



Workshop Introduction

Ken Bloom
US CMS Tier-2 Workshop
March 8, 2010

At long last, beam:

- ➔ From a computing perspective, not much data, ~few TB minbias
- ➔ But the distributed computing analysis system responded well

Our real test comes over the next two years:

- ➔ Collisions with little interruption through 2011
- ➔ Luminosity below design by x20-50, but it's really about live time
- ➔ Enough data to make hundreds of real measurements driven by thousands of real physicists, who are looking for quick turnaround
- ➔ Current estimates show T2's over-subscribed in this period

So everything we have done so far is only practice, really.

The usual questions persist:

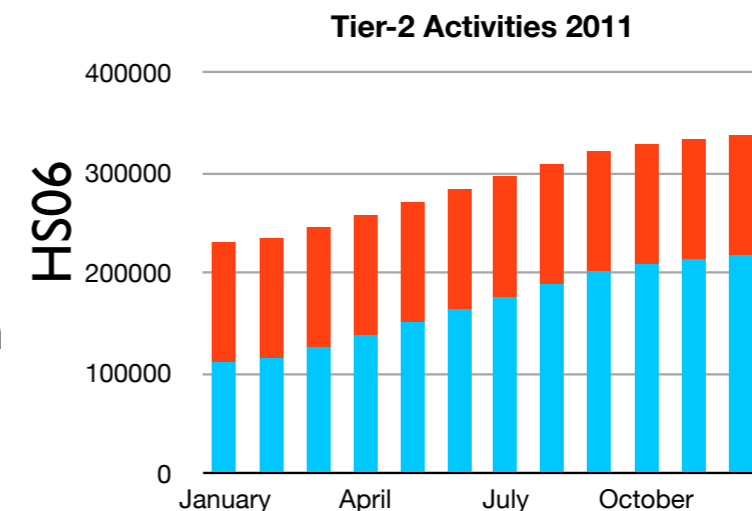
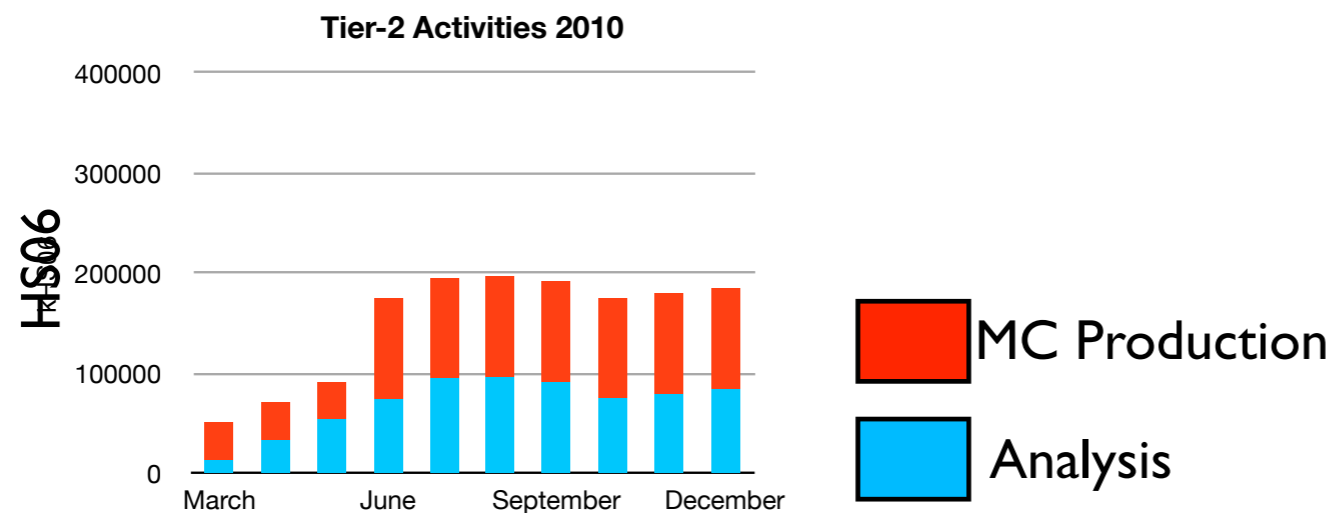
- ➔ How well are we doing?
- ➔ What can we be doing to better serve the physics program?
- ➔ What technical steps can we take to make life better?

Ian Fisk at last week's Computing Resources Board:



