Fermilab Site Report
Spring 2014 HEPiX Workshop

Keith Chadwick
Scientific Technical Architecture Group
Fermilab

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

Fermilab

Acknowledgements

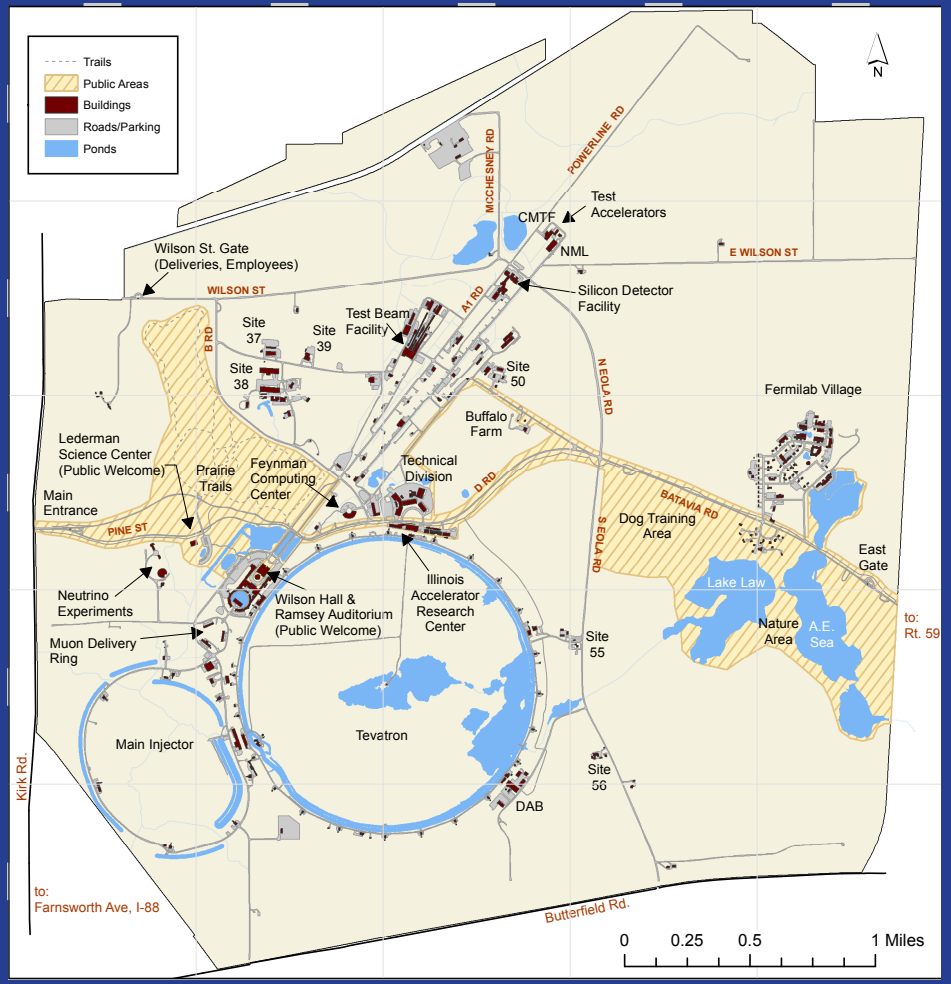
I am reporting on the accomplishments of the many individuals and organizations of the Fermilab Computing Sector,

Please credit them with all the accomplishments and assign to me any misunderstandings.

Fermilab

Some quick facts about Fermilab:

- We are located in Batavia, IL (~45 miles/~60 km west of Chicago).
- Is the premier US based HEP lab,
- Provides services to the US HEP program,
- Hosts experiments in the Energy, Intensity and Cosmic frontiers,
- Is a strong collaborator with the Open Science Grid (OSG) and the World-wide LHC Computing Grid (WLCG),
- Hosts the US Tier 1 facility for the CMS collaboration,
- Is the home of Scientific Linux.



Selected Organizational Changes since October 2013

Rob Roser appointed acting CIO and Computing Sector Head,

- In addition to serving as the Scientific Computing Division Head.

As of Monday 29-Apr-2014, I am now serving in the ITIL Service Continuity Manager and Service Availability Manager roles,

- In addition to my existing STA/EA roles…

Computing Facilities

FCC:

- The FCC3 computer rooms continue to operate without any disruptions since commissioning in December 2010,

GCC:

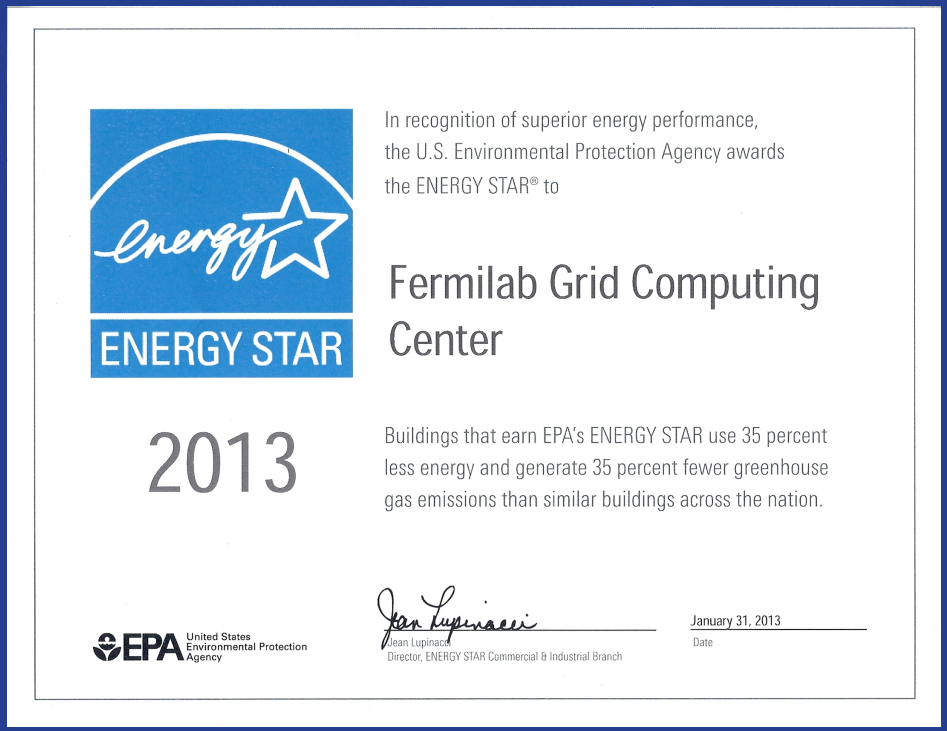
- GCC-[A,B,C] had ~8 hours of scheduled downtime,
- GCC-B and GCC-C were impacted by one "Load Shed" event in the summer of 2013,
- GCC-C also had an additional 8 hour downtime to deal with an internal electrical panel issue.

