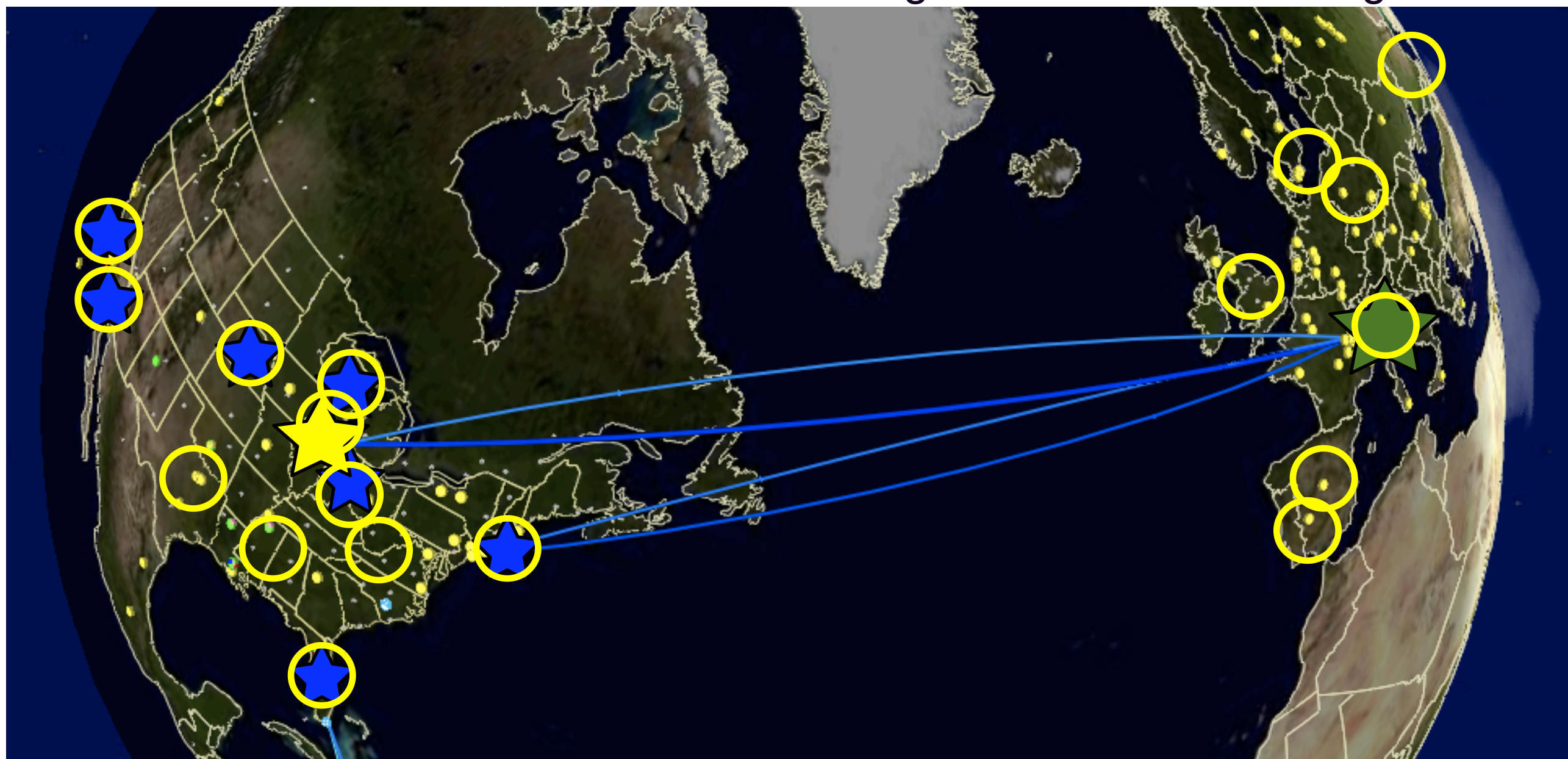# CMS Status and Needs

Open Science Grid Council Meeting

L.A.T.Bauerdick/Fermilab

Worldwide LHC Computing Grid

★ OSG connected to Grids in Europe and Asia

★ U.S. CMS: Tier-1 at Fermilab, 7 Tier-2 centers, ~9 Tier-3 centers

★ Grid connectivity to access world-wide distributed data at 7 Tier-1 centers

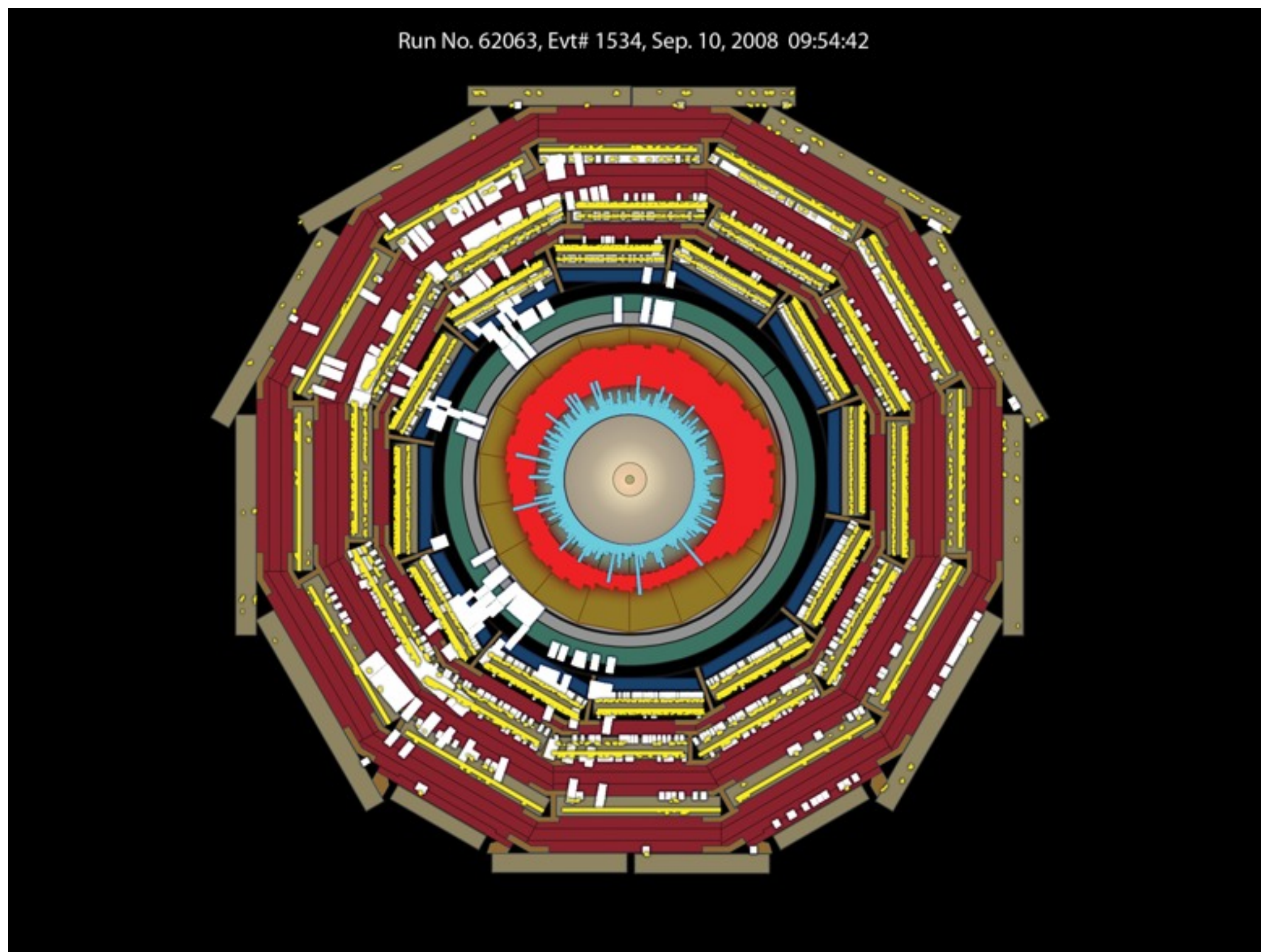★ Put to real-world tests with Data Challenges and first data taking