Tier-2s

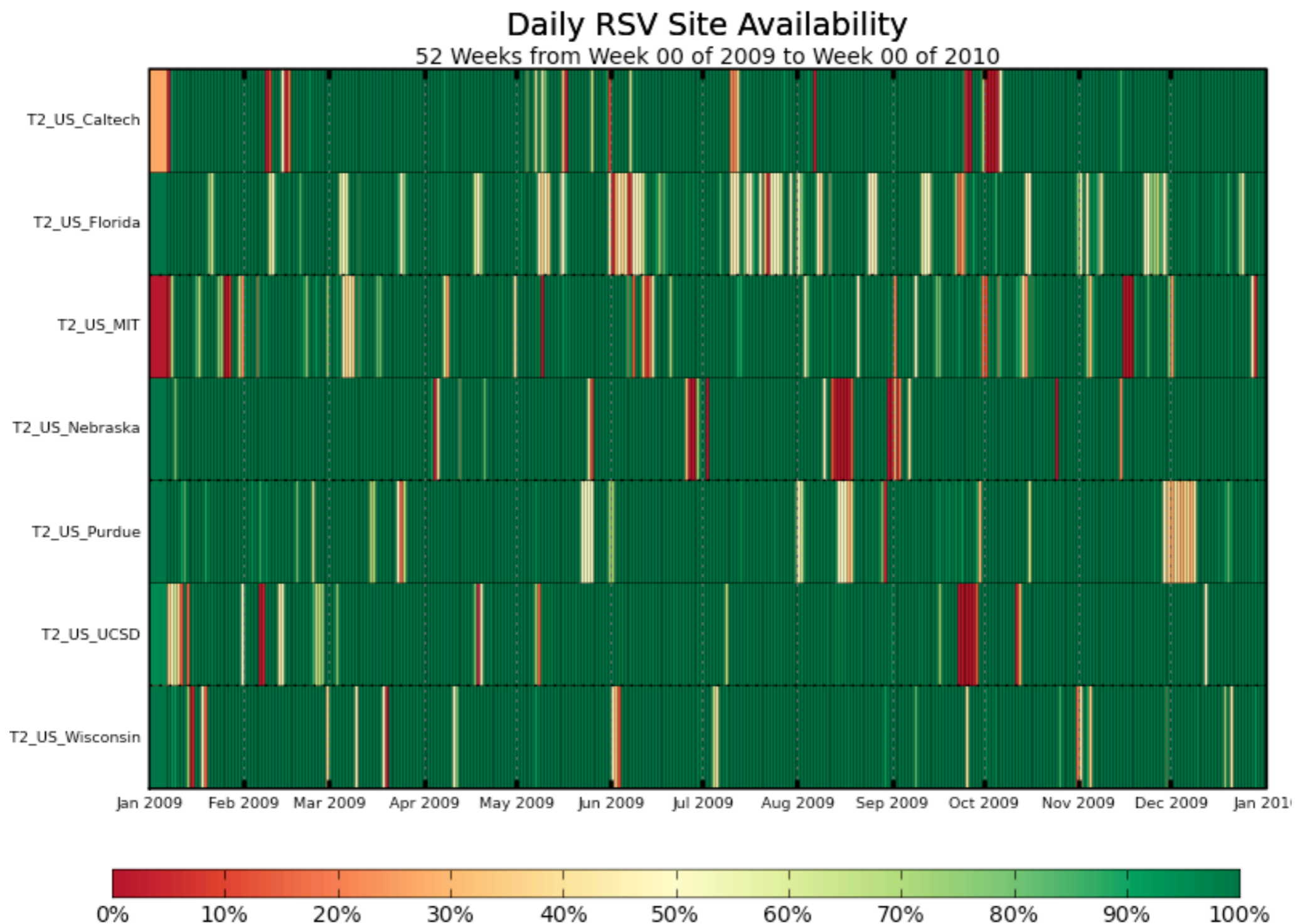
- ▶ CMS is striving to maintain a balance between analysis and simulation at the Tier-2s
- ▶ Using the model used previously for analysis CPU we can almost maintain 50% analysis and 50% simulation at the Tier-2s (if Tier-1s contribute)

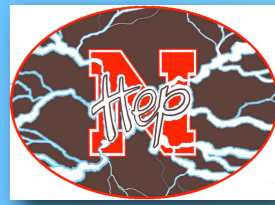


Relies on a migration to summarized data as outlined in the CMS ECoM report

- ▶ Roughly a 50%-60% increase in the Tier-2 CPU is required in 2011 to accommodate the data analysis and MC production needs

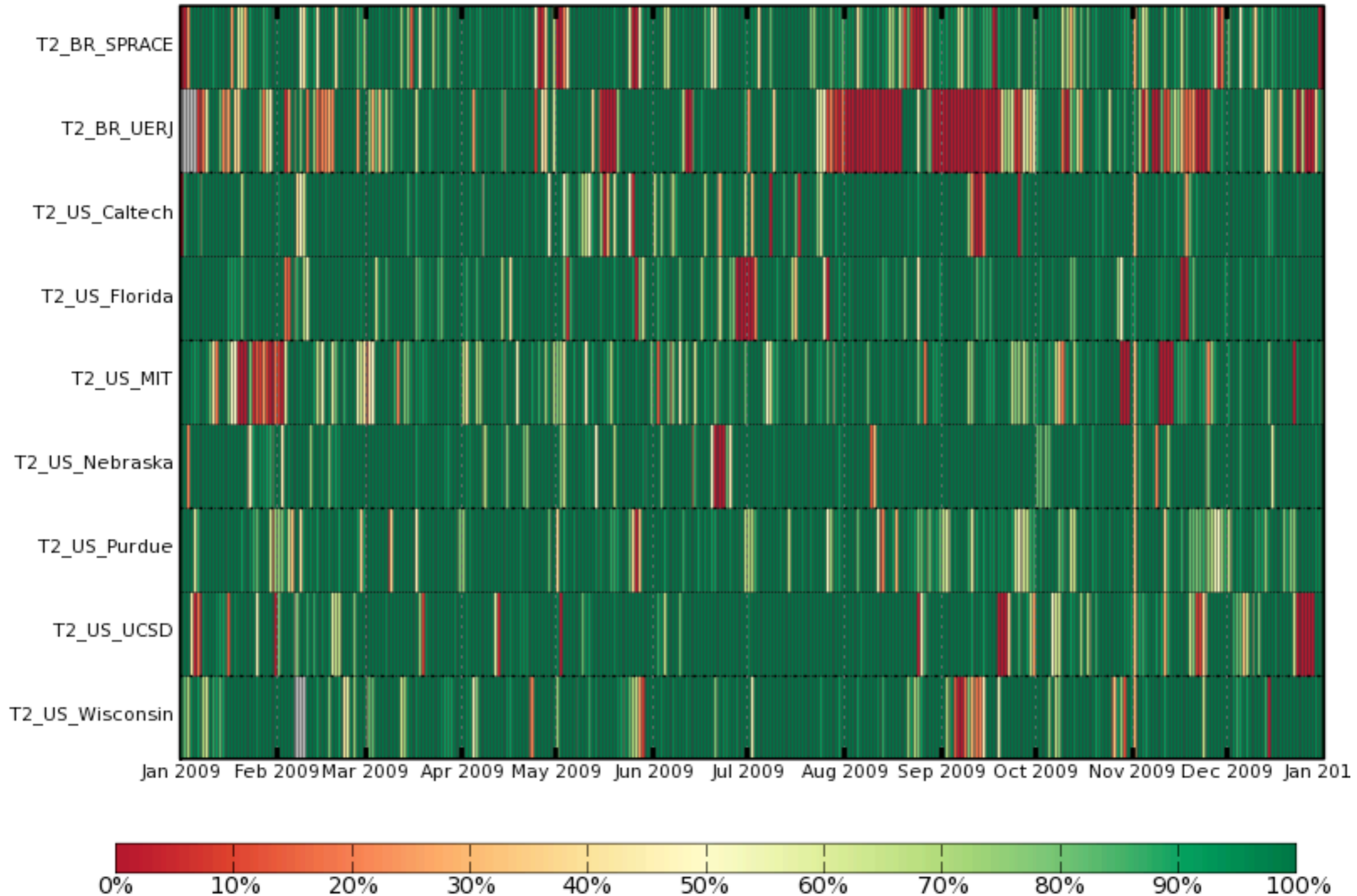
This is what is reported to WLCG for availability/reliability, and that in turn gets reported to the funding agencies.