LCC:

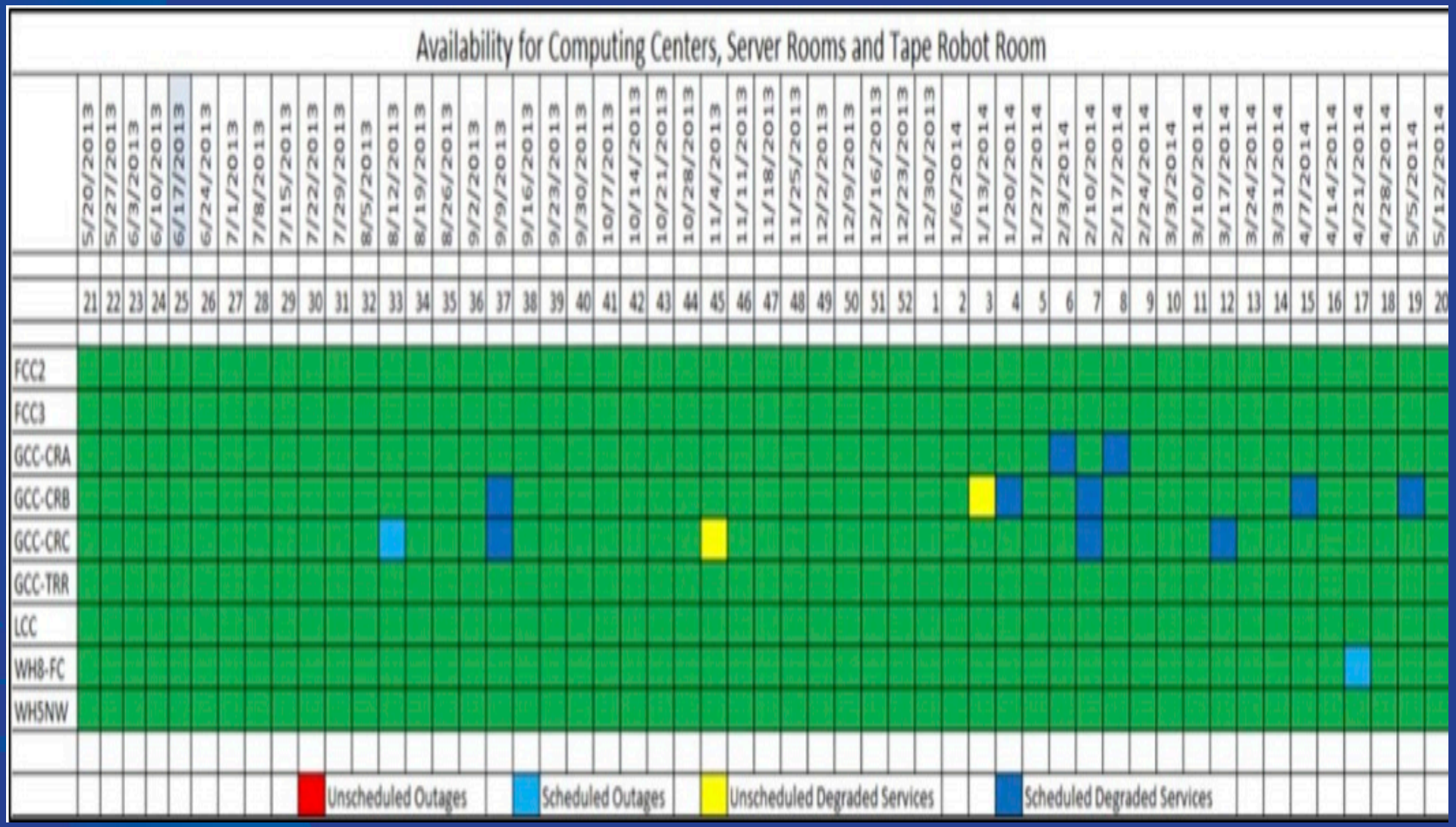
- LCC had >99.9% uptime.

The GCC computer center received an energy star certificate for 2013.

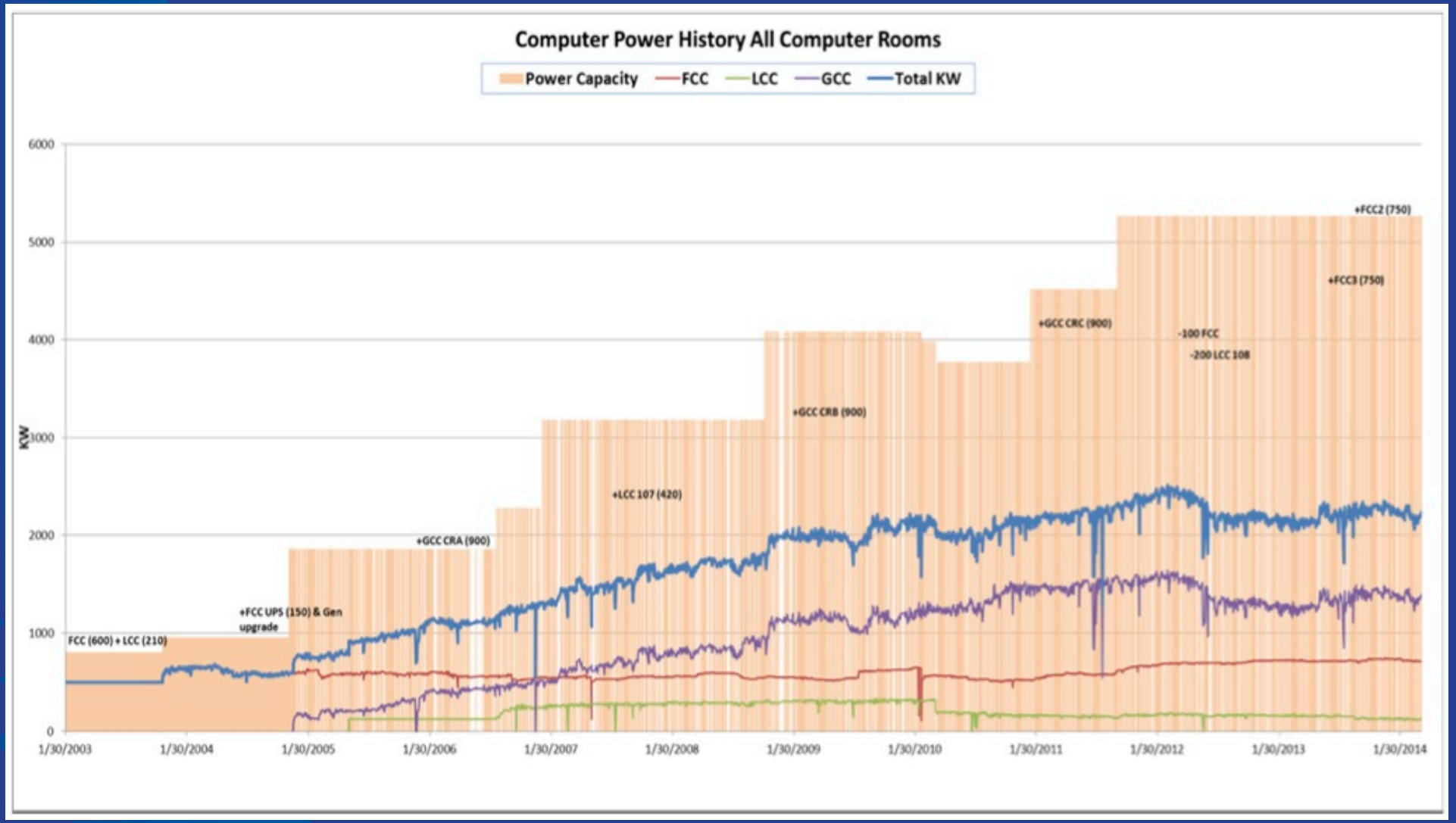
- This is the third year in a row that this building has received the award.



Computer Center Availability



Computing Center (Power) Capacity and Utilization



GCC Cooling Upgrade in Progress

Summer 2011:

- Soaker hoses on the concrete pad => load sheds.

Summer 2012:

- Berm Removal => load sheds.

Summer 2013:

- Rearrange equipment in computer rooms => 1 load shed,

Fall 2013:

- Upgrade cold aisle containment.

Facilities personnel are making ***excellent*** progress on the project to relocate the heat exchangers from the concrete pad behind the building to on top of the building:

- Moving ~4 heat exchangers/week (2 per computer room),
- Work should be completed by early June.

