

Wisconsin CMS T2 Site Report



D. Bradley, S. Dasu, A. Mohapatra, T. Sarangi, C. Vuosalo
(**HEP Computing Group**)

Outline

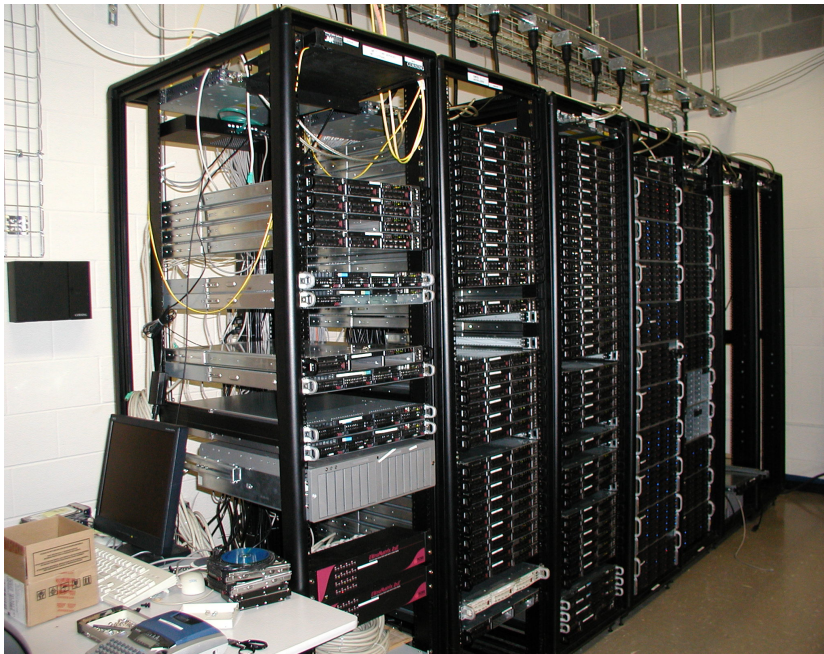
- Infrastructure
- Resources
- Management & Operation
- Future Evolution
- Summary

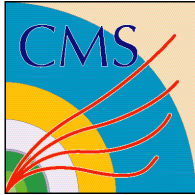


Infrastructure



- ✓ **2 machine rooms, 16 racks**
- ✓ **Power supply – 650 kW**
- ✓ **Cooling**
 - **Chilled water based air coolers and POD based hot aisles**

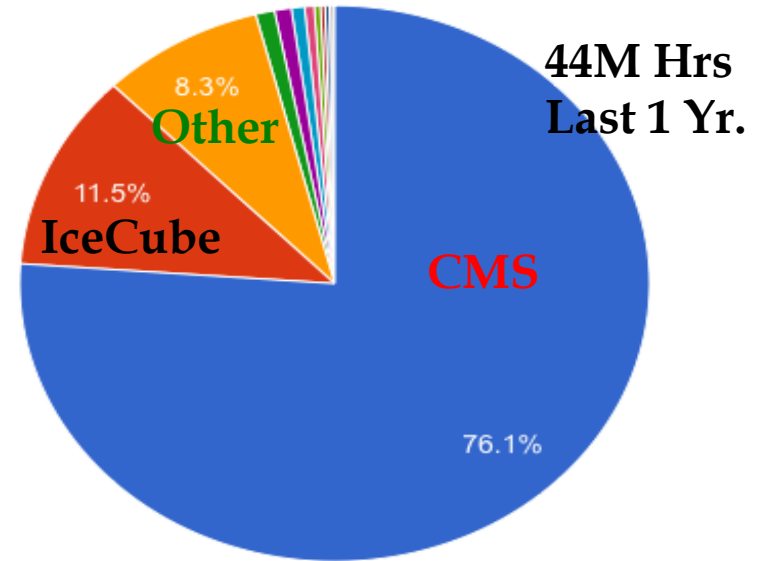




Compute / Storage Resources



- ✓ **Compute (SL6 OS)**
 - T2 HEP Pool – 6700 cores
 - Dedicated to CMS
 - CHTC Pool – 15000 cores
 - Opportunistic ~ 1500



