



# Workshop Introduction

Ken Bloom  
US CMS Tier-2 Workshop  
March 3, 2009

## Expected LHC/CMS schedule:

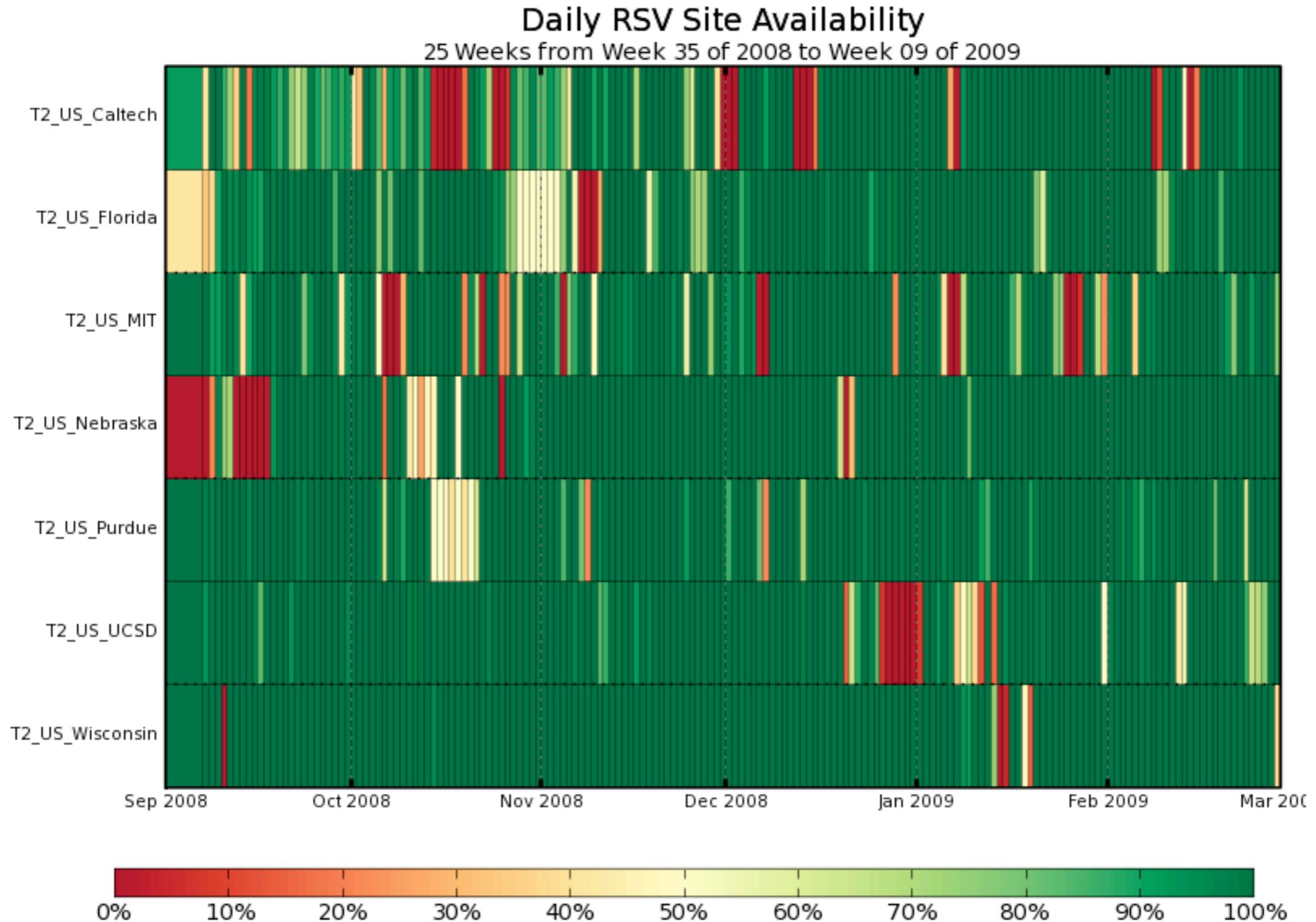
- ➔ Beams circulating again in late September, collisions in late October
- ➔ Data run for about a year (44 weeks), to accumulate  $200 \text{ pb}^{-1}$  at 10 TeV
  - A longer run than might have been expected a few months ago!
- ➔ Data size goes like running time, not integrated luminosity

Can expect a huge influx of users as soon as collisions start, with a lot of activity on “basic” datasets (e.g. high-momentum leptons)

Once we start, it's hard to see how it slows down, barring catastrophe.