Goal for 2014 – "No load sheds":

- Will update the load shed process just in case,
- Time will tell if the engineering was successful…

2011:

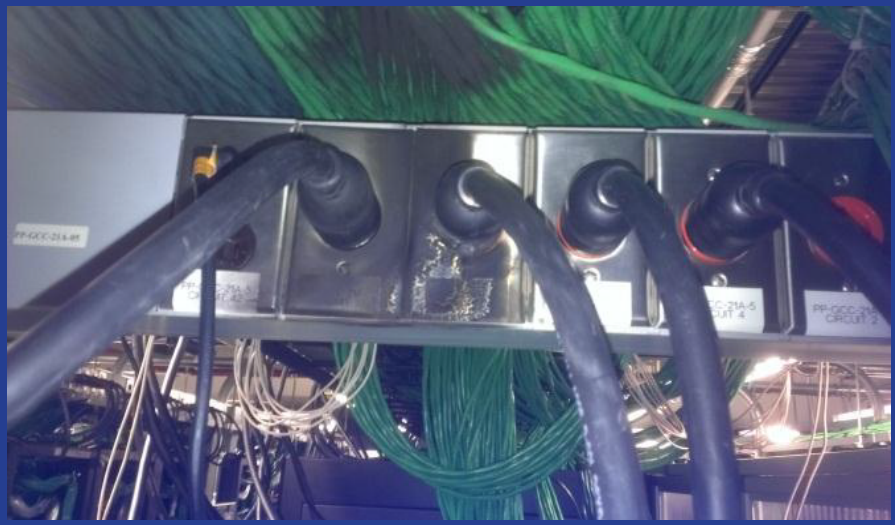
15-May-2014:

Incidents with PDU Cords

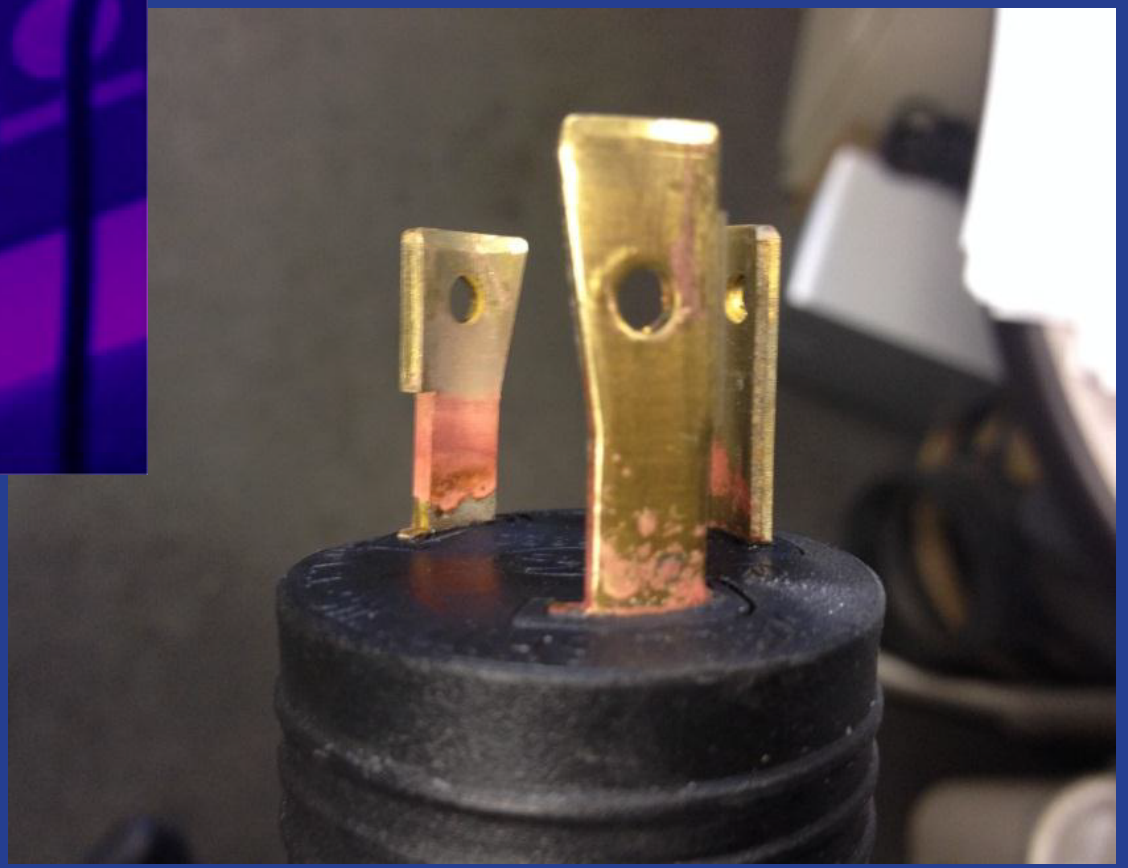
Recently, the Data Center Facilities personnel identified a pattern of incidents with both L5-30 120 VAC and L6-30 208 VAC PDU cords/receptacles.

- On Thursday, January 16, 2014 shortly after 6:00PM, IT monitoring software reported four of sixteen computers in Rack GCC-CRB-3041 were not responding.
- Security camera footage later revealed a small flash event originating from the PDU (Power Distribution Unit) plugs at this time. The camera footage also revealed a second flash event at 6:40PM.
- This event caused the ceiling VESDA system to alarm and the Fire Department was dispatched to GCC-B.
- The Fire Department responded and found no reason for the alarm and the system was reset.
- The next day at 9:00AM, the incident (four servers off) was reported to Computing Sector Facility Operations.

A total of six events occurred between 16-Jan-2014 and 20-Feb-2014.



Thermographic Imaging is providing an early "in situ" warning



Other Factors & Recommendations

The affected circuits were operating within the 80% capacity guidelines,

- E.g. - No more than 12A of load on a 15A branch circuit.

The cords are not associated with a specific PDU manufacturer (we use PDUs made by different manufacturers for 120 VAC and 208 VAC),

- The PDUs and cords are UL listed,
- There are some indications that the cords may share a common subcomponent manufacturer.

Typical server installations have been operating in GCC since 2005, with server refreshes on the average of 4-7 years,

- There have been no previously reported events with PDU cords/plugs,
- The PDU cords in question have been in service for at least two equipment "refresh" cycles,
- Recommendation to replace PDUs (and cords) every two equipment refresh cycles.

Periodic thermographic scanning is essential to provide early warnings prior to PDU cord failure,

- Once a cord is identified as showing signs of impending failure, then the plug on the PDU cord ***AND*** the receptacle must be replaced,
- Perform visual inspection of the plug any time it is removed from the receptacle.

ISO 20K / ITIL Service Management

Formal ISO 20K Audit in 2013 was successful,

- Also performing self audits across the various ITIL roles.

The Fermilab Service-Now installation has been upgraded to the "Dublin" release,

- Upgrade took longer than expected, since we had skipped the previous release.

Additional Scientific Services are in the process of being "onboarded" to full ISO 20K,

- Upgrading from Incident, Problem and Change/ Release (aka "ITIL-Lite").

Existing SLAs and OLAs are being reviewed and updated,

- Part of our "self audit" responsibilities.

Individual Technical Statements of Work (TSWs) are being established between the Computing Sector and Experiment Collaborations.

- Experiment specific SLA needs are being documented as part of this process.

ITIL Roles and Personnel:

Fermilab IT Service Management

Network

The offsite network links were moved from the ESnet CHIMAN 10GE network to the new ESnet ChiExpress 100GE infrastructure on Friday 9-May-2014.