First Beam-induced Events in CMS
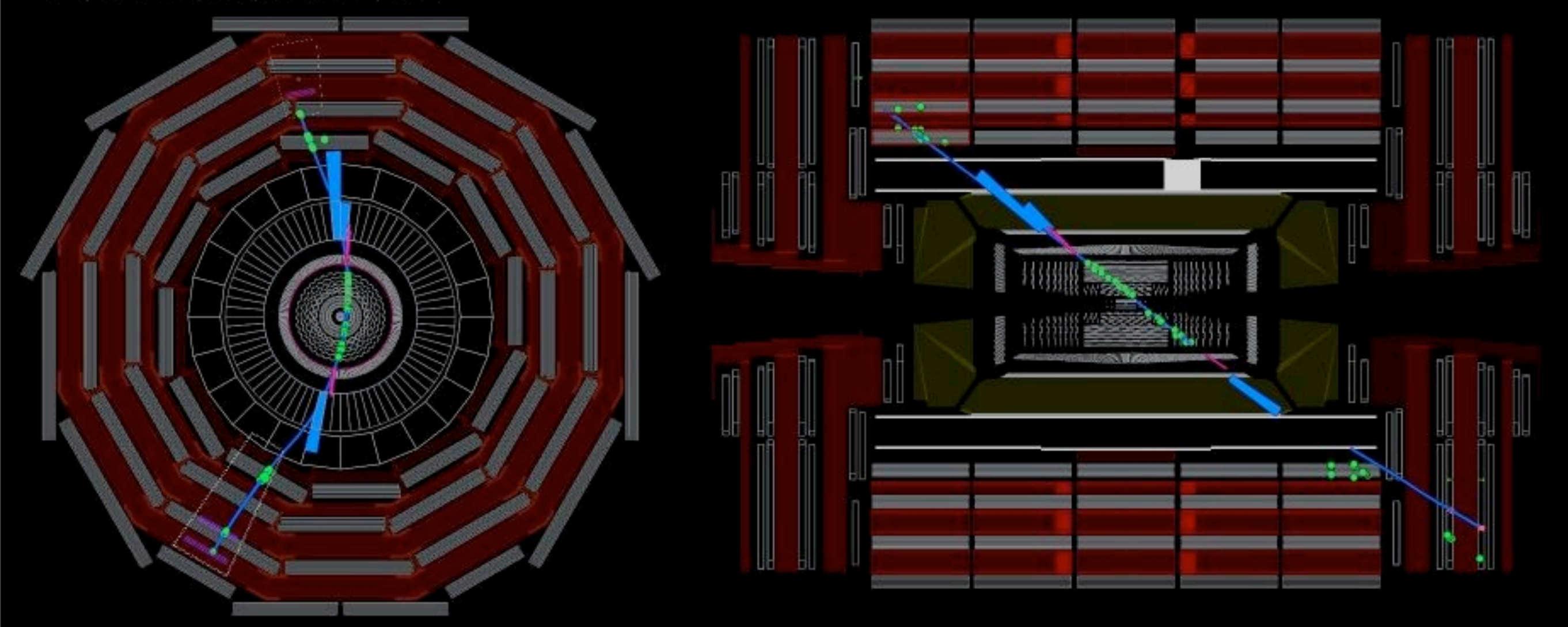
~ 2.10⁹ protons on collimator ~150 m upstream of CMS

