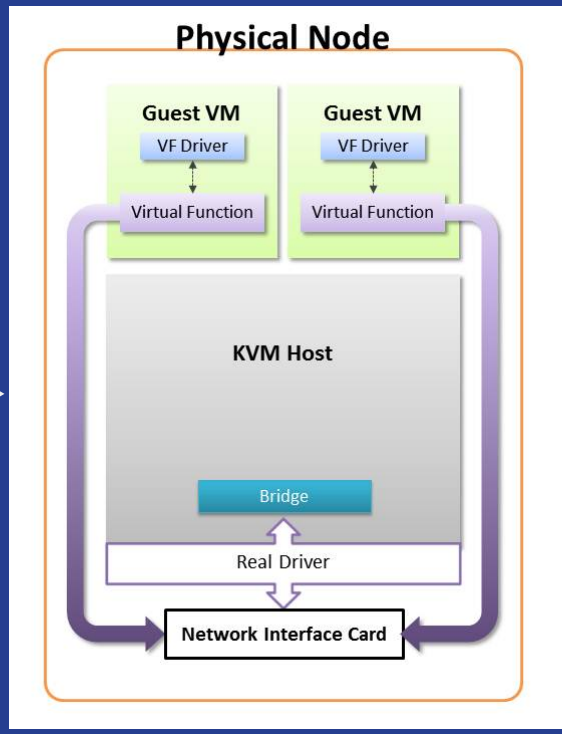
The project evaluated techniques to expose IB cards of the host to the FermiCloud VM

Approach:

- Now: software-based resource sharing
- Future: hardware-based resource sharing

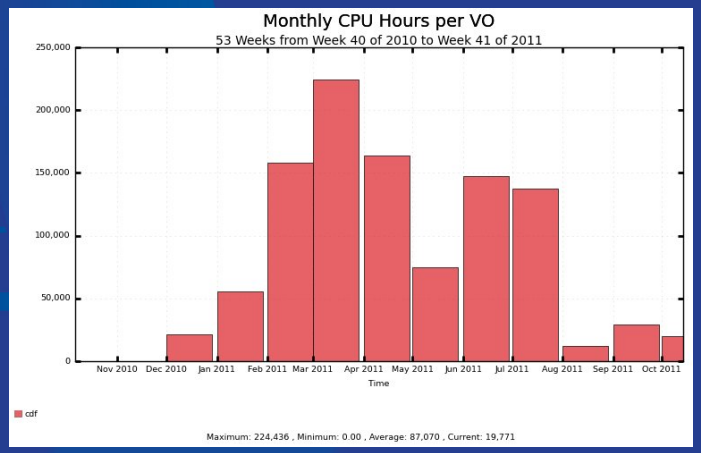
Dr. Hyunwoo Kim has been leading the effort to provide InfiniBand support on virtual machines.





KISTI and OSG

- KISTI has participated in an ambitious data reprocessing program for CDF
- Data moved via SAM and job submitted through OSG interfaces
- We look forward to further collaborations
  - Resource federation
  - Opportunistic resource usage



Fermilab

Conclusions

- Fermilab future is well planned and full of activities on all Frontiers
- The Grid and Cloud Computing department contributes with
  - FermiGrid operations
  - FermiCloud commissioning & operations
  - Distributed Offline Computing development and integration
- We are thrilled about continuing our collaboration with KISTI
  - Successful visiting program: FermiCloud enhancement & Grid knowledge exchange
  - Remote Support for CDF and OSG interfaces
  - SuperComputing 2011 Posters on Collaboration