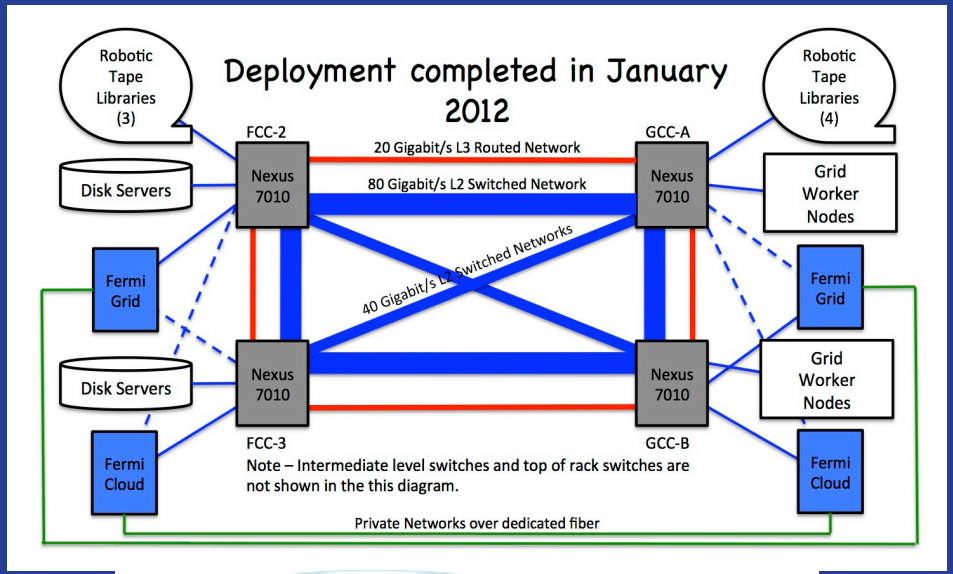
The networking staff has conducted "near real-life" failover tests of the redundant network infrastructure,

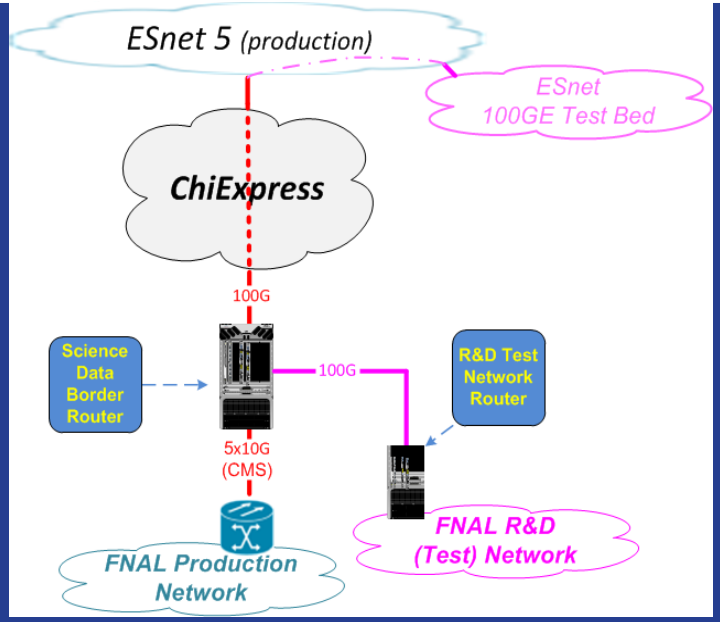
- The test evaluation was "successful with issues",
- Several items/areas were identified for improvement.

Network registration has been moved from an old stand-alone web application into a Service-Now based workflow.

The reorganization of the IPv6 and the 100 Gb/s testbed continues,

- Slow progress since enhancements and maintenance of the production network takes priority.





Authentication and Identity Management

Fermilab has joined Eduroam!

SHA2 on Fermilab KCA,

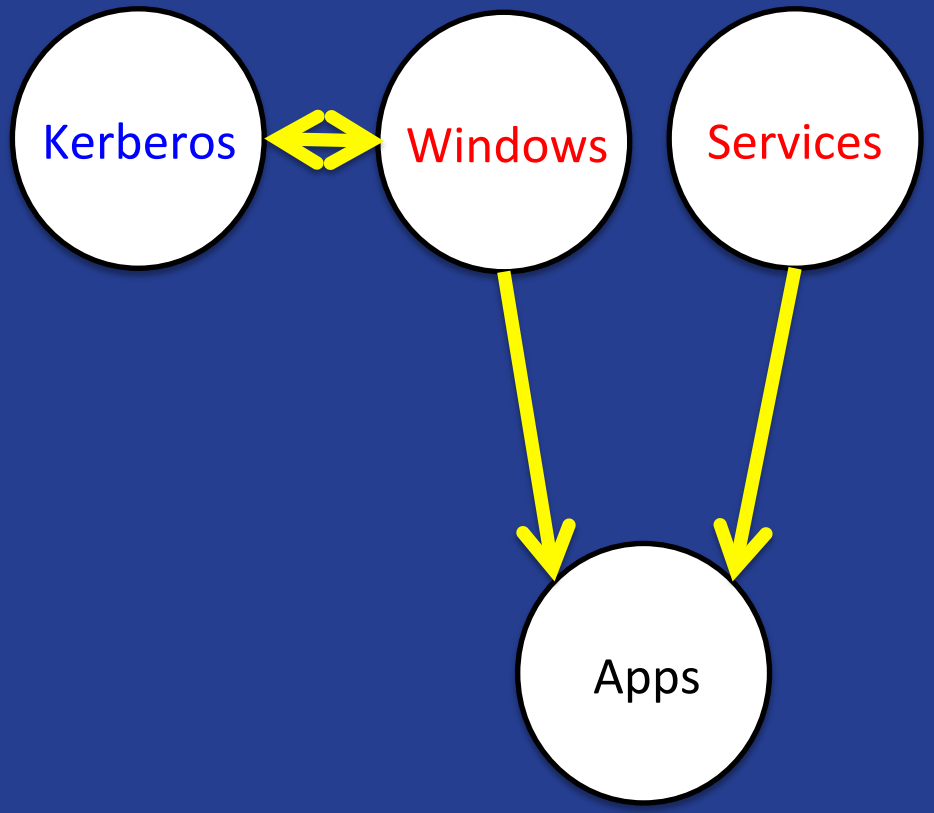
- Waiting for release of Heimdal 1.6

Deployment of new "Apps" domain is underway,

- One way trust of Windows and Services domains,
- No accounts are in Apps domain.

COTS password reset tool deployed for the Windows and Services domains,

- Tool does not support Kerberos domain.



HPC & Lattice QCD

Significant number of LQCD users using Globus Online.