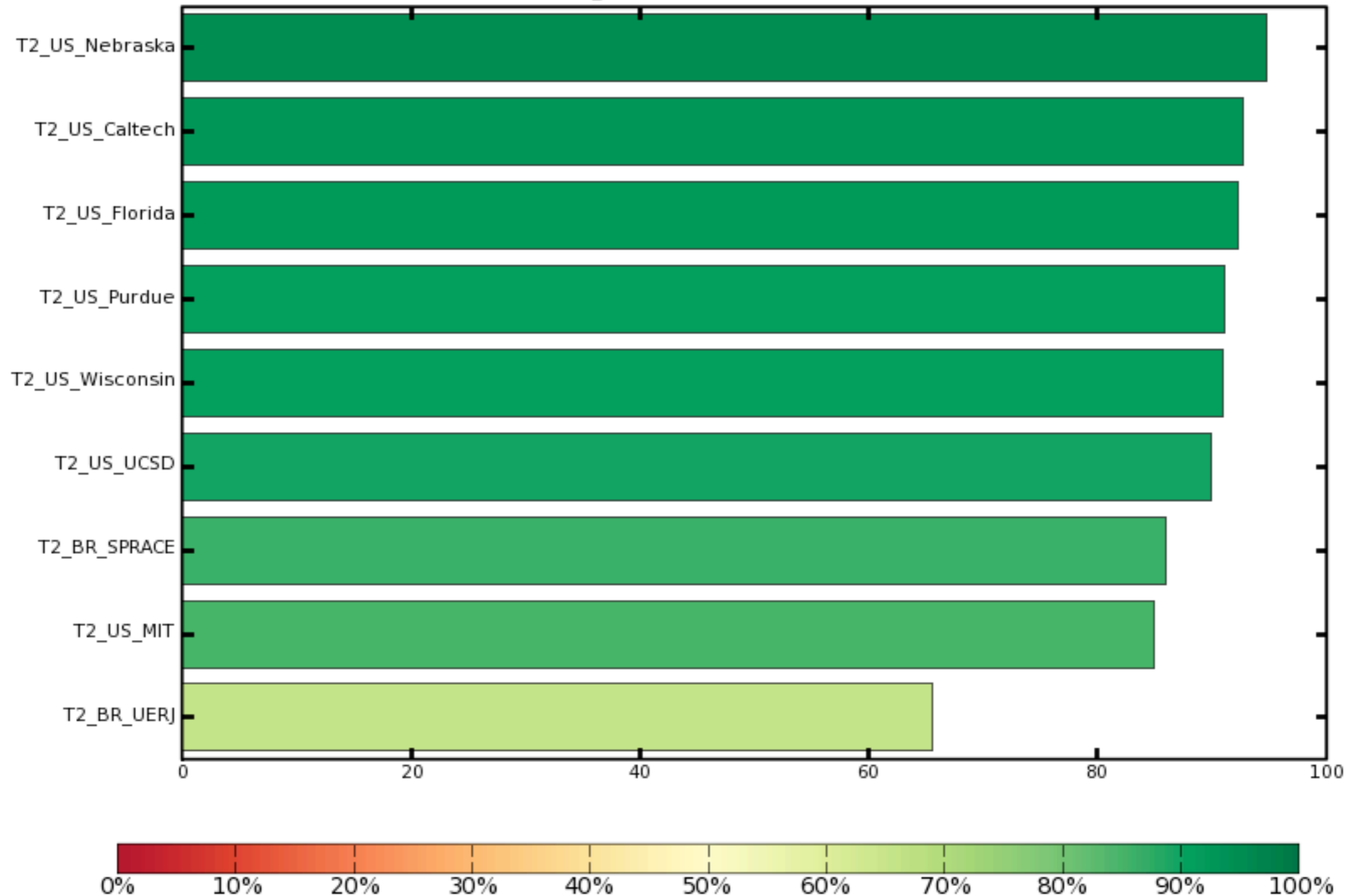
Site Availability

365 Days from Week 00 of 2009 to Week 00 of 2010



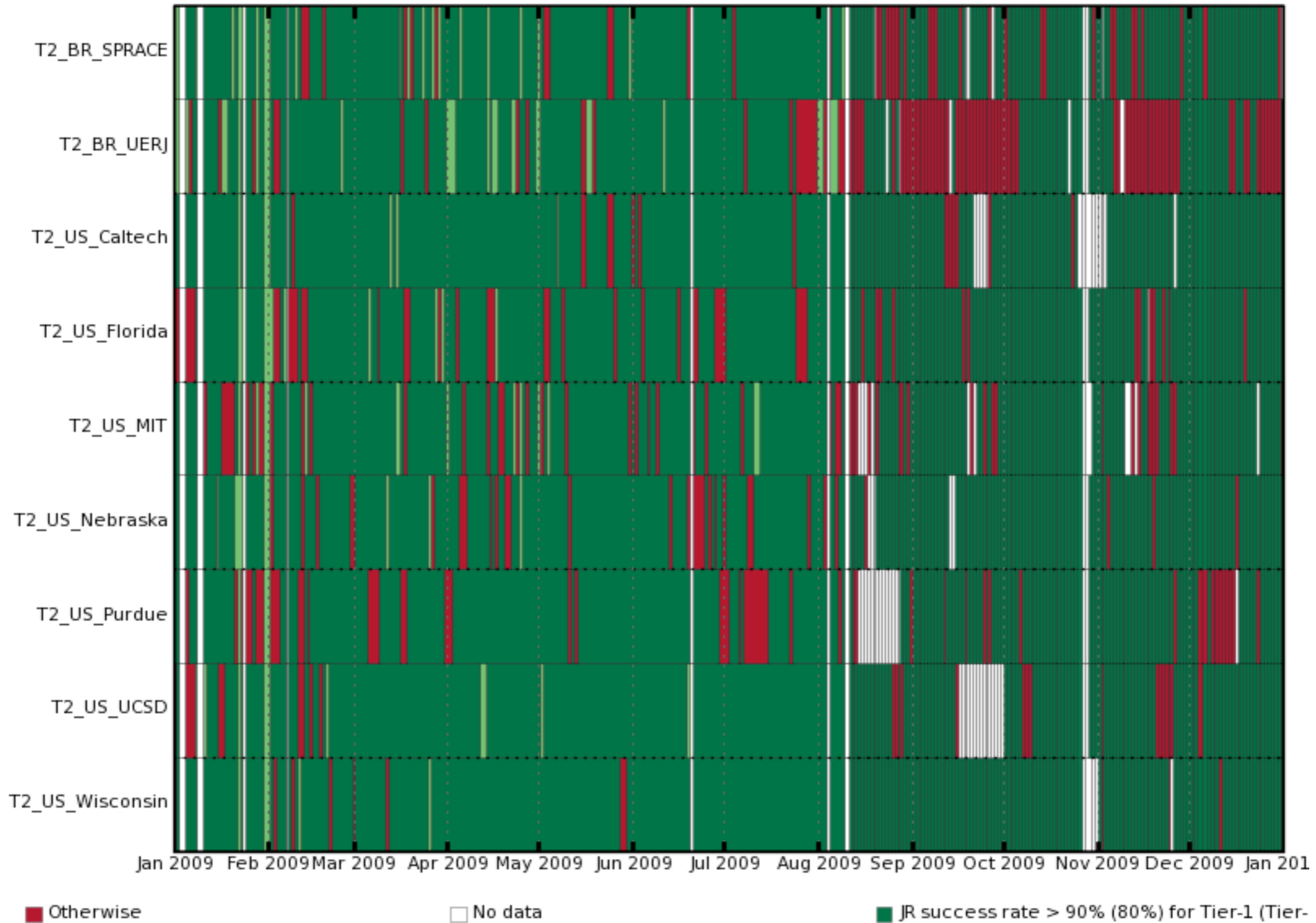
This is quite good, but only a component of what CMS