- ✓ **Storage FS (Hadoop)**
 - OSG released hadoop-2.0
 - 4.5 PB raw → 2.25 useable (with replication factor = 2)





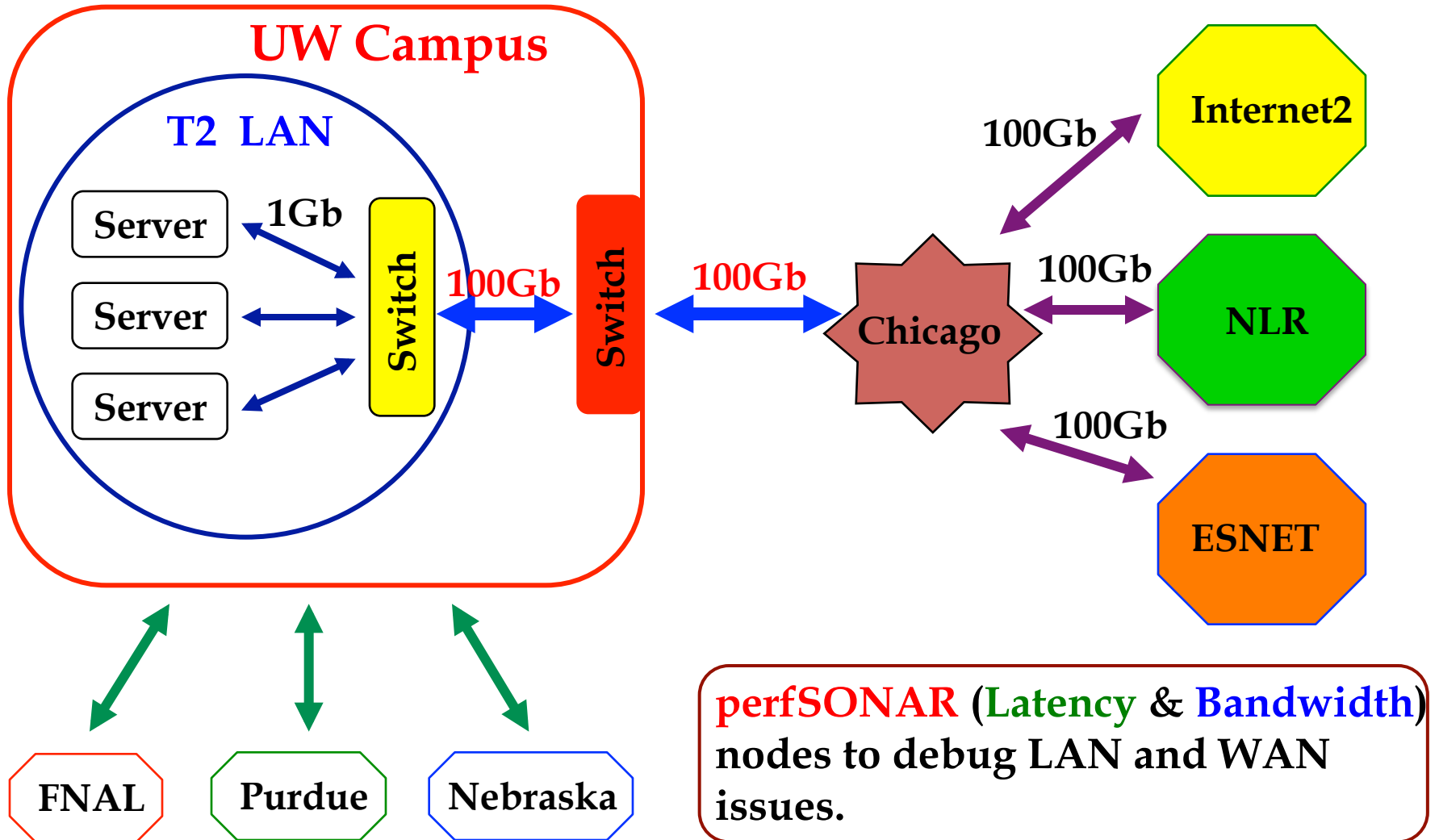
Job Scheduling and Multicore

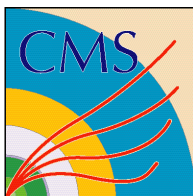


- ✓ CEs : Gram-CE → HTCondor-CE
- ✓ HTCondor as batch scheduler
- ✓ Full support for the pilot role (Production/Analysis fair share settings are removed).
- ✓ Supports `condor_ssh_to_job` (for debugging).
- ✓ Opportunistic resources are heavily used by CMS jobs (mostly analysis) regularly.
- ✓ Other VOs use the T2 slots opportunistically (in the absence of enough CMS jobs).
- ✓ Small fraction of nodes support multicore jobs as a test. More nodes can be configured if needed.