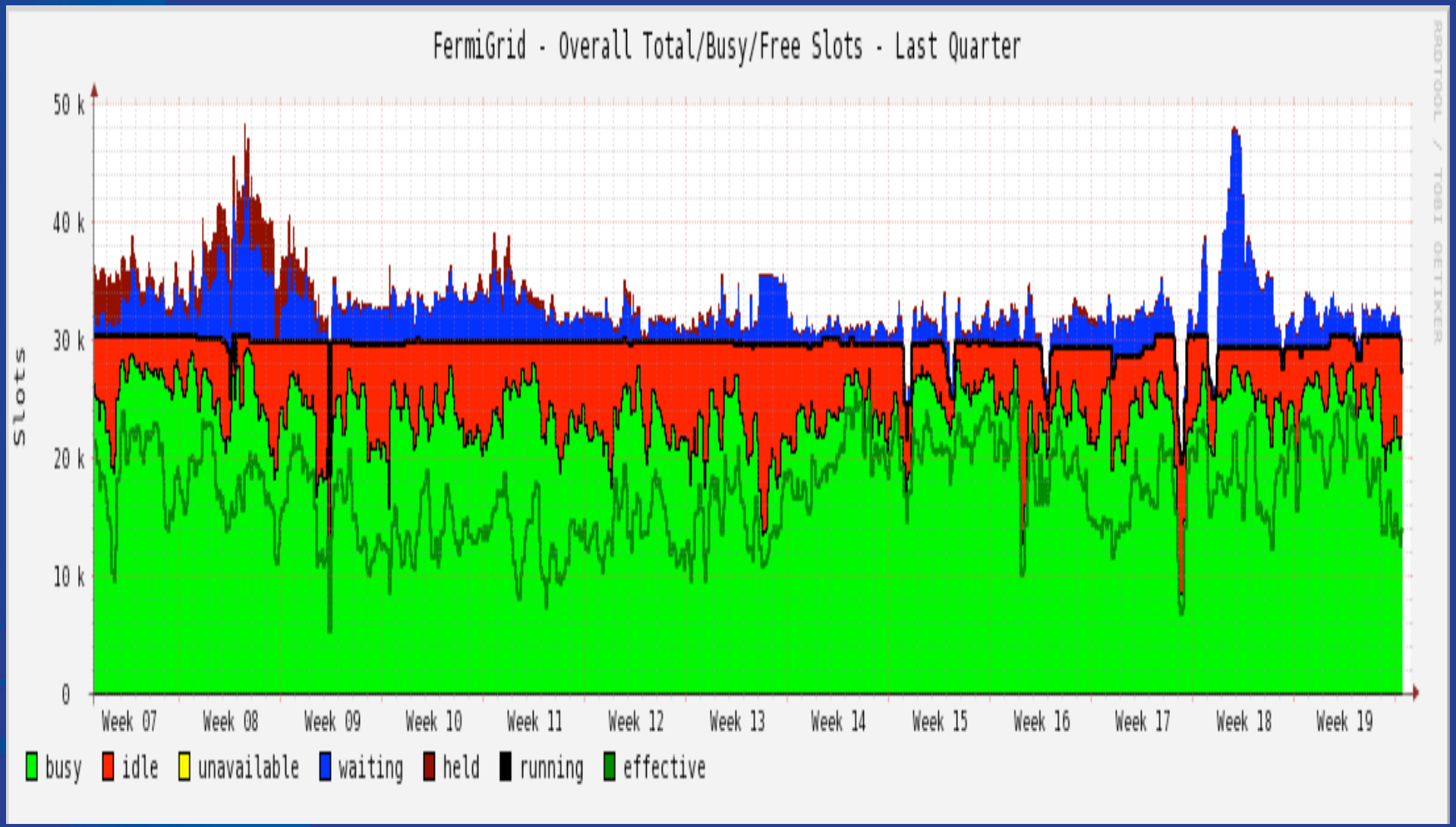
J/psi cluster turned off on 19-May-2014.

"Hybrid" HCP Cluster Acquisition currently underway:

- ≥ 10 TF sustained performance Conventional,
- ≥ 29 TF sustained performance GPU.

Vendor responses are in the process of being evaluated on a "best value" basis.

Fermilab Campus Grid Utilization



Current Fermilab Campus Grid Statistics (as of May 2014)

Cluster(s)	Batch System	Job Slots	Raw Occupancy	Effective Utilization
CDF	HTCondor	5,088	73.1	60.0
CMS T1	HTCondor	11,807	90.3	71.0
D0 (Merged)	PBS	5,720	60.8	35.5
GP Grid	HTCondor	7,672	81.0	60.3
———		———	———	———
Overall-Today		30,257	79.1	59.4
Fall 2013 HEPiX		24,530	79.4	62.8

FermiGrid Service Availability "Tiles"

Automatically generated from the service availability data already being collected by the FermiGrid Service Monitoring.

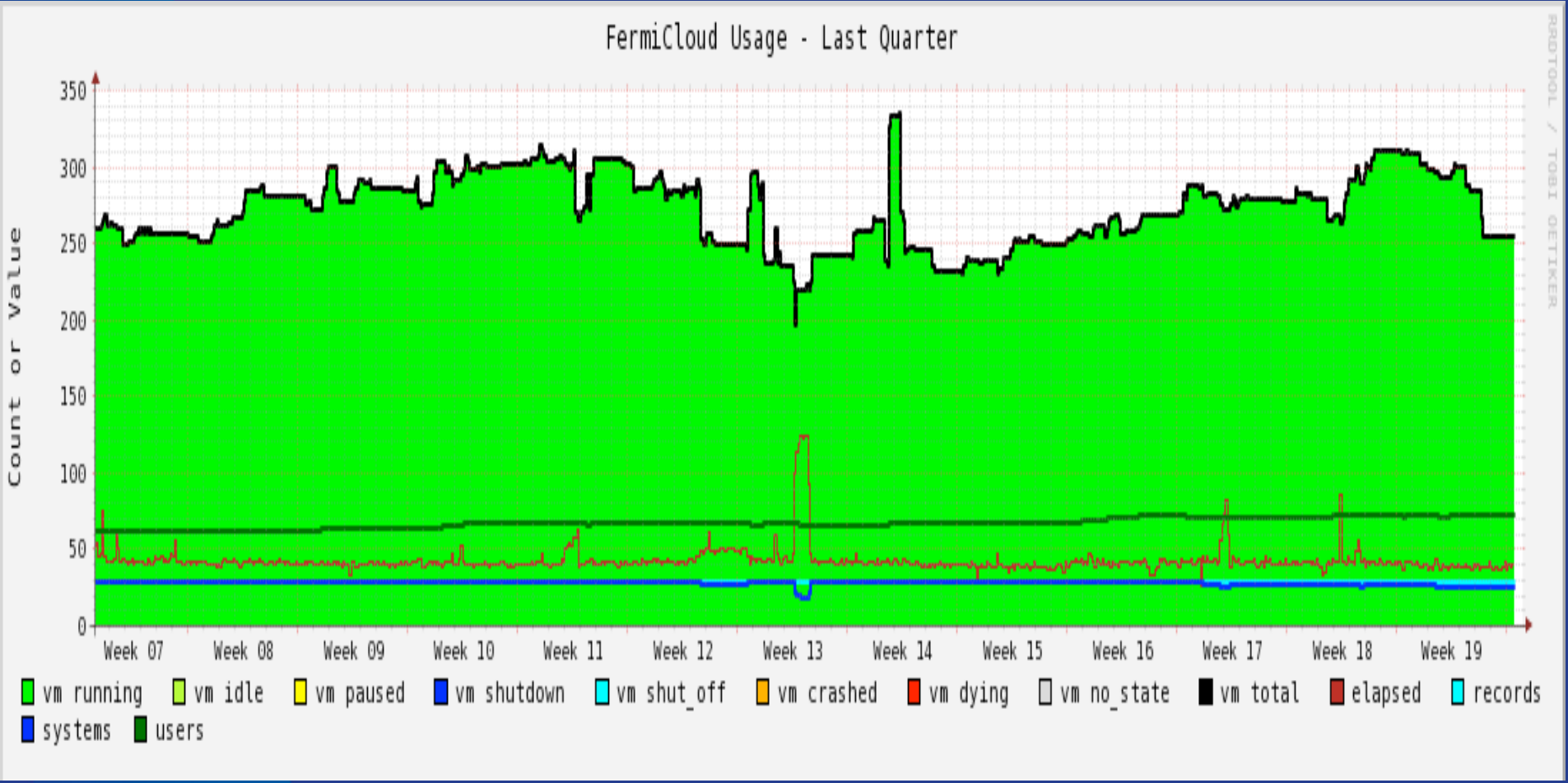
Daily Availability – Past 30 Days:

- Summary Report by Service Category
- Detail Report by Individual Service

Weekly Availability – Past 52 Weeks:

- Summary Report by Service Category
- Detail Report by Individual Service

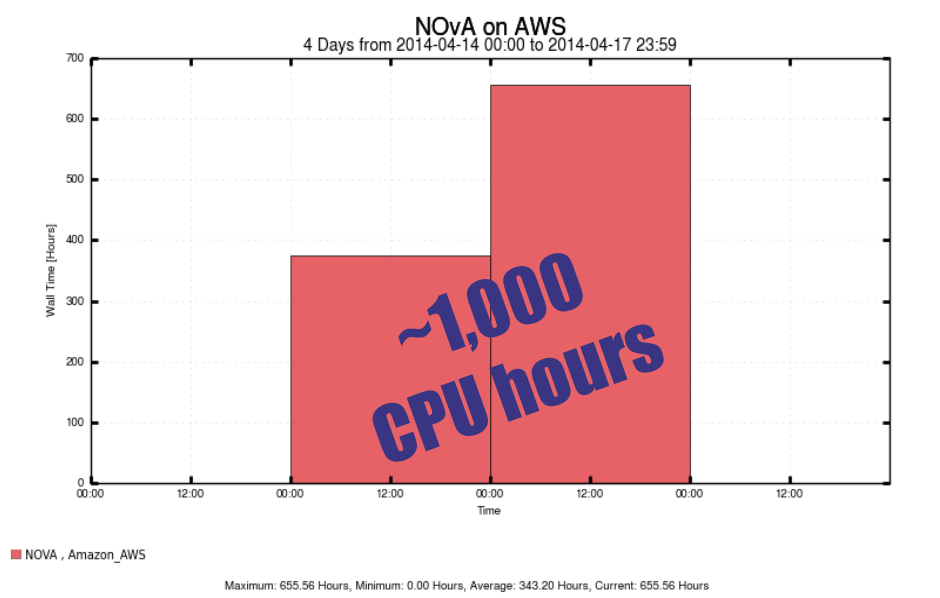
FermiCloud Utilization



See Steve Timm's talk...

NOvA Monte-Carlo via GlideinWMS on AWS

Number of Jobs	1088
Success	1047
Failure	41
File upload	1
File download	1
NOvA executable	33
Art non-0 exit code	1
Geant4 failure	1
Hang, external kill required	1
Average execution time	~56 minutes
Overhead	~6 minutes/job
Output data	~326 GB
Experimental data	~152 GB
Log files	~170 GB



Maximum # of VMs: 100

CPU cost: $86 (blended price of ~ $0.066 per CPU hour including overheads)

I/O cost: $39

Total cost: $125

Scientific Linux March of Progress

RHEL 5 -> Scientific Linux 5

- 5.0, 5.1, 5.2, 5.3, 5.4, 5.5,
- 5.6, 5.7, 5.8, 5.9, 5.10.

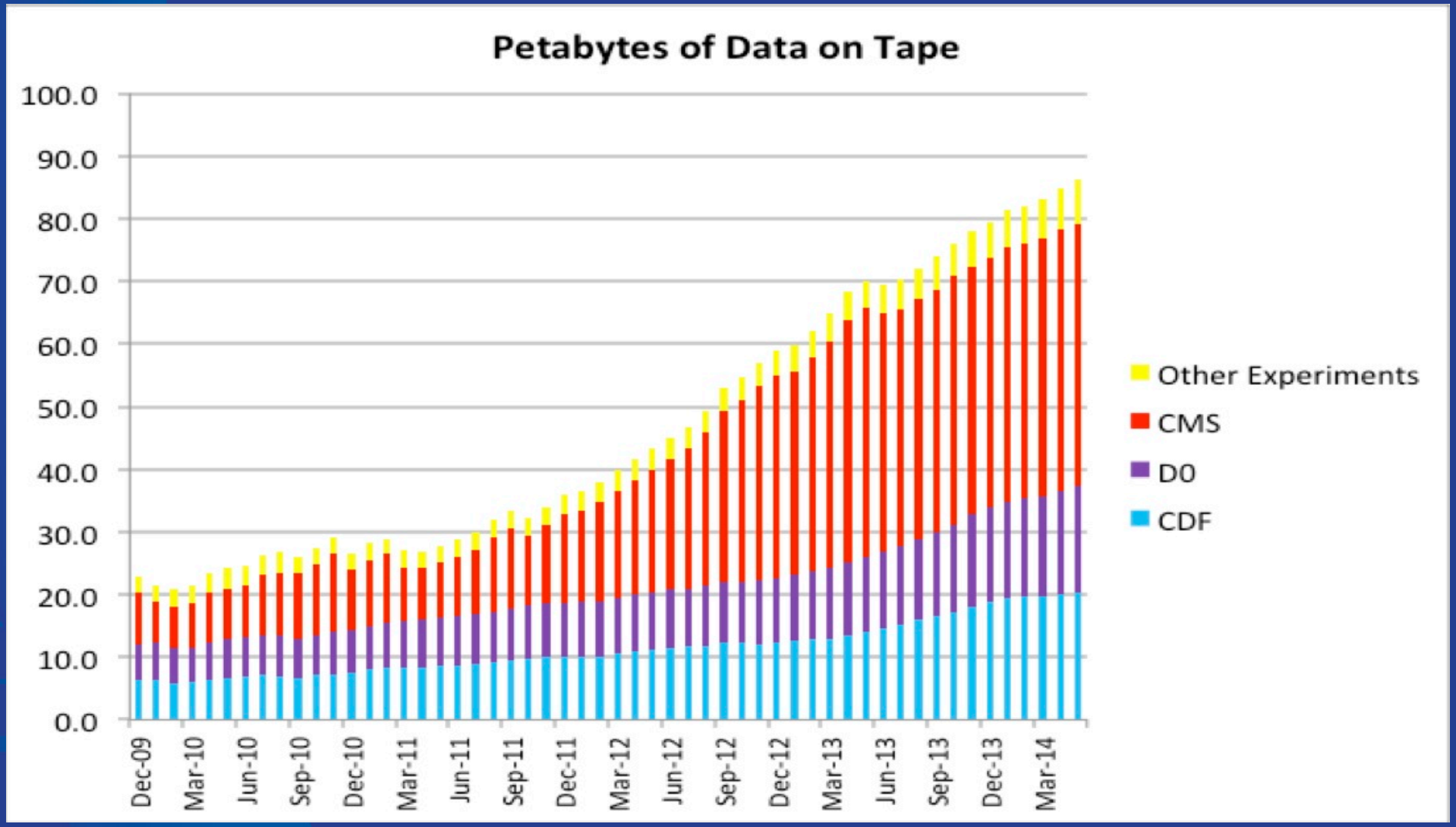
RHEL 6 -> Scientific Linux 6

- 6.0, 6.1, 6.2, 6.3, 6.4, 6.5, ...

RHEL 7 -> Scientific Linux 7

- See Connie Sieh's SL talk...

Enstore - Data on Tape



dCache Storage

FY13 Purchase of ~6.5 PB has been deployed:

- Nexsan E60 with 60 x 4 TB drives (raw),
- CMS deployed ~2 PB to to EOS,
- "Public" dCache (used by Intensity Frontier) upgraded from 0.18 PB to ~4 PB
  - Intensity Frontier experiments moving data access from BlueArc to dCache

CMS split their ~15 PB of existing dCache:

- ~13 PB non-tape backed,
- ~2 PB tape backed.

Run II:

- CDF dCache at 1.6 PB,
- D0 SAM cache at 1 PB,
- As part of their long term preservation strategy, D0 is looking into moving to dCache.

Altogether a total of approximately 24 PB disk cache.

Notable Security Events

Windows XP End-of-Life,

- The number of Windows XP systems has dramatically trended down since January 2014 (currently ~80 systems),
- Windows XP systems have off site web access blocked via site web proxy rules since the official Microsoft EOL date.

Mac OS X 10.6 "Snow Leopard" End-of-Life,

- Warning notices began 1-May-2014, users had until 21-May-2014 to retire or upgrade to a supported OSX version (either 10.7 "Lion", 10.8 "Mountain Lion" or 10.9 "Mavericks"),
- As of 21-May-2014 any remaining Mac OS X 10.6 systems will have their web access blocked via site web proxy rules.