uses to measure site readiness....

Site Availability, 2009-01-01 - 2010-01-01



Status of SiteComm JR

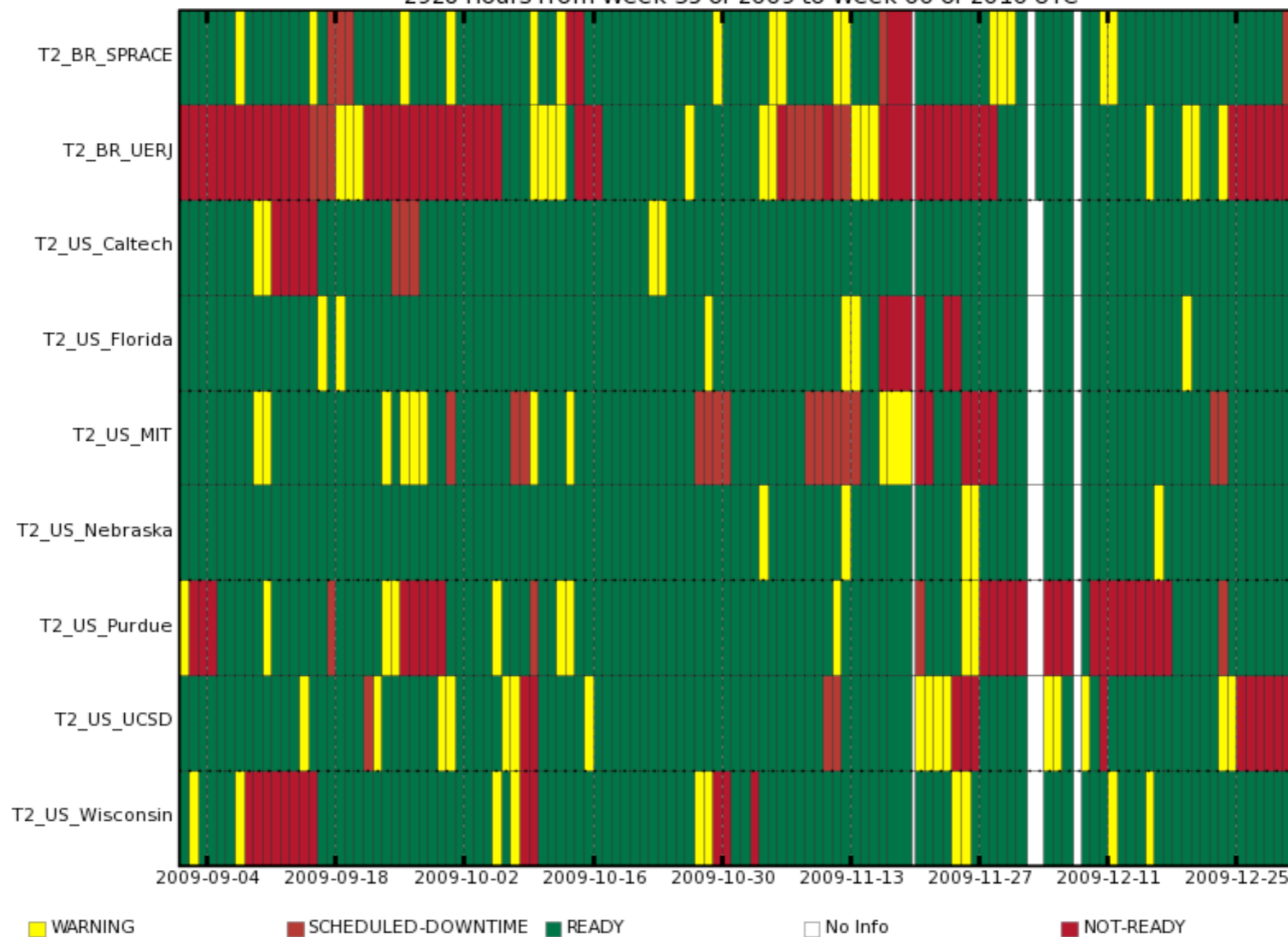
8760 Hours from Week 00 of 2009 to Week 00 of 2010 UTC