**Sept. 10, 2008
Beam Splash**

Run No. 62063, Evt# 1534, Sep. 10, 2008 09:54:42
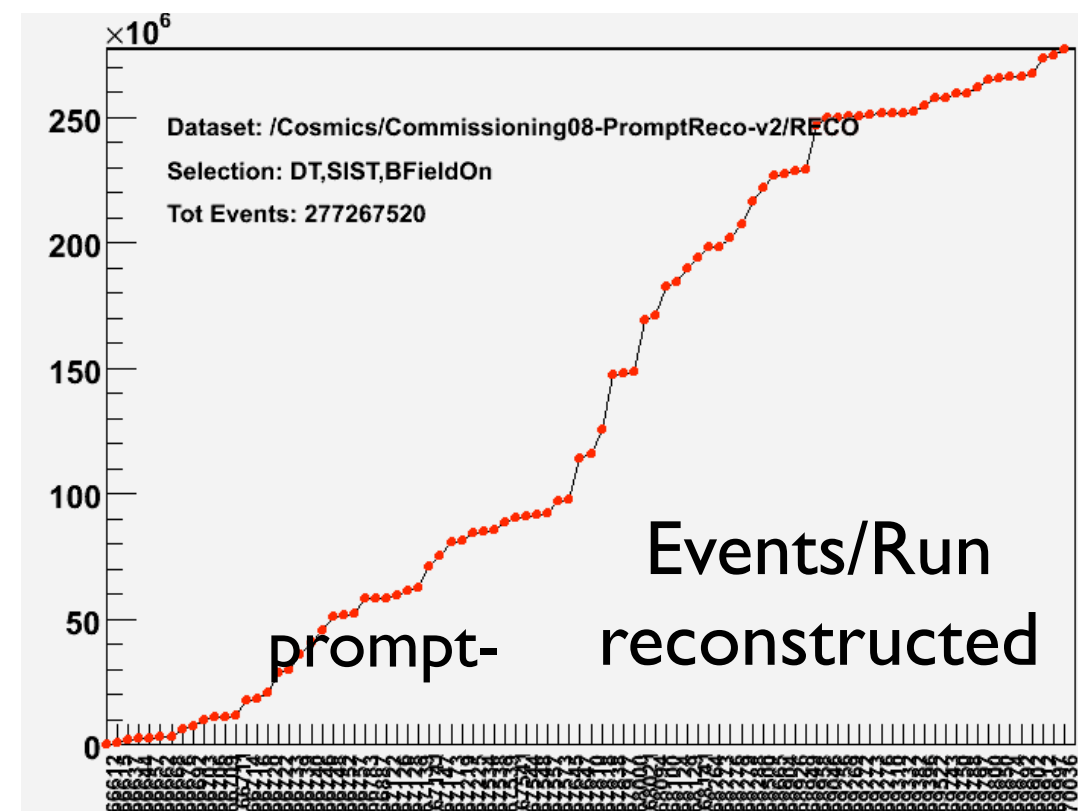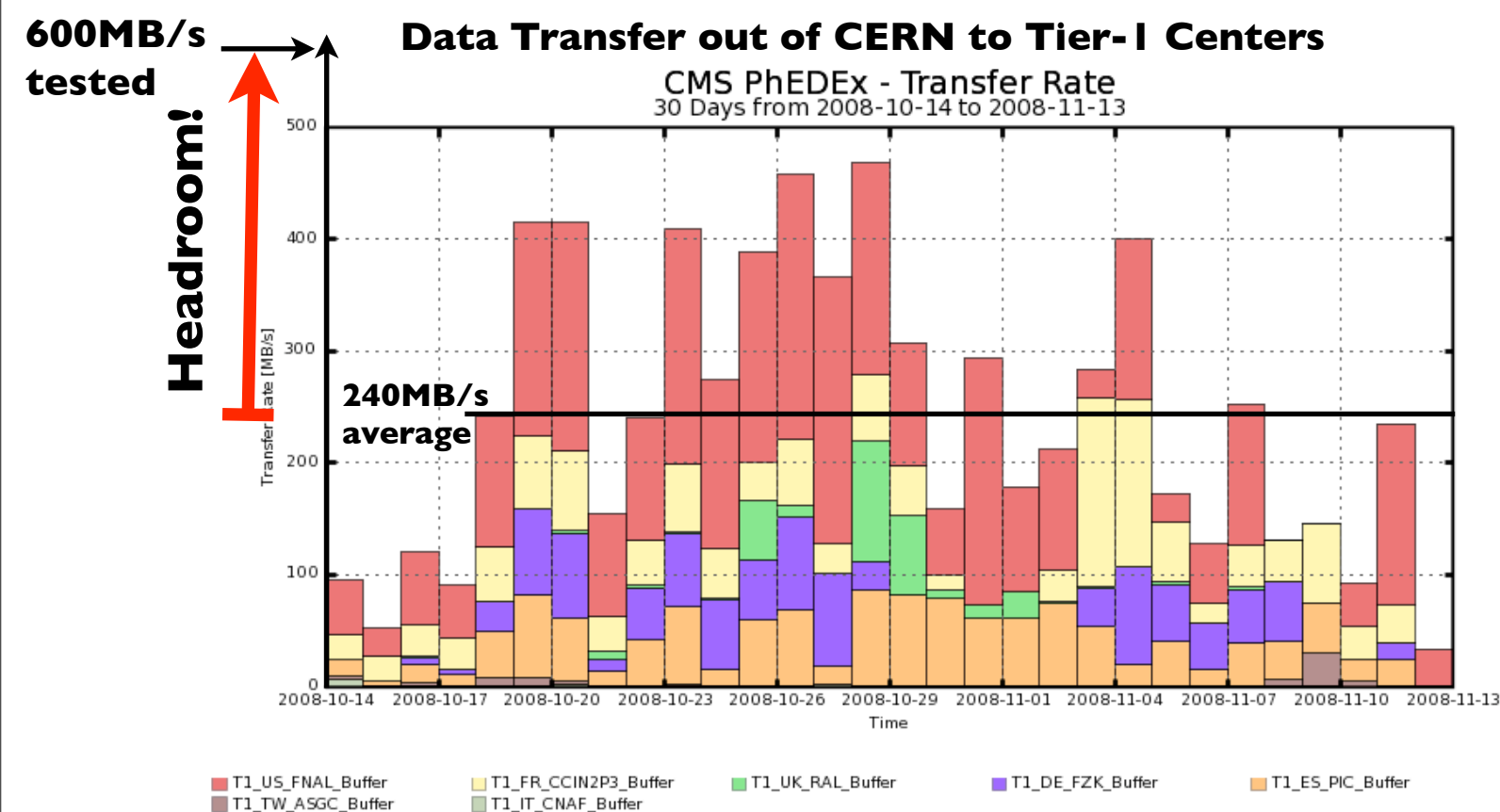
Run 66748, Event 8900172, LS 160, Orbit 167345832, BX 2011

✦ Throughout 2008, S&C systems getting successively ready for data

★ data/computing challenges, global runs, first beam, cosmics run w/ field

✦ S&C services behave rather well under data taking conditions

★ initial beam data taking fully successful

★ chain of data and work flows from HLT farm through Tier-0/1/2/3 works

★ continuous Cosmic data taking at 300kz for 4 weeks, 355M events

✦ 277M events available with full detector and magnetic field, reprocessed at Tier-1s

600MB/s tested

Headroom!

240MB/s average

> Continuous Cosmics data taking from Oct. 17 to Nov. 11 with the complete detector

> Collected over 355 Million Cosmics events ready for analysis

▶ 277 Million events with magnetic field on

▶ Gained valuable lessons not only about the detector

Data Transfer Out of CERN to Tier-1 Centers

CMS PhEDEx - Transfer Rate
30 Days from 2008-10-14 to 2008-11-13

Transfer Rate [MB/s]

Time

T1_US_FNAL_Buffer T1_FR_CCIN2P3_Buffer T1_UK_RAL_Buffer T1_DE_FZK_Buffer T1_ES_PIC
T1_TW_ASGC_Buffer T1_IT_CNAF_Buffer

Daily CMS PhEDEx transfer volume...

By destination site for non-tape storage only
Sources matching 'T0' and destinations matching 'T1'
37 days from 2008-10-12 to...

×10⁶

Dataset: /Cosmics/Commissioning08-PromptReco-v2/RECO

Selection: DT,SIST,BFieldOn

Tot Events: 277267520

Events/Run
prompt-reconstructed

Data Serving on the Grid for Physics Analysis

✦ Fermilab well established as the major CMS data center, serving CMS processing needs and supporting U.S. data analysis