Heartbleed,

- Surprisingly small number of vulnerable systems,
- All vulnerable systems have been patched (and rekeyed where appropriate).

Other

Human Capital Management Upgrade - Replace various home grown solutions with Workday Cloud Service,

- Lots of effort…
- Will shortly be going live.

Kronos (Time Reporting) Upgrade -> latest release of Kronos.

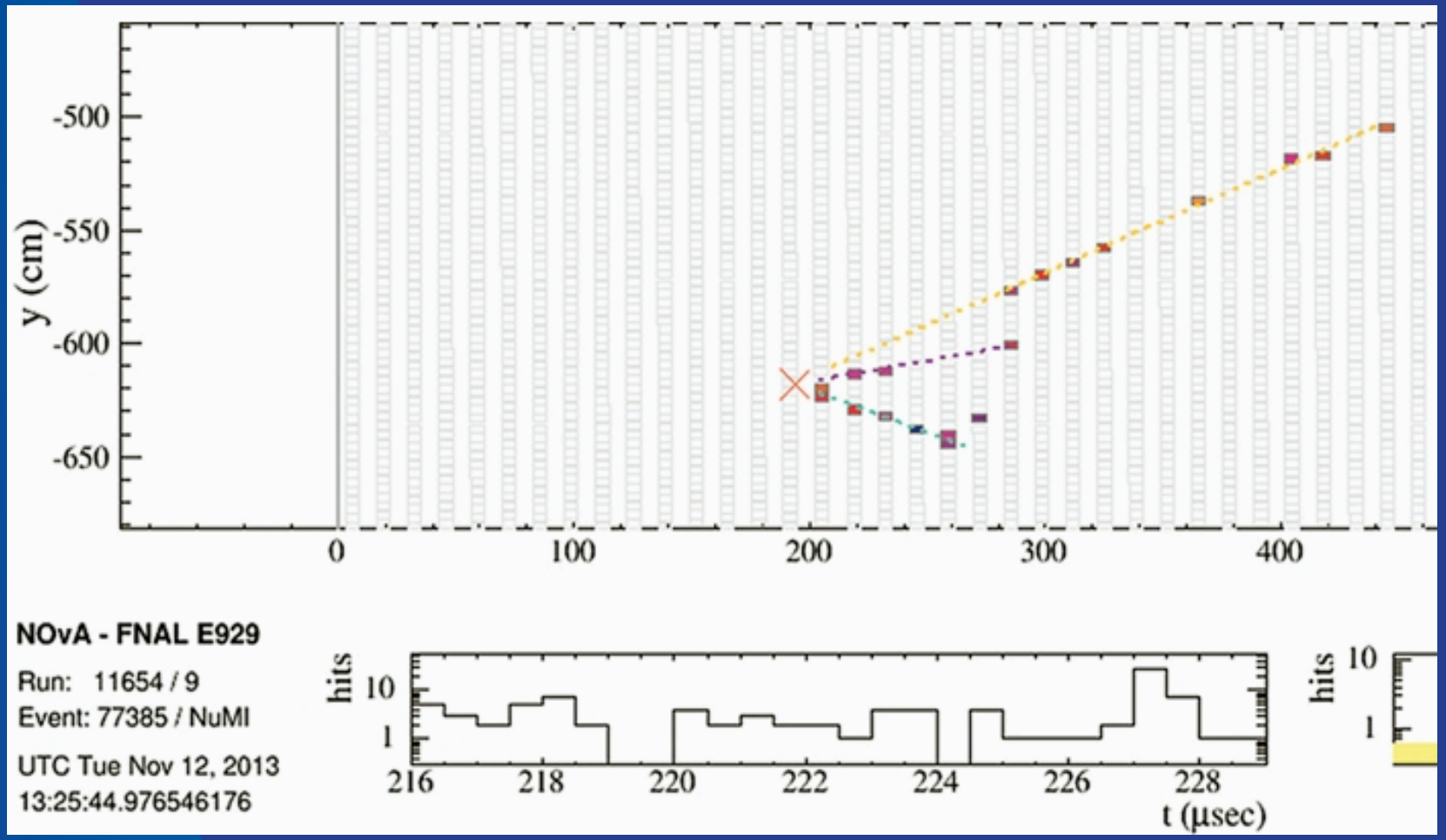
SharePoint 2010 -> SharePoint 2013 (aka "fermipoint.fnal.gov"),

- SP2010 Meeting Workspaces being converted into SP2013 Meeting Apps,
- Migration started on 14-Mar-2014 with mysites.fnal.gov and web.fnal.gov,
- Migration of CCD department sharepoint pages uncovered various issues with domain trusts,
- Migration paused for ~six weeks while issues with domain trusts & ADFS were sorted out,
- Migration resumed on 12-May-2014,
- Current target for Computing Sector (CCD+SCD) migration completion is 30-Jun-2014.

DOE has announced that they will end funding for ReadyTalk teleconferencing at the end of the current fiscal year (30-Sep-2014),

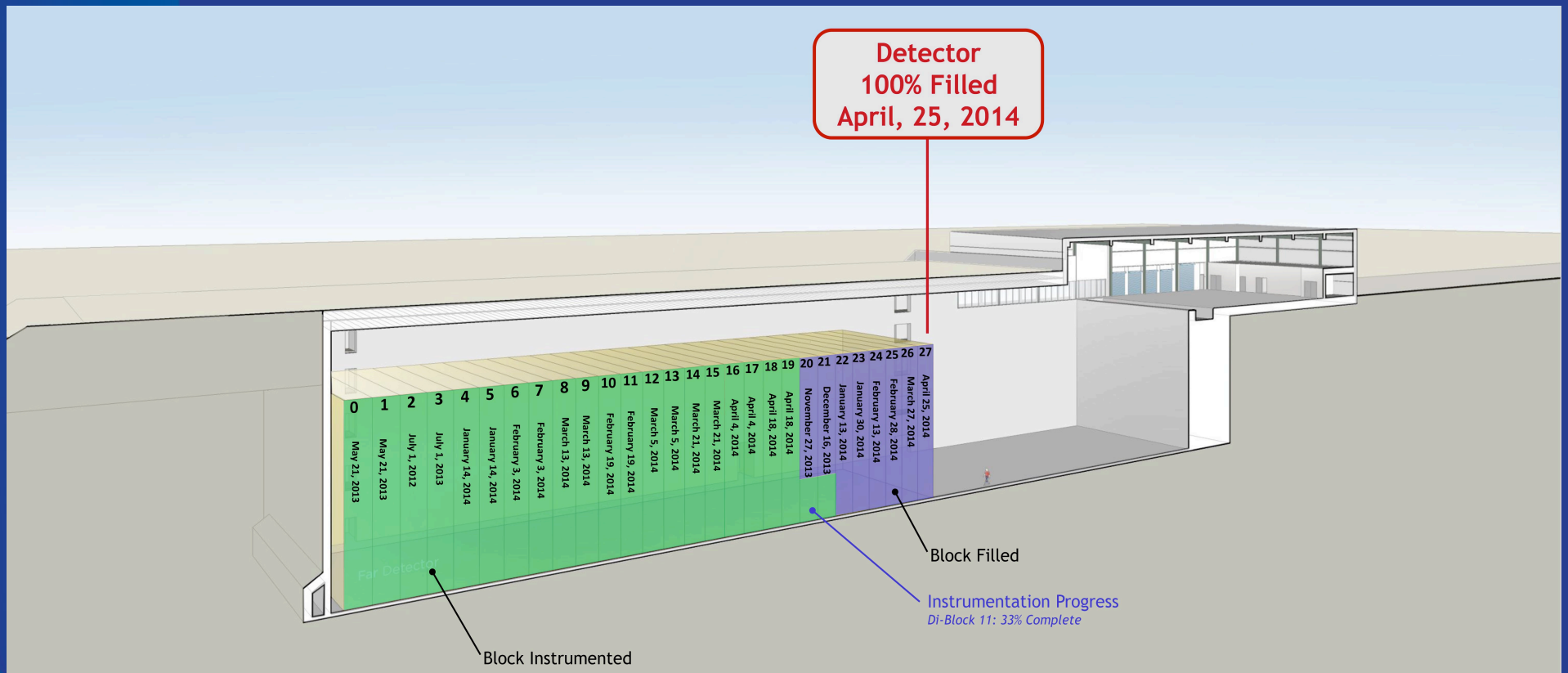
- Fermilab is currently reviewing the potential options…