(This is only since September 1.) This is a composite of SAM, robot and (as of December) transfer tests, and it's the metric that gets shown to CMS management these days. Need to keep an eye on this.

Status of SiteReadiness Status

2928 Hours from Week 35 of 2009 to Week 00 of 2010 UTC



Somewhat better more recently, but not at all sites....

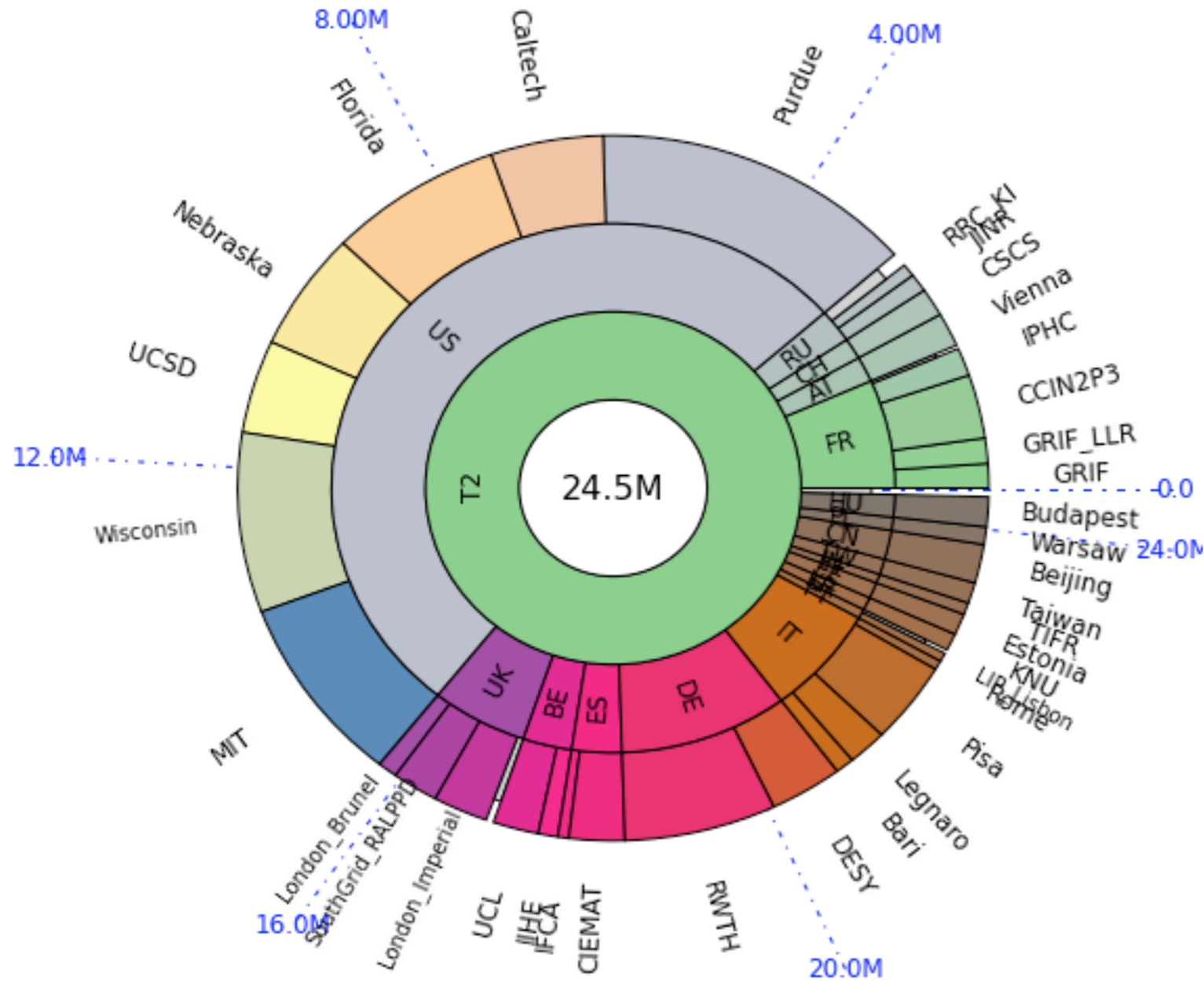
Status of SiteReadiness Status inc. good links