★ serving data to Tier-2, Tier-3 and LPC-CAF physics users

Data Served Out From Fermilab into the Grid

Thursday, March 5, 2009
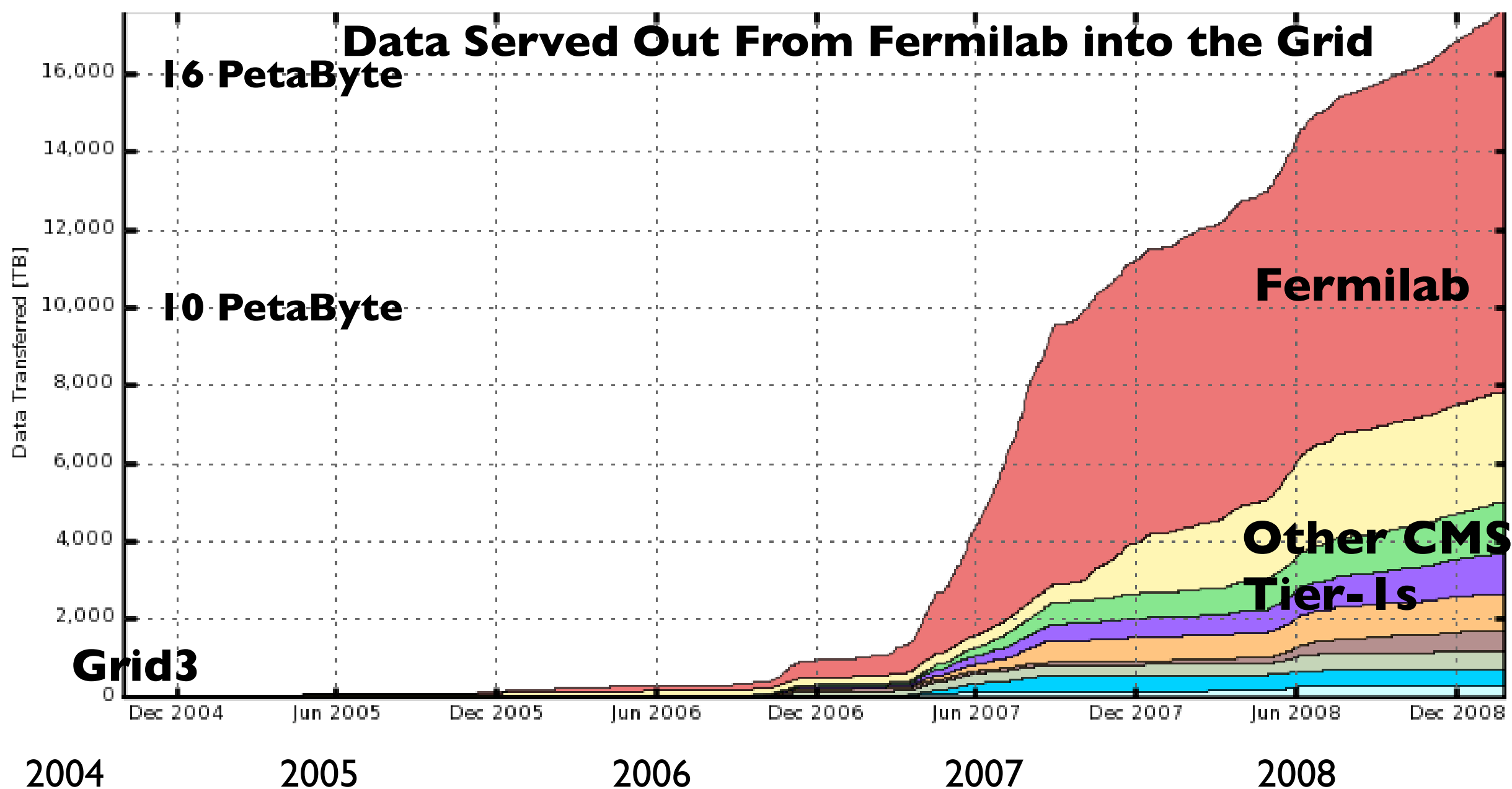
U.S. Computing Infrastructure: Tier-2 Centers

✦ **U.S. Tier-2s extensively used for simulation and analysis processing.**
  - ★ all the US Tier-2 sites reached the nominal capacity by this summer
  - ★ FY09 Tier-2 ramp to double storage and increase CPU by 50%

✦ **Tier-2 program great success thanks to very engaged sites**
  - ★ Nominal Resources per site:

US-CMS Tier-2 Summer 2008	CPU T2	1MSI2k	Tier-2 Processing Nodes
	Disk T2	200TB	dCache (200MB/s IO)
	Network	2.5-10Gb/s	WAN Networking
	People	2FTE	Supported Tier-2 Operations

✦ **Organized sites in CMS to support physics analysis services for users**
  - ★ associated sites with specific physics/detector performance groups
  - ★ allocated disk space for groups, and disk space for individual users

Using Tier-2s for CMS Physics

- Tier-1 to Tier-2 data transfers improved dramatically over last year
  - Improvements in full mesh of Tier-1 to Tier-2
  - Have successful demonstrations of 200MB/s - 400MB/s to Tier-2 sites
  - High transfer rates enable Tier-2 storage as dynamic physics cache
- Simulation at Tier-2s is a well established operation
  - Reached design rates for production
  - Good spread across grids OSG and EGEE
- Analysis moving away from CERN/Tier-1 down to Tier-2 sites
  - Analysis use and adoption of distributed computing tools is expanding
  - Rates of job submissions have steadily increased
  - Some work left to do to automate and harden the infrastructure
- Results of "Tier-2 round table" with physics customers:
  - feedback from each of the physics groups and individual users solicited
  - overall very positive feedback, individual issues are being dealt with
  - assessed list of services that would make user experience better