Network Configuration

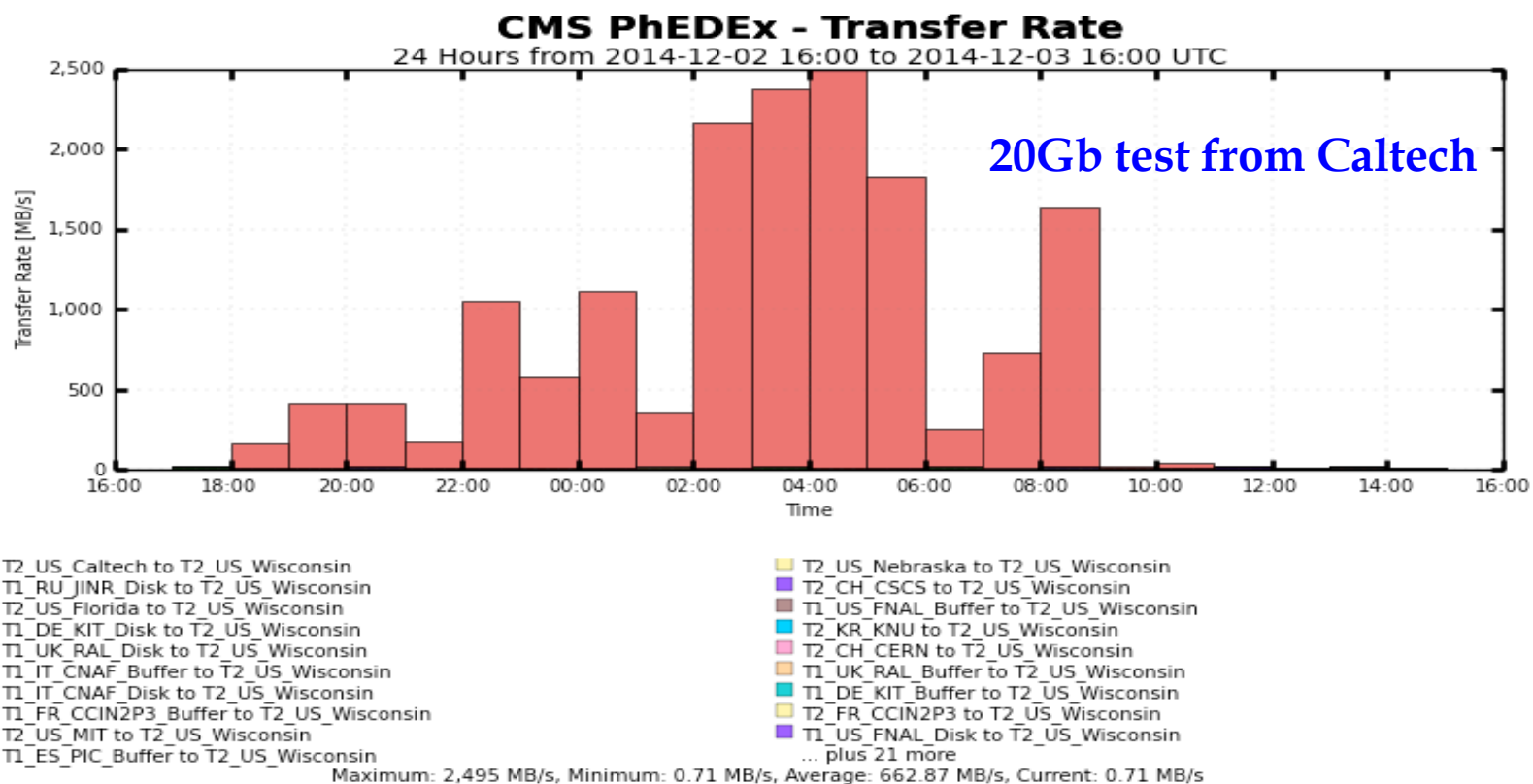


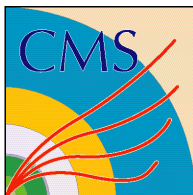


100Gb WAN Upgrade



- ✓ WAN → 100Gb (Strong support from UW network team)
- ✓ Machine rooms interconnection → 80Gb
- ✓ Within machine rooms → 1Tb





IPv4 → IPv6



- ✓ A total of 350 machines (Compute and Storage Nodes)
 - Dual Stack IPv4/v6 Network to the outside world
 - Static IPv6 address configuration

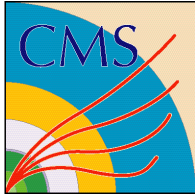
- ✓ OSG services work with IPv6-only and Dual Stack mode :
 - IPv6 is enabled for GridFTP servers and works in IPv4/v6
 - Xrootd (non-OSG release) has also been tested to work with IPv6-only and Dual Stack mode
 - **Hadoop, SRM communications haven't been tested with IPv6**



Software and Services



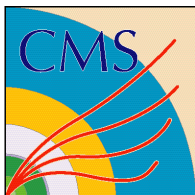