1416 Hours from Week 00 of 2010 to Week 09 of 2010 UTC



Time [Success] (Merge and Processing jobs)

sorted by Site matching T2

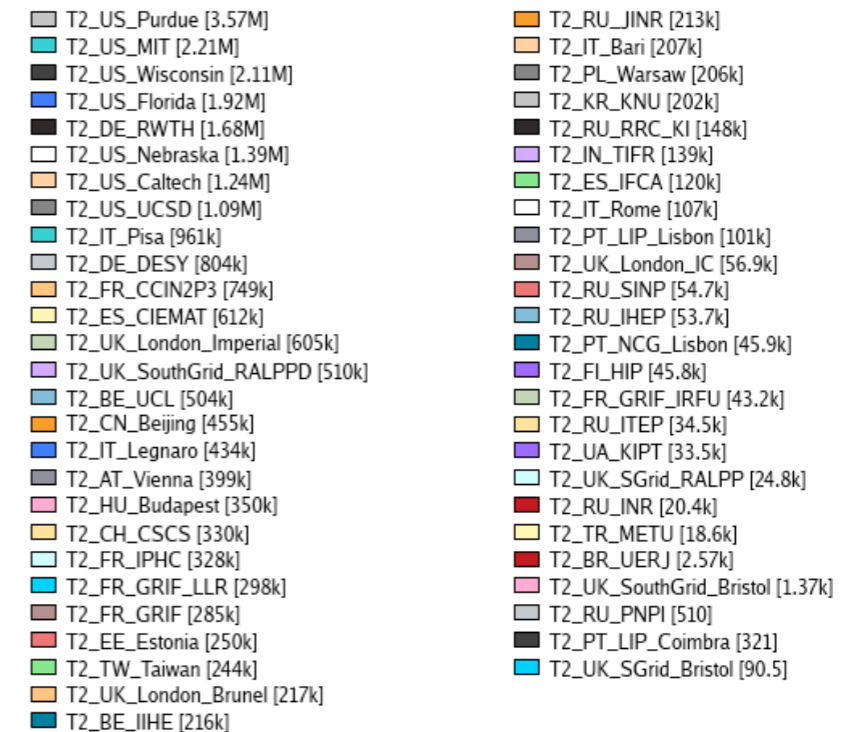
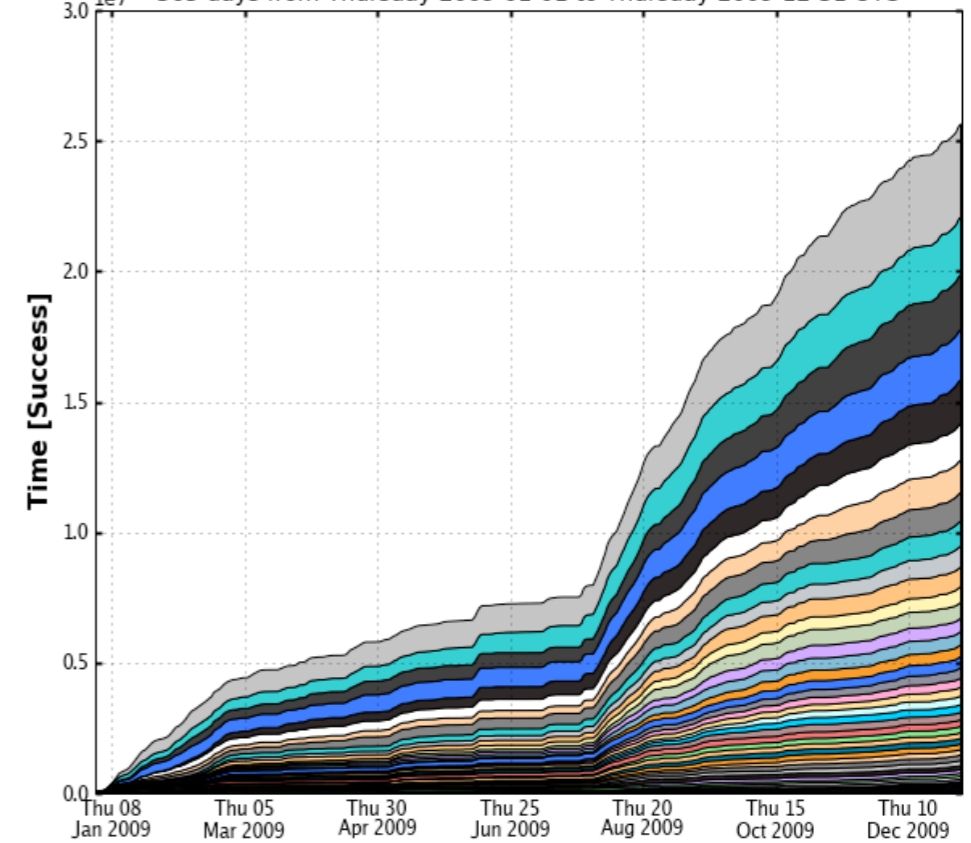


No legend

Time [Success] (Merge and Processing jobs)