NOvA Observes First Neutrino November 12, 2013



http://www.symmetrymagazine.org/article/april-2014/no%CE%BDas-first-neutrino

NOvA Assembly Progress



14 kilotons = 28 NOvA Blocks

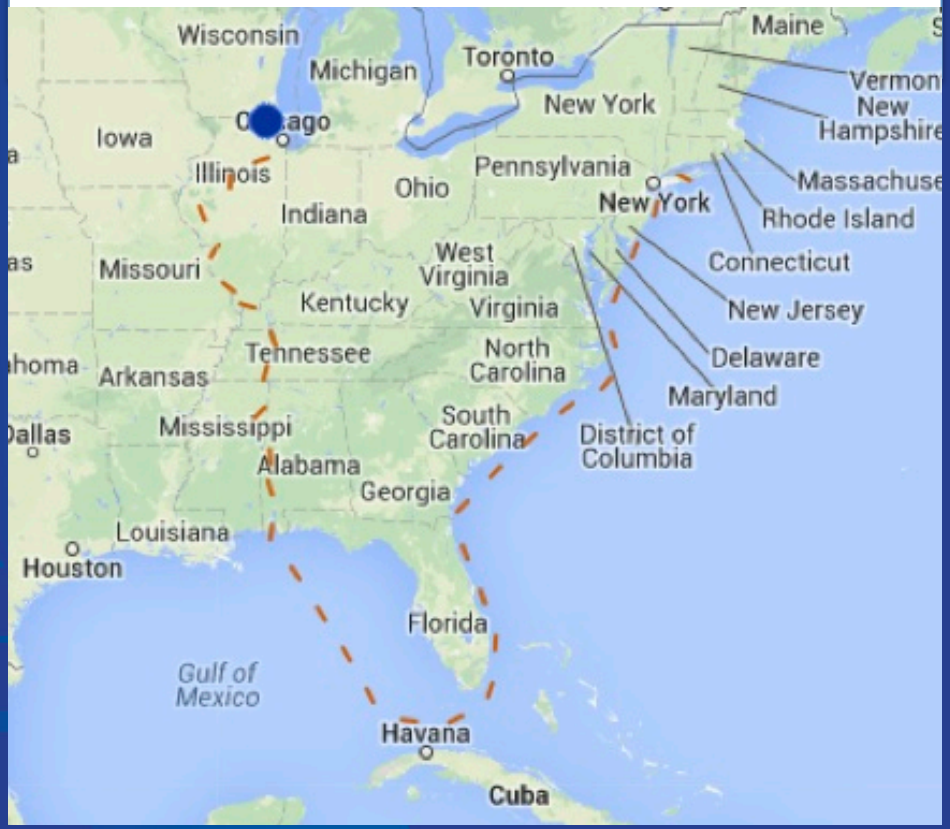
28 blocks of PVC modules are assembled and installed in place

28.0 blocks are filled with liquid scintillator

20.66 blocks are outfitted with electronics

Muon g-2

Move of the magnet from BNL to FNAL during June/July of 2013:



Beneficial occupancy of the new building at Fermilab for the Muon g-2 experiment was granted on 10-Apr-2014.

Summary

It is an exciting time to be supporting science at Fermilab across all of the frontiers of science:

- Energy Frontier
- Intensity Frontier
- Cosmic Frontier

And there is no danger of running out of work supporting and extending the computer systems and services that provide the foundations of supporting science.

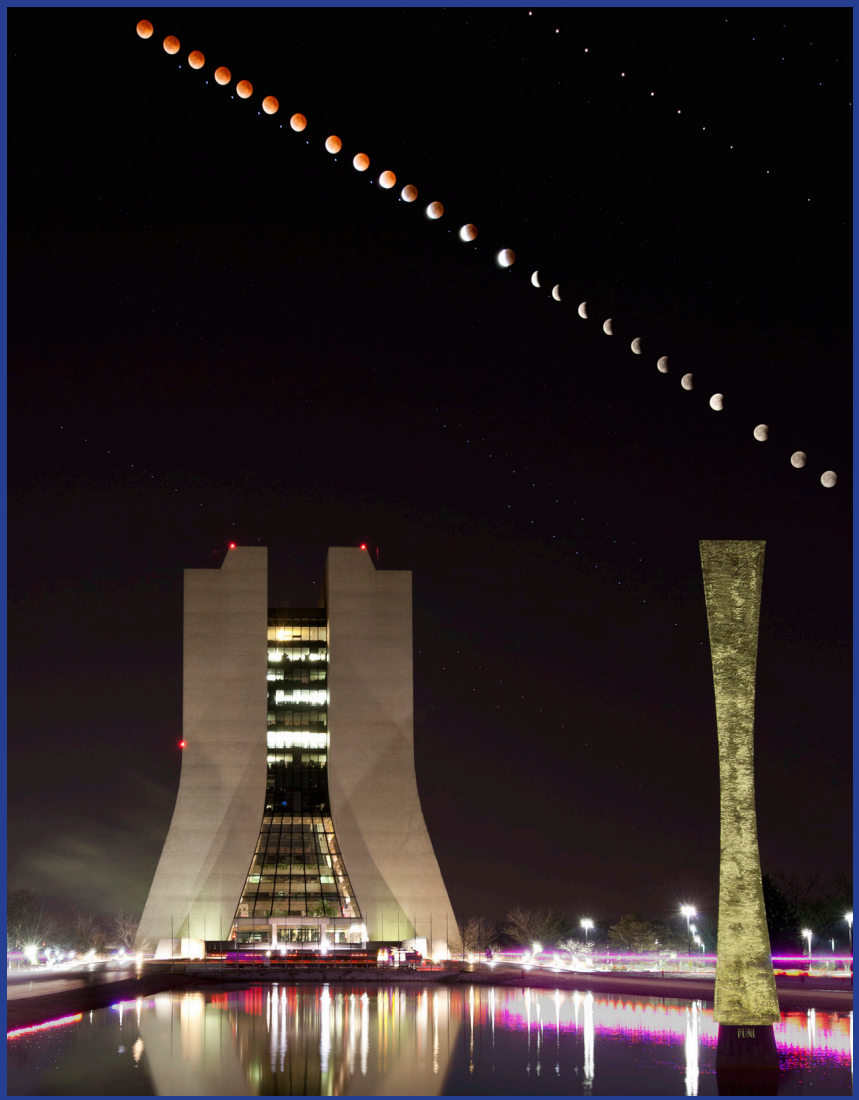
Lunar Eclipse ("Blood Moon")

This picture of the lunar eclipse on the (early) morning of 15-Apr-2014 is composed of many exposures of differing lengths [taken between 2:45 and 4:45 a.m. by Marty Murphy, AD]

The bright star near the moon is Spica. The planet Mars is the brightest dot across the top right corner.

Because the exposures are of different lengths, some stars are visible throughout, and some are not.

As the moon starts to brighten, nearby stars appear to diminish in its light.



Thank You!

Any Questions?