- ✓ **File systems & proxy service**
 - AFS, NFS, CVMFS, Frontier/Squid
- ✓ **Job batch system**
 - HTCondor
- ✓ **OSG software stack**
 - Globus, GUMS, glexec, SEs, HTCondor-CE.
- ✓ **Storage and Services**
 - Hadoop (hdfs), SRM, GridFTP, Xrootd
- ✓ **Cluster management & monitoring**
 - Puppet, ganeti, Nagios, Ganglia



Cluster Management & Monitoring



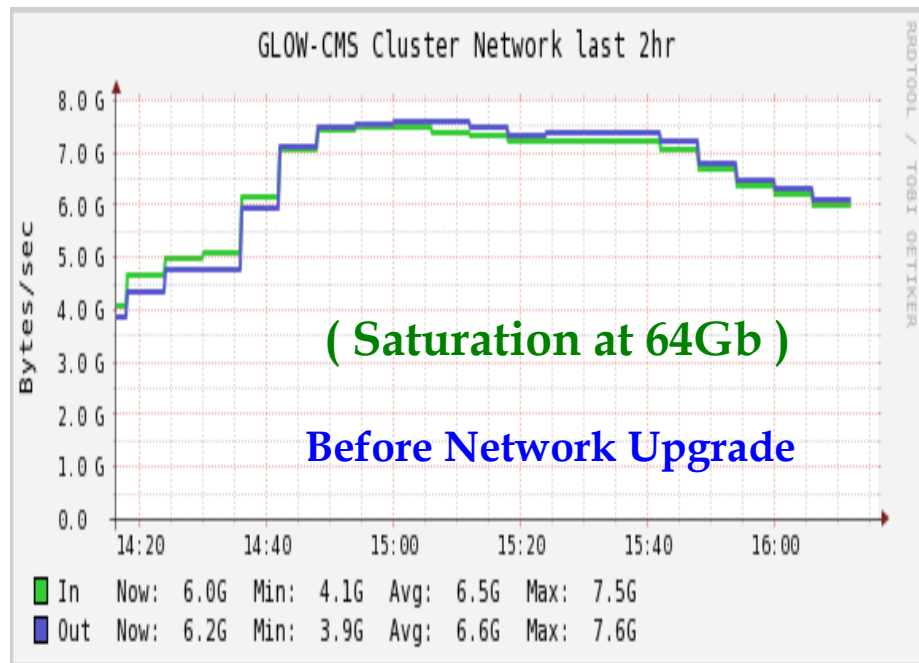
- ✓ **Puppet**
 - Manages all aspects of software deployment
- ✓ **Ganeti (VM Manager/Debian)**
 - DRBD, KVM as the underlying tech.
- ✓ **Nagios**
 - Hardware, disks temp, etc...
- ✓ **Ganglia**
 - Services, memory, cpu/disk usage, I/O, network
- ✓ **OSG and CMS dedicated tools**
 - RSV, SAM, Hammer Cloud, Dashboard
- ✓ **Network latency/bandwidth (perfSONAR)**