University Computing: Access to U.S. CMS Facilities and CMS Data

- ✦ Number of Tier-3s is getting >> than number of Tier-2s
  - ★ Tier-3s require particular support, from CMS, OSG, community
    - ✦ S&C providing support as part of Grid services, with OSG
      - ✦ OSG has well-defined model how to integrate new Tier-3s
      - ✦ USCMS support CMS-specific services, like software installs
    - ✦ USCMS has hired dedicated Tier-3 support personnel
      - ✦ had started the effort with 25% of an FTE during the fall
- ✦ Strategy for including University computing facility relies on OSG
  - ★ large interest at Universities to bring their resources online to the Grid
  - ★ OSG is providing what sites need to succeed on the Grid
  - ★ setting up OSG enables sites to get access to CMS data and resources at Tier-2 and Tier-1
- ✦ OSG enabler to include University computing ad campus grids into CMS infrastructure, and to allow University scientist access to the large samples of CMS data
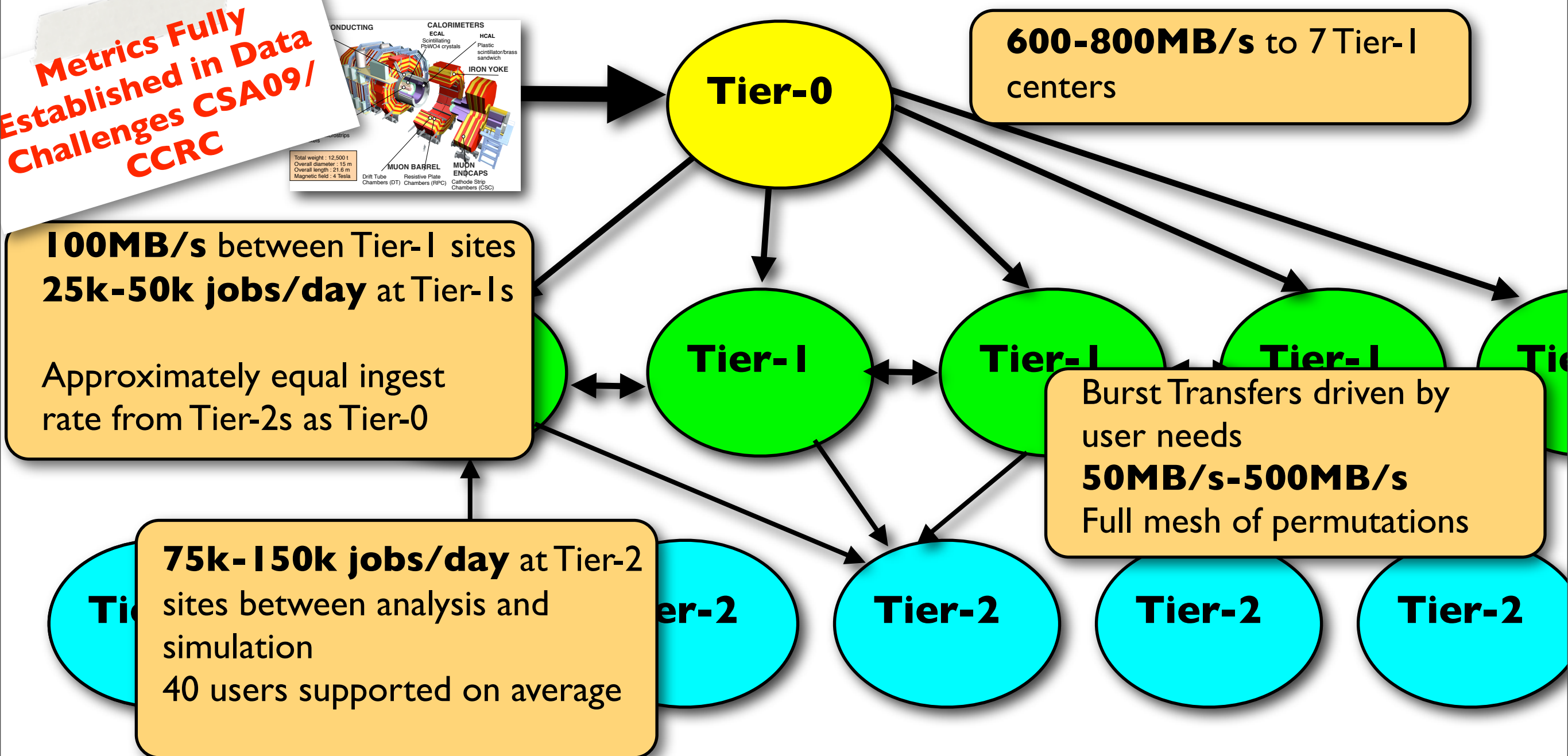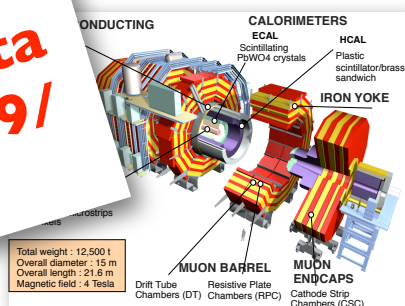
CMS Computing Metrics

✦ Reminder of CMS Computing Model:

★ user submit jobs, workload management system submits to where data is

★ data movements are triggered by operators, physics organizers, users

Metrics Fully Established in Data Challenges CSA09/ CCRC

Tier-0

600-800MB/s to 7 Tier-1 centers

100MB/s between Tier-1 sites
25k-50k jobs/day at Tier-1s

Approximately equal ingest rate from Tier-2s as Tier-0

Tier-1   **Tier-1**   **Tier-1**   Ti

Burst Transfers driven by user needs
50MB/s-500MB/s
Full mesh of permutations

75k-150k jobs/day at Tier-2 sites between analysis and simulation
40 users supported on average

Tier-2   **Tier-2**   **Tier-2**   **Tier-2**

Getting Data to Tier-2s for Physics Analysis

**Transfers
Fermilab —> Tier-2 sites
As tested in CCRC08**

CMS PhEDEx - Transfer Rate
72 Hours from 2008-05-27 00:00 to 2008-05-30 00:00 UTC