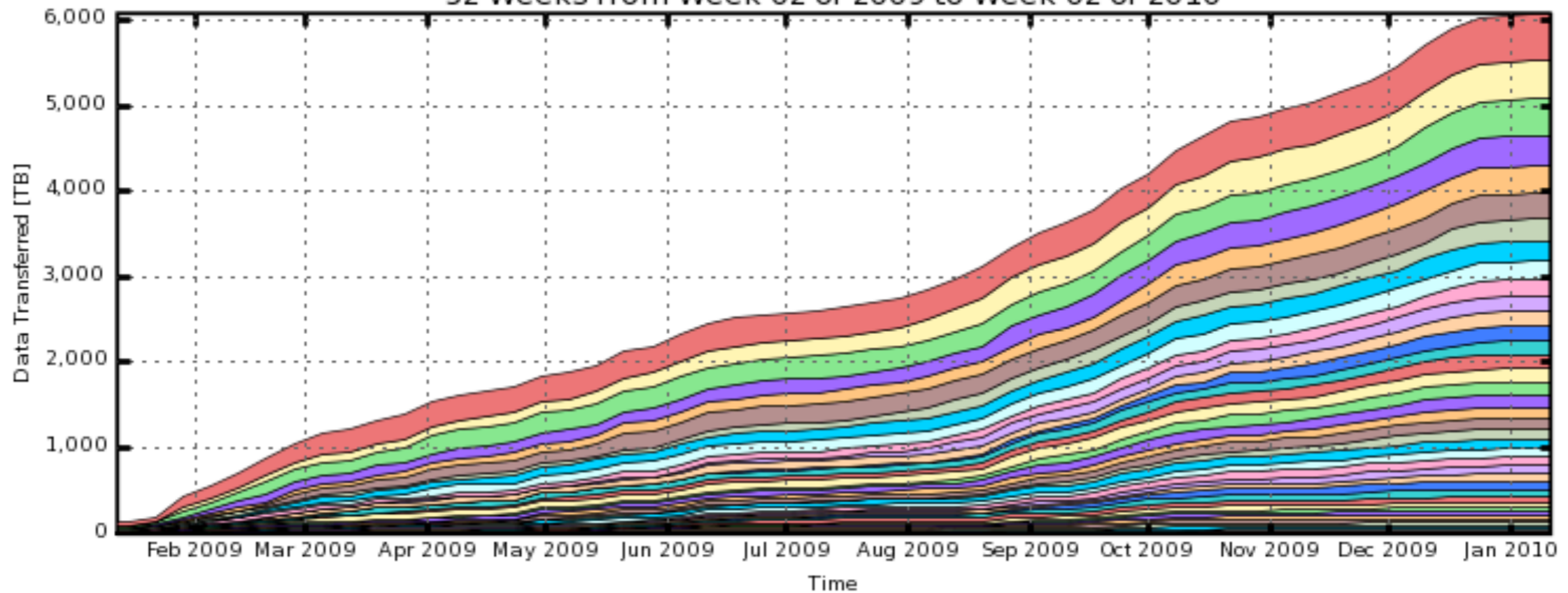
sorted by Site matching T2

365 days from Thursday 2009-01-01 to Thursday 2009-12-31 UTC



T1 to T2 transfers -- a measure of ability to make data available for analysis.

CMS PhEDEx - Cumulative Transfer Volume
52 Weeks from Week 02 of 2009 to Week 02 of 2010