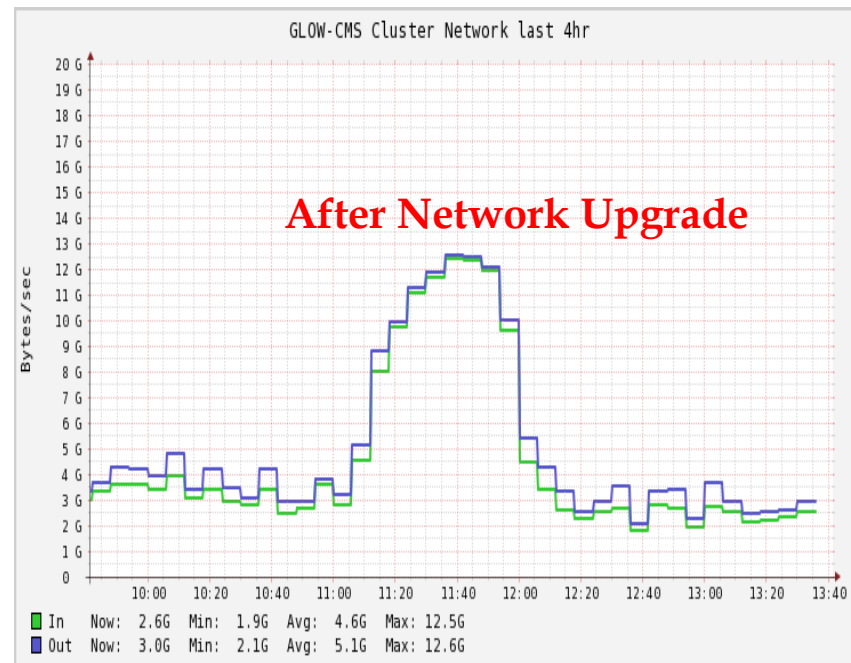
AAA Scale Tests



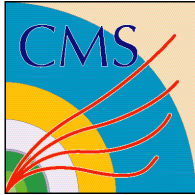
- ✓ **AAA scale test using HTCondor:**
 - **Wisconsin T2 cluster provides 10K condor job slots for the scale test (running parallel to main condor pool)**



10K File Read Test From/To Wisconsin (08/08/14)



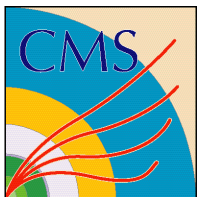
10K File Read Test (01/02/15)