600 MB/s

250-300MB/s

- ✦ Goals reached
  - ★ Tier-1 —> Tier-2
    50-500MB/s in bursts
- ✦ Grid transfers work in general
  - ★ Still working to get every individual link up to speed
  - ★ great progress due to site and link commissioning work over this year
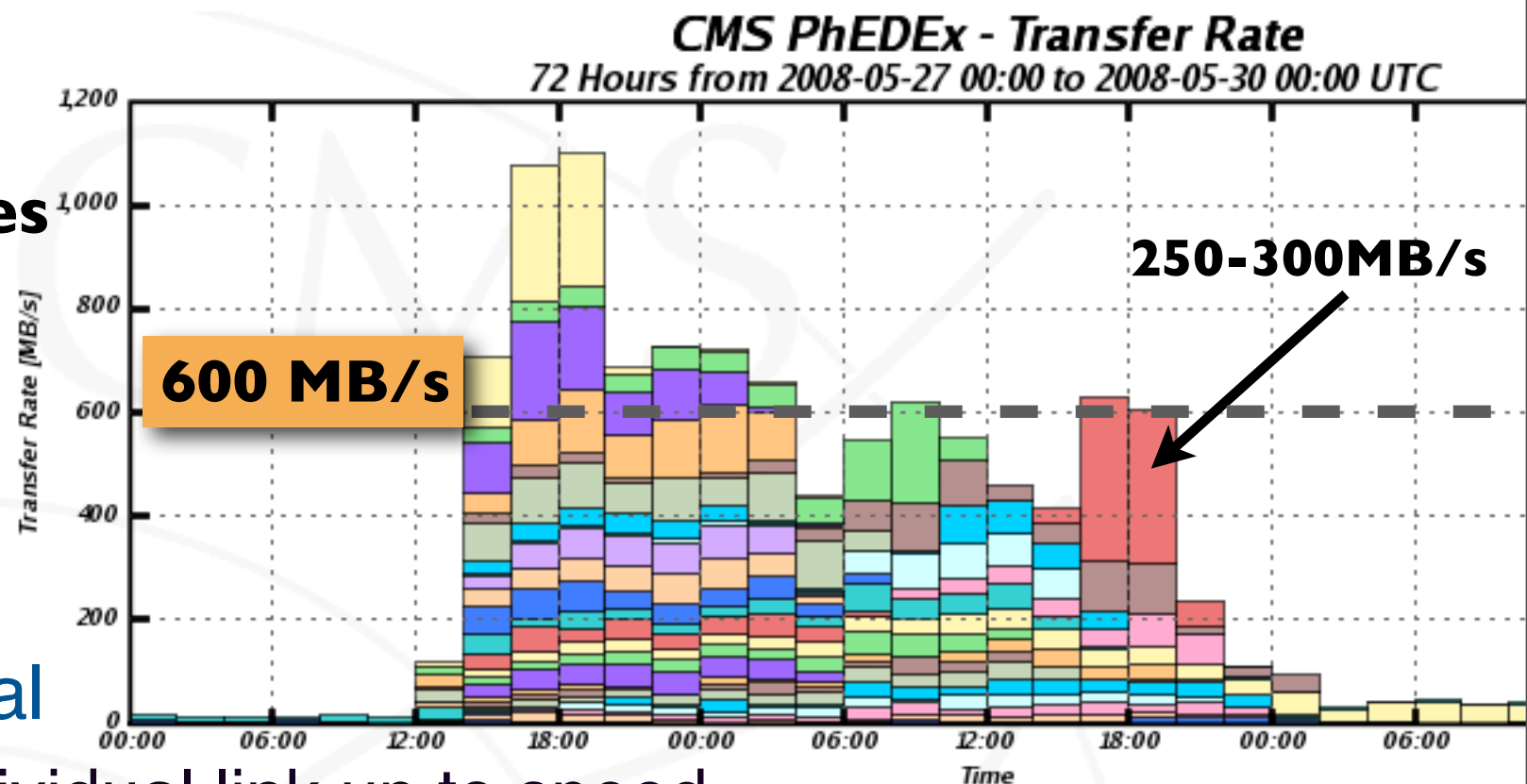- ✦ Broad utilization of analysis tools by user
  - ★ Nearly 800 individuals submitted CRAB jobs in 2008
  - ★ Have reached nearly 100k jobs by 80-100 individual users per day
    - ✦ reaching 50k jobs for individuals as demonstrated in computing challenge
  - ★ Still working to improve the overall user experience on grid
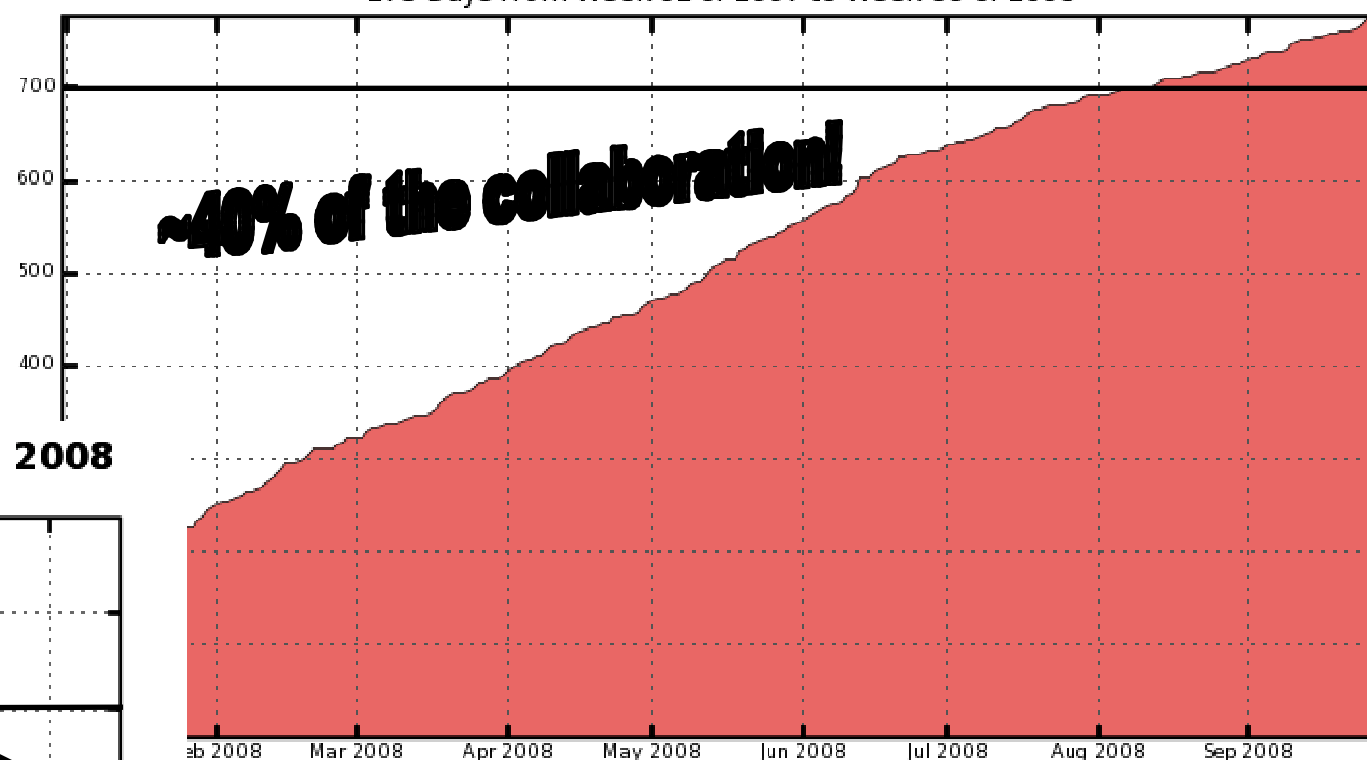   - ✦ but doing ok according to user feedback

Distributed Data Analysis

★ CMS Tool CRAB, with the glide-in WMS

✦ Good adoption by the collaboration: 40% of collaboration using it!

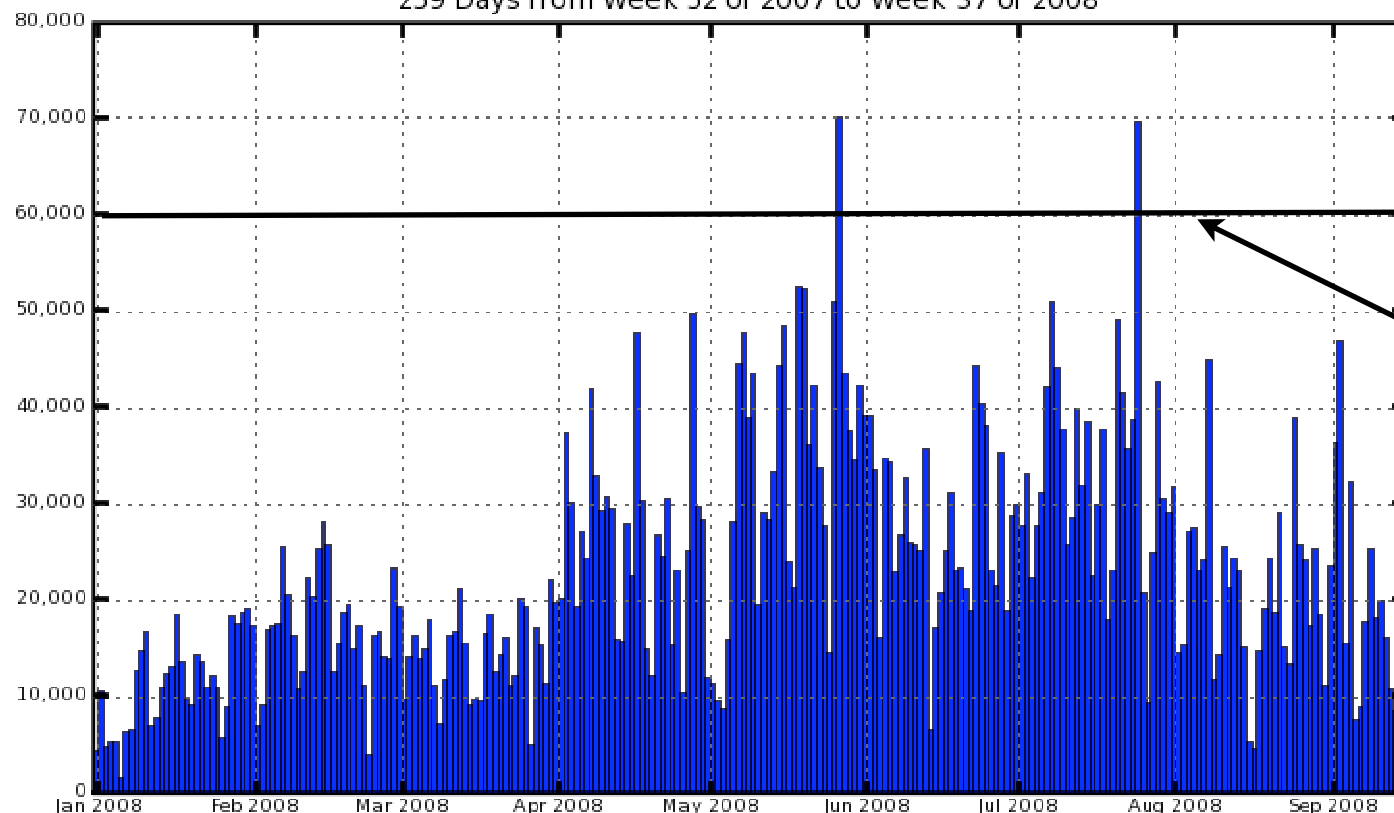**Crab distinct users from the beginning of 2008**
273 Days from Week 52 of 2007 to Week 39 of 2008