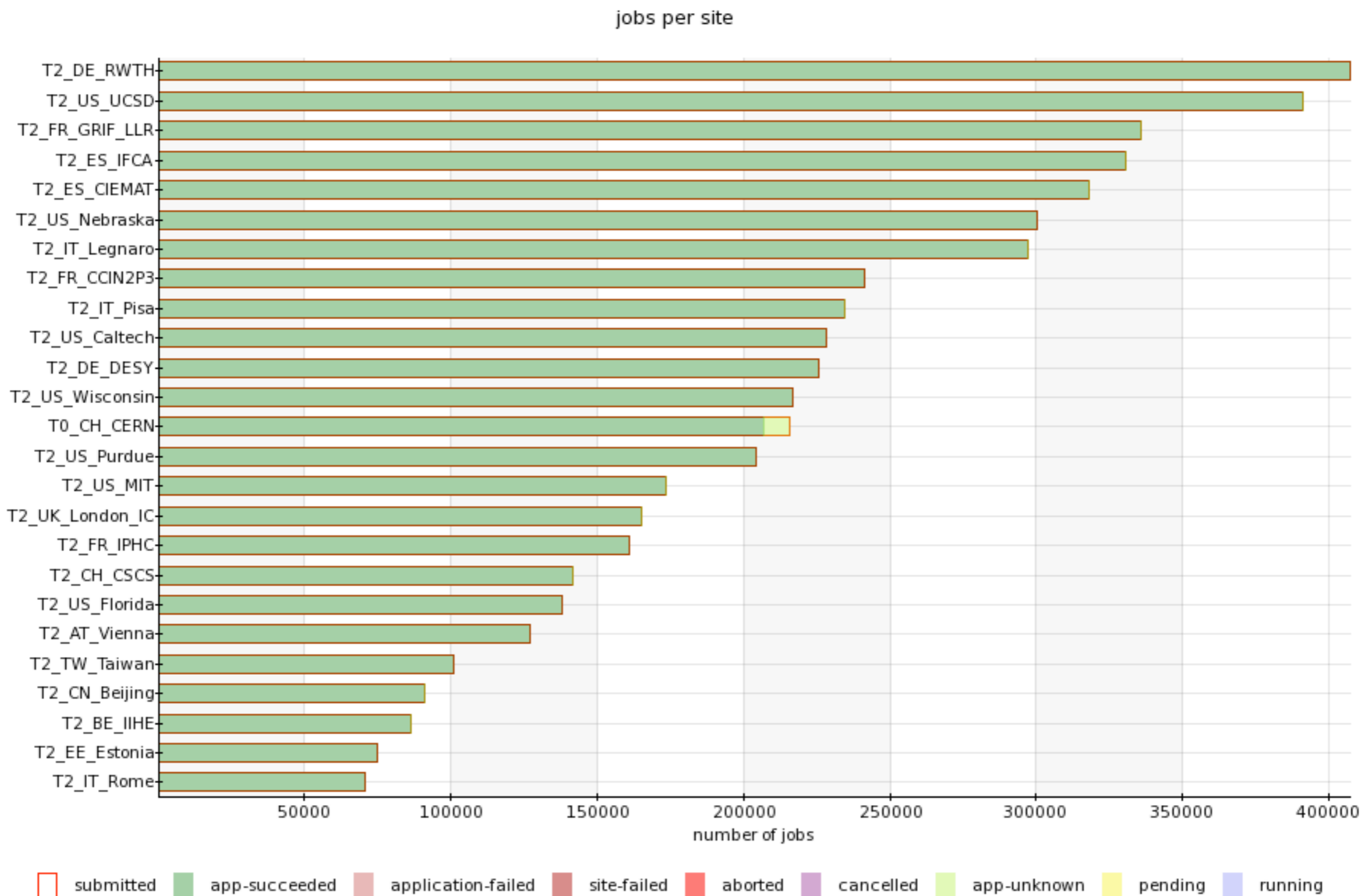


- | | | | | |
|-------------------|---------------|--------------------|---------------------|---------------------|
| T2_US_Purdue | T2_US_Caltech | T2_US_Nebraska | T2_US_Wisconsin | T2_US_UCSD |
| T2_US_Florida | T2_TW_Taiwan | T2_UK_London_IC | T2_IT_Legnaro | T2_IT_Pisa |
| T2_IT_Rome | T2_DE_RWTH | T2_FR_CCIN2P3 | T2_UK_London_Brunel | T2_DE_DESY |
| T2_FR_GRIF_LLQ | T2_US_MIT | T2_ES_CIEMAT | T2_BE_IIHE | T2_ES_IFCA |
| T2_UK_SGrid_RALPP | T2_FR_IPHC | T2_BE_UCL | T2_KR_KNU | T2_AT_Vienna |
| T2_IT_Bari | T2_RU_JINR | T2_FI_HIP | T2_BR_UERJ | T2_FR_GRIF_IRFU |
| T2_CN_Beijing | T2_EE_Estonia | T2_IN_TIFR | T2_TR_METU | T2_HU_Budapest |
| T2_PT_LIP_Lisbon | T2_UA_KIPT | T2_RU_SINP | T2_RU_ITEP | T2_UK_SGrid_Bristol |
| T2_RU_RRC_KI | T2_BR_SPRACE | T2_RU_IHEP | T2_PT_NCG_Lisbon | T2_RU_PNPI |
| T2_PK_NCP | T2_RU_INR | XT2_PT_LIP_Coimbra | | |

Total: 6,088 TB, Average Rate: 0.00 TB/s

Not shown: huge successful effort to implement T2-T2 links!

Totally successful analysis jobs at all CMS sites:



Site reports today will focus on plans for meeting our 2010 hardware deployment goals:

- ➔ 570 TB of data hosting (+40%)
- ➔ 7760 HS06 of CPU (+30%)

This is a relatively gentle ramp compared to past years.

But we really don't know how the computing model will perform in this first year of steady data-taking -- we should be prepared to deploy as much as budget allows this year.



Today: serving (US) physicists



The most “public” thing we do is serve the computing needs of physicists.

Another question for site reports: is the site-group association model “working”? What interactions are you having with physics groups? What are your observations on Analysis Operations so far?

A more technical issue: as stated in the past, resources that we deploy that are in excess of our commitment to “global” CMS are owned by US CMS, and US physicists should have priority over others on them. We owe it to our “constituents” to implement this.

- ➔ Discussion today led by Burt on how to do the implementation
- ➔ Let’s use the time to nitpick it and figure out how to work at as many sites as possible
- ➔ And then implement it ASAP?
- ➔ (A little tricky -- US people have equal access with everyone else to some resources, but priority access on another set....)



Serving US physicists: comment



I personally think that it is a priority to get “all” CMS data into the US for as long as it is feasible.

- ➔ “All” means anything anyone might have an interest in running on; I presume these are the secondary datasets.
- ➔ Probably the physics groups will be subscribing most of these, but we should fill in any gaps.
- ➔ We probably want this data at multiple sites, for redundancy.

I will try to coordinate this, along with CMS AnaOps, and ask sites and data managers to be flexible about data hosting.



Last year we saw the emergence of a viable SE based on Hadoop, and a strong desire by some sites to move towards it (and away from dCache) as quickly as possible.

We implemented a review and approval procedure that appears to have been a success and a template for the future. Today:

- ➔ Overview of Hadoop status, answers to review questions
- ➔ Overview of Lustre status -- is it ready to move into review process?

Bold question: given that there are alternatives that people are happy with, and that might be more suitable for T2 sites, do we want to use dCache at all?

- ➔ dCache was really designed as a disk cache with tape behind it
- ➔ But there is a lot of operational experience, lots of support available from FNAL, proof that it can be used at a facility on the scale of the FNAL T1
- ➔ Discussion today....

- ➔ Gotten gLexec working?
 - This *will* become a requirement of CMS, although time scale unclear
 - Let's get it all tested out sooner rather than later
 - As far as I know, works at Caltech, SPRACE, UCSD

- ➔ Made sure that you kept SiteDB, Savannah, OIM, etc. up to date on details of your site, including current pledge (1.5 MSI2k, 400 TB) and personnel?
- ➔ Made sure that I have your most up-to-date deployment numbers?
- ➔ Deployed perfSONAR on a suitable machine?
- ➔ Ever used this interesting tool to monitor dataset usage at your site?
- ➔ Used the dashboard to identify “site failures” and debug problems?
- ➔ Checked that you are publishing your space usage correctly in GIP?

Thanks in advance to:

- ➔ Our hosts at FNAL for handling all of the site logistics
- ➔ The OSG for organizing the all-hands meeting, and for supporting the platform that our CMS work rests on
- ➔ All of our speakers for doing the work to make good presentations
- ➔ All of you for coming to visit the FNAL mother ship
- ➔ Everyone involved with the T2 project -- you make my job easy.