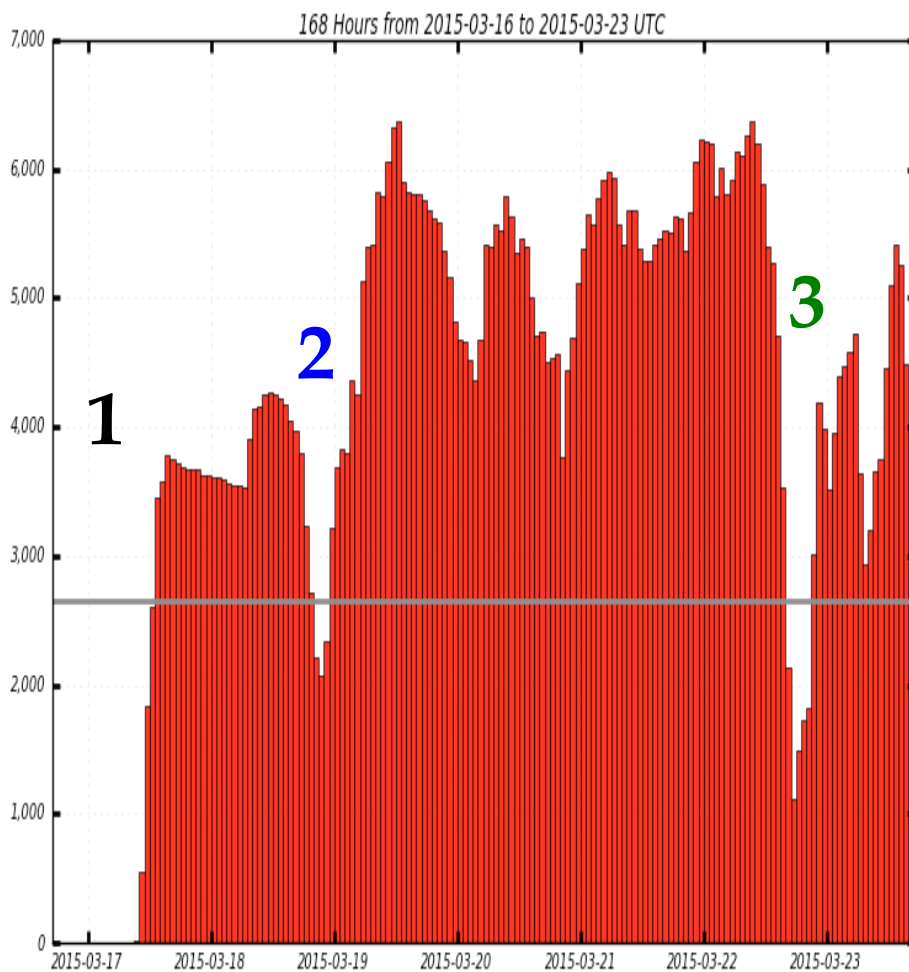
Future Evolution



- ✓ No scheduled downtimes in 2015.
- ✓ Sufficient power supply and cooling capacity for next couple of years.
- ✓ Conversion of (chilled water) air coolers to POD based hot isles for the 2nd machine room without service interruption.
- ✓ Growth and availability of opportunistic resources as campus clusters grow in size.
- ✓ Network bandwidth expansion with minimal interruption.
 - WNs with 1G NICs → 10G NICs
 - Existing 10G lines → 40G lines