700 people →

~40% of the collaboration!

Analysis jobs terminated per day from the beginning of 2008
259 Days from Week 52 of 2007 to Week 37 of 2008

60k jobs

Thursday, March 5, 2009

# CMS Largest User VO within OSG

Weekly Use of OSG (Wall Clock Hours)
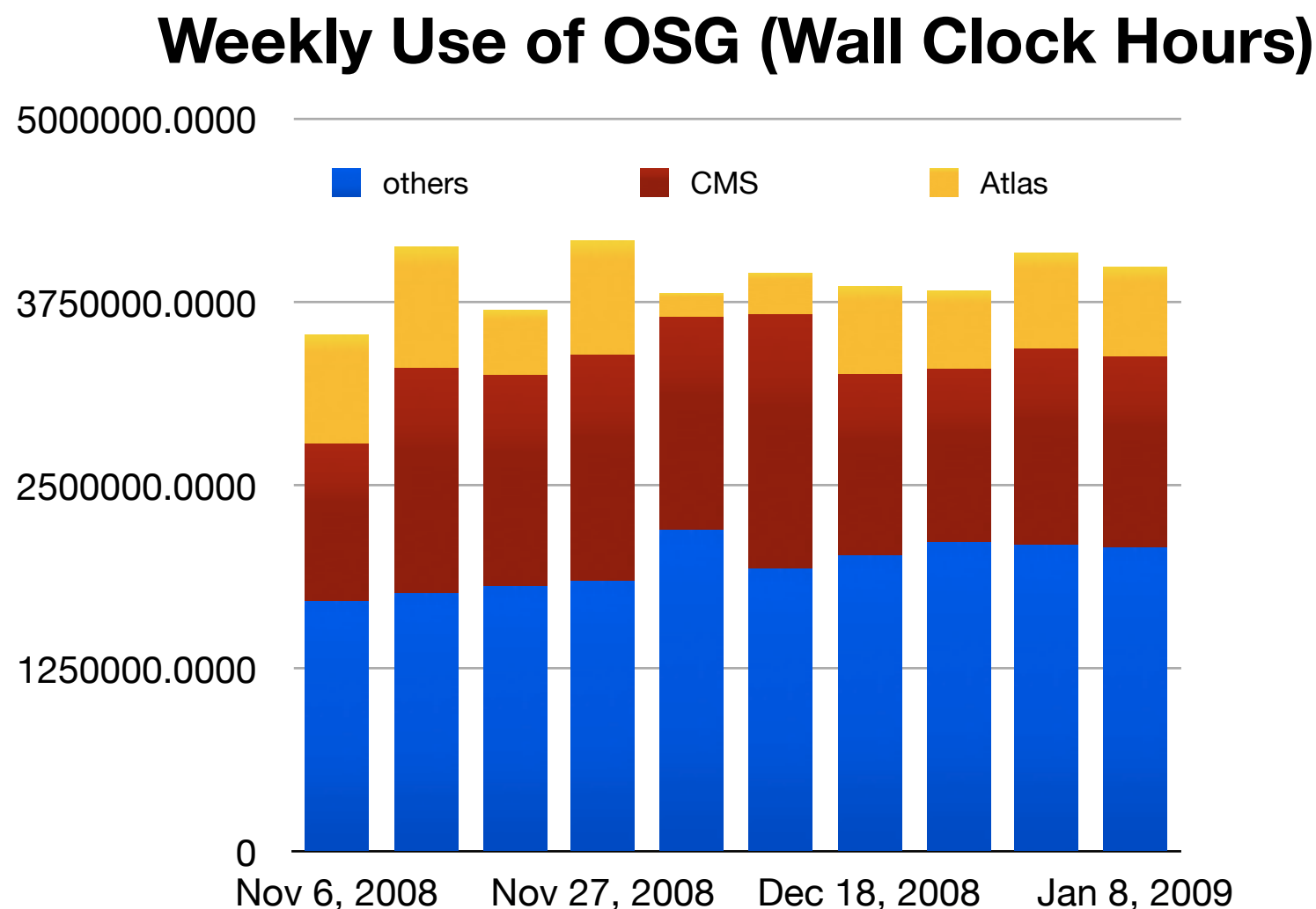
CMS did a massive production of simulated events and
a reconstruction processing of cosmics data
over the months of Nov 2008 - January 2009

✦ Start of Data Taking in October, and cosmic data taking in ~July

★ then a looong run for ~44 weeks

✦ in Spring 2009: end-to-end exercise to verify that luminosity, conditions and data quality information are consistently handled