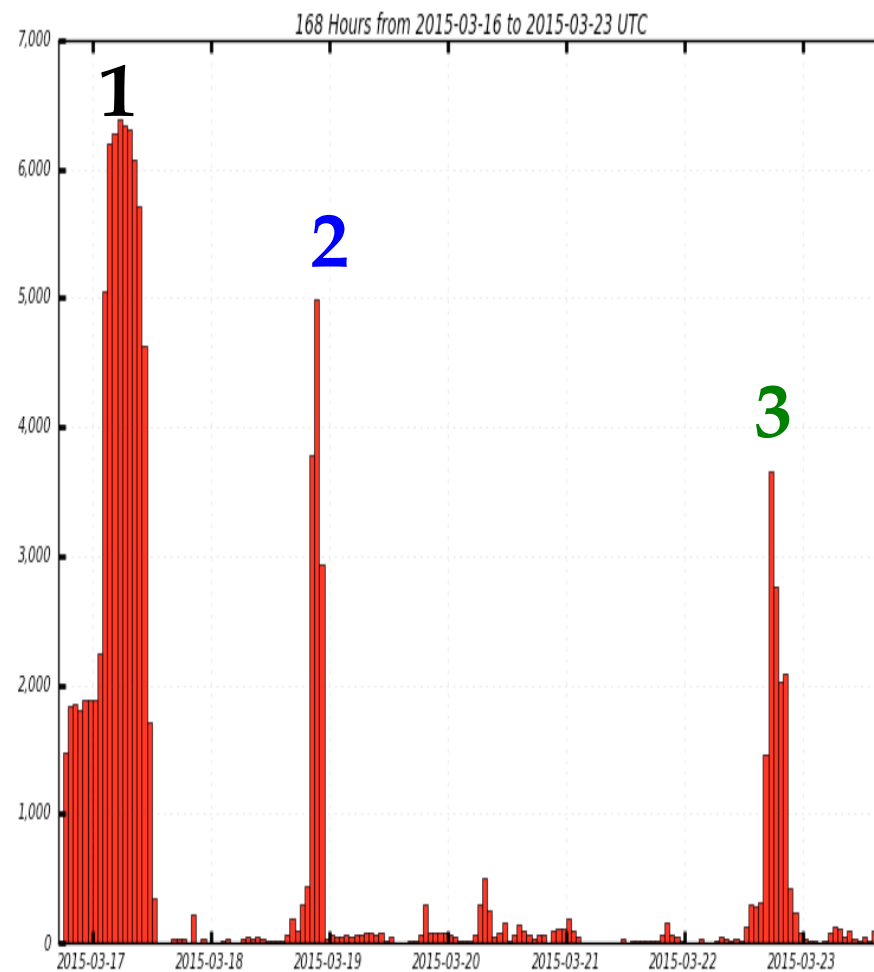


Global Pilots & Production Jobs



Running

T2_US_Wisconsin



Pending

T2_US_Wisconsin



Summary

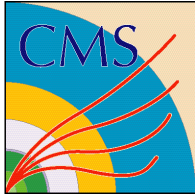


- ✓ The site is in good health and performing as expected.
- ✓ Trying our best to maintain the high availability/reliability.

T2_US_Wisconsin

Site Readiness Status: R R R R R R R R R R R R R R R R R

Daily Metric:	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	E		
Maintenance:	Up	Up	Up	Up	Up	Up	Up	Up	Up	Up	Up	Up	Up	Up	Up	Up	Up	Up	Up	Up	Up		
HammerCloud:	99%	100%	100%	100%	100%	100%	99%	100%	100%	100%	99%	100%	100%	100%	100%	100%	100%	100%	100%	100%	77%		
SAM Availability:	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	85%	100%	84%		
Good T2 links from T1s:	12/14	11/14	11/14	12/14	13/14	14/14	14/14	13/14	14/14	14/14	14/14	14/14	14/14	14/14	14/14	14/14	13/14	14/14	13/14	12/14	12/14		
Good T2 links to T1s:	14/14	14/14	14/14	14/14	14/14	14/14	13/14	14/14	14/14	14/14	14/14	14/14	14/14	14/14	14/14	14/14	14/14	14/14	14/14	13/14	13/14		
Active T2 links from T1s:	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13		
Active T2 links to T1s:	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13		
Waiting Room:	out	out	out	out	out	out	out	out	out	out	out	out	out	out	out	out	out	out	out	out	out		
	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	
	Mar																						



Thank You !



Questions / Comments ?



Available Resources



Gen	CPU Class	Slots*	HS06/slot**	Total HS06	Storage (TB)
g12	8 x 2.66 GHz Xeon X5355	208	8.31	1861	44
g14	8 x 3.00 GHz Xeon E5450	192	9.79	1801	60
g16	8 x 3.00 GHz Xeon E5472	400	10.17	4393	-
g18	16 x 2.4 GHz Opteron 6136	496	9.30	4612	189
g19	24 x 2.2 GHz Opteron 6174	696	8.51	5926	184
g20	16 x 2.67 GHz Xeon E5640	128	8.35	1069	167
g22	24 x 2.3 GHz Opteron 6176	408	8.65	3529	528
g23	24 x 2.3 GHz Opteron 6176	384	8.77	3368	300
g24	24 x 2.6GHz Opteron 6238	240	8.30	1993	310
g25	24 x 2.6GHz Opteron 6344	24	9.25	222	29
g26	32 x 2.2GHz Xeon E5-2660	960	9.64	9254	1207
g27	40 x 2.2GHz Xeon E5-2660V2	1200	9.64	11568	360
g28	40 x 2.2GHz Xeon E5-2660V2	1400	9.64	13496	228
s15	8 x 2.60 GHz Xeon	-	-	-	246
s17	8 x 2.60 GHz Xeon	-	-	-	398
s21	8 x 2.40 GHz Xeon	-	-	-	340
Total	-	6736	9.0	67092	4590