★ putting it all together, running through the standard workflows at all tiers
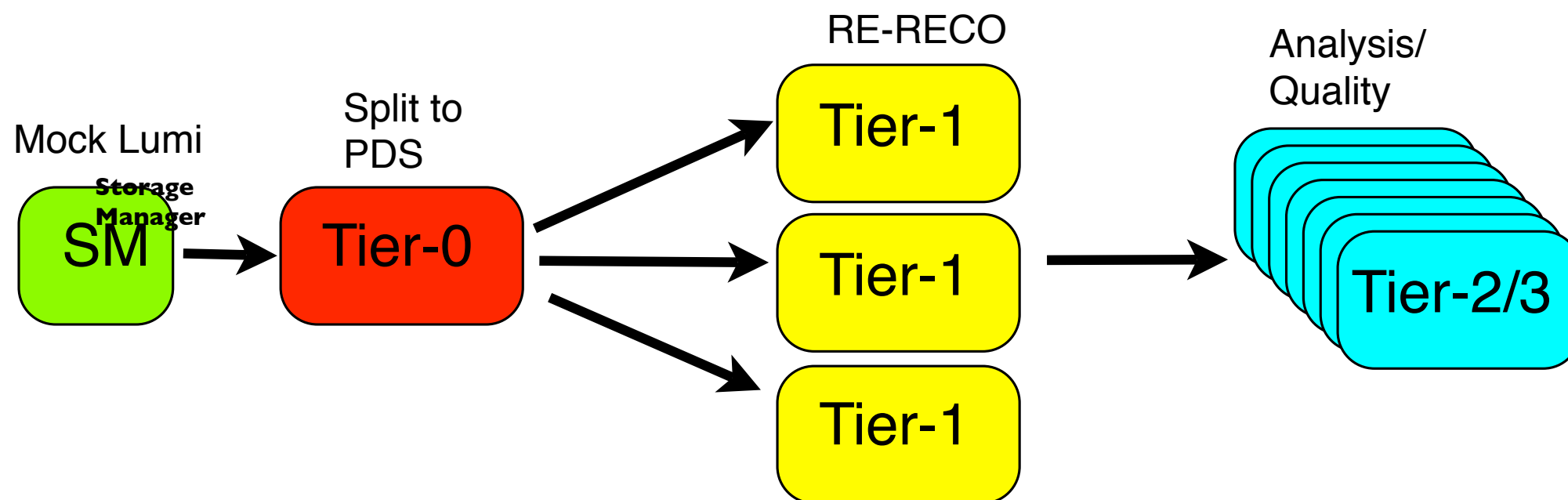
✦ Moving towards Analysis Operations

★ task force started to assess effort and tools required —> needs to OSG

★ transition to analysis operations in spring —> load on OSG

Specific Needs:
Operations of Services

★ Security context. This includes operational, policy, software, validation.

★ BDII at the GOC.

★ GOC ticketing system.

★ Accounting and RSV, and related reporting to WLCG

★ Software cache and software lifecycle management via the VDT, including the OSG cache, as well as deployment testing, incl. vtb and itb.

★ Variety of validation, including VOMS and GIP.

★ Registration and administrative services

★ Support, including OSG-storage (office hour, ticketing), and CE deployment.

CMS identified the BDII as special among these because it is a single point of failure which can make the entirety of OSG disappear from view.

Requested scalability testing of BDII, resulting in a clear understanding of operational parameters, now and in the future. (ongoing)

Specific Needs: Accounting and Reporting

✦ Accounting, reliability, and availability reporting
  - ★ OSG is responsible for aggregating the reporting for CMS, and present CMS towards WLCG for problem resolution in WLCG accounting.
  - ★ OSG prepare a monthly report for the funding agencies to offset any misrepresentations by WLCG.

✦ OSG to engage in WLCG, and be pro-active about changes in WLCG requirements.
  - ★ E.g. transition to SI2006, space accounting, deployed capacity accounting.

★ *All goals are being met by OSG.*

Tier-2 Reliability and Accounting

★ Availability and Reliability is being published since the end of the summer

Federation	Site	CPU Count	Reli-ability	Avail-ability	Reliability History		
					Oct-08	Nov-08	Dec-08
T2_US_Caltech (USA, Caltech CMS T2)							
	cit_cms_t2	N/A	97 %	97 %	68 %	85 %	91 %
T2_US_Florida (USA, Florida CMS T2)							
	uflorida-hpc	N/A	96 %	96 %	94 %	92 %	100 %
	uflorida-ihepa	N/A	100 %	100 %	99 %	100 %	100 %
	uflorida-pg	N/A	100 %	100 %	99 %	98 %	100 %
T2_US_MIT (USA, MIT CMS T2)							
	mit_cms	N/A	77 %	77 %	87 %	93 %	90 %
T2_US_Nebraska (USA, Nebraska CMS T2)							
	nebraska	N/A	99 %	99 %	92 %	97 %	92 %
T2_US_Purdue (USA, Purdue CMS T2)							
	purdue-lear	N/A	N/A	N/A	N/A	N/A	N/A
	purdue-rcac	N/A	97 %	97 %	98 %	93 %	95 %
	purdue-steele	N/A	99 %	99 %	86 %	96 %	97 %
T2_US_UCSD (USA, UC San Diego CMS T2)							
	ucsdt2	N/A	88 %	88 %	100 %	98 %	78 %
T2_US_Wisconsin (USA, U. Wisconsin CMS T2)							
	GLOW	N/A	96 %	89 %	100 %	100 %	100 %

Specific Needs:
Site Operations Improvements

✦ Ability to announce downtimes of both CE & SE.

✦ Ability to ban single FQAN at a site.

✦ Integrate all existing monitoring functionalities of CE & SE into one site-wide "monitoring portal".

✦ Ability to have multiple CEs per cluster viewed as one as in EGEE.

✦ Ability to reject new jobs while in drain-off.

Adequate progress made, but none of these is completed.

(completion = deployed and in operations at CMS sites)

Additional operational goals:

Both are met.

Reinstallation of CE in 4-8 hours
Reinstallation of SE in 6-12 hours

WMS Issues

✦ OSG support during Production roll-out of glideinWMS. (ongoing)

✦ Help from "scalability, reliability, usability" area in OSG with condor testing. (done)

✦ glideinWMS not to submit glideins to sites that are off in BDII. (incomplete)

Storage Issues

- ✦ A lot of progress, some open issues, and: do we need a road-map?
  - ★ LCG client tools (done)
  - ★ Achieve 2Hz SRM request rate (done)
    - ✦ CMS realized change in needs: 20-100Hz
  - ★ Additional ops tools for dCache ops.
    - ✦ Locally configurable space usage monitor (scheduled for next release)
  - ★ Procedures for operations problems to feed back into development cycle. (scheduled for end of year 3)
  - ★ Successfully deploy BestMan at 5-10 OSG sites for production use with PhEDEx. (partially done)
  - ★ Mutually compatible gLiteUI <-> OSG client.
    - ✦ (partially done via lcg client tools deployment as part of OSG client)
- ✦ future storage systems for Tier-2, Tier-3?
  - ★ need to chart out a way forward
  - ★ require more manageability, reliability, usability
  - ★ the Nebraska Hadoop "coup" shows what is feasible

Summary/Conclusions

- ✦ After several stress tests and data challenges, numerous large scale simulation exercises, a (all too) short experience with beam-induced data, and extensive cosmic runs, the S&C systems and operations teams are ready for sustained data taking, processing and analysis
  - ★ Work remains in the areas of usability, automation, reliability
- ✦ Software and Computing Systems useable for a large # physicists
  - ★ Software is functional, remaining issues are being addressed
  - ★ We have significant computing resources deployed
    - ✦ that are being used for detector commissioning and physics studies
  - ★ CMS OSG Tier-2 and Tier-3 sites are working well and are being used
    - ✦ University sites started to plug into the system
    - ✦ —> system tests show that the system works for many users
- ✦ We are ready to operate software and computing 24/7
  - ★ Still, sizable development, integration, commissioning ahead
- ✦ OSG has been fully engaged and is major contributor to CMS