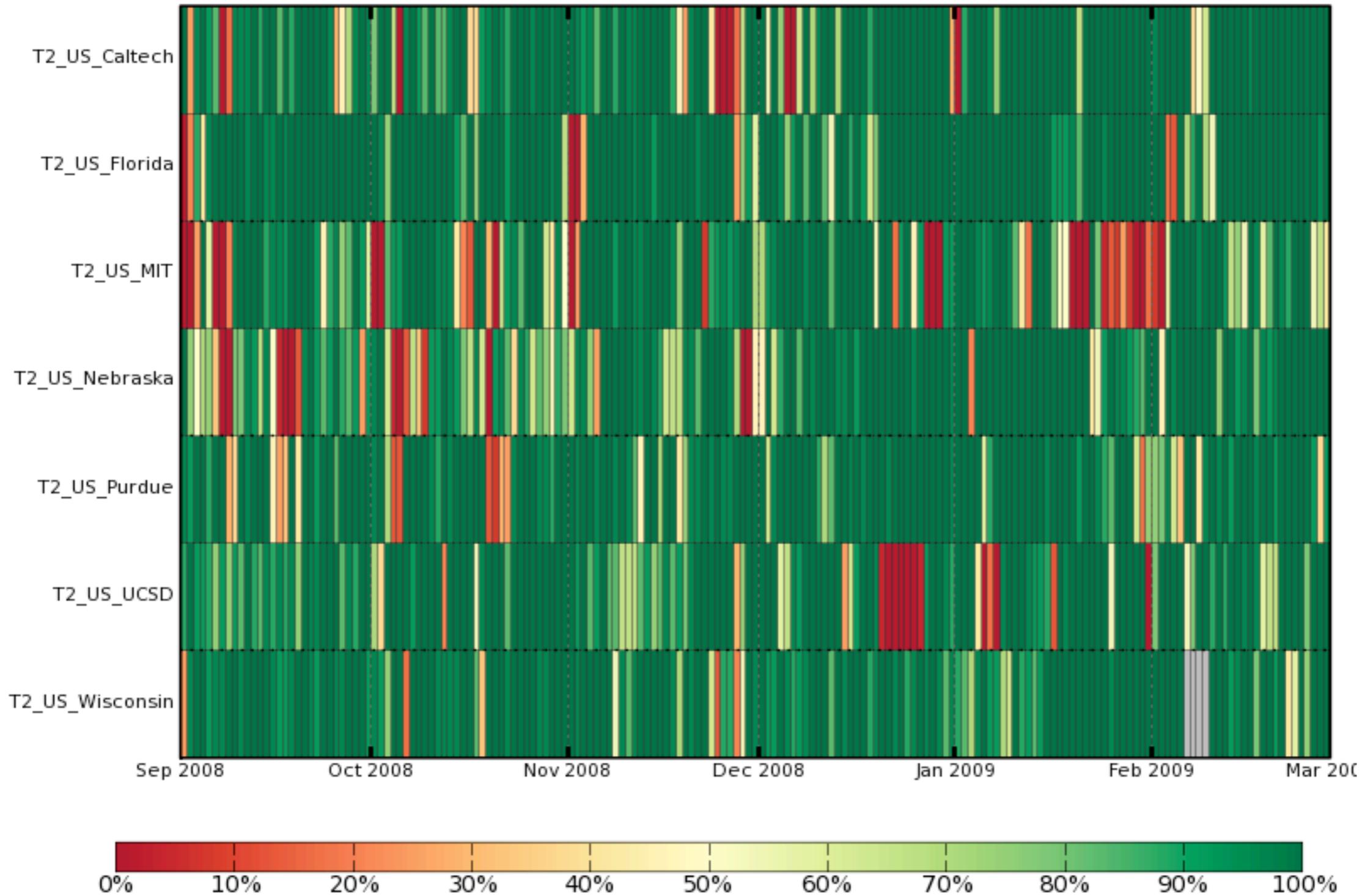
- ➔ Lots of user support!
- ➔ Lots of interaction with affiliated physics groups!
- ➔ Maintaining stable operations while not killing ourselves will be crucial.
- ➔ Essential questions for us: What are we doing well and what are we doing poorly in operations? How can we make improvements, e.g. do we need more support from FNAL/OSG? Why do different sites perform differently?

How much of this is just “normal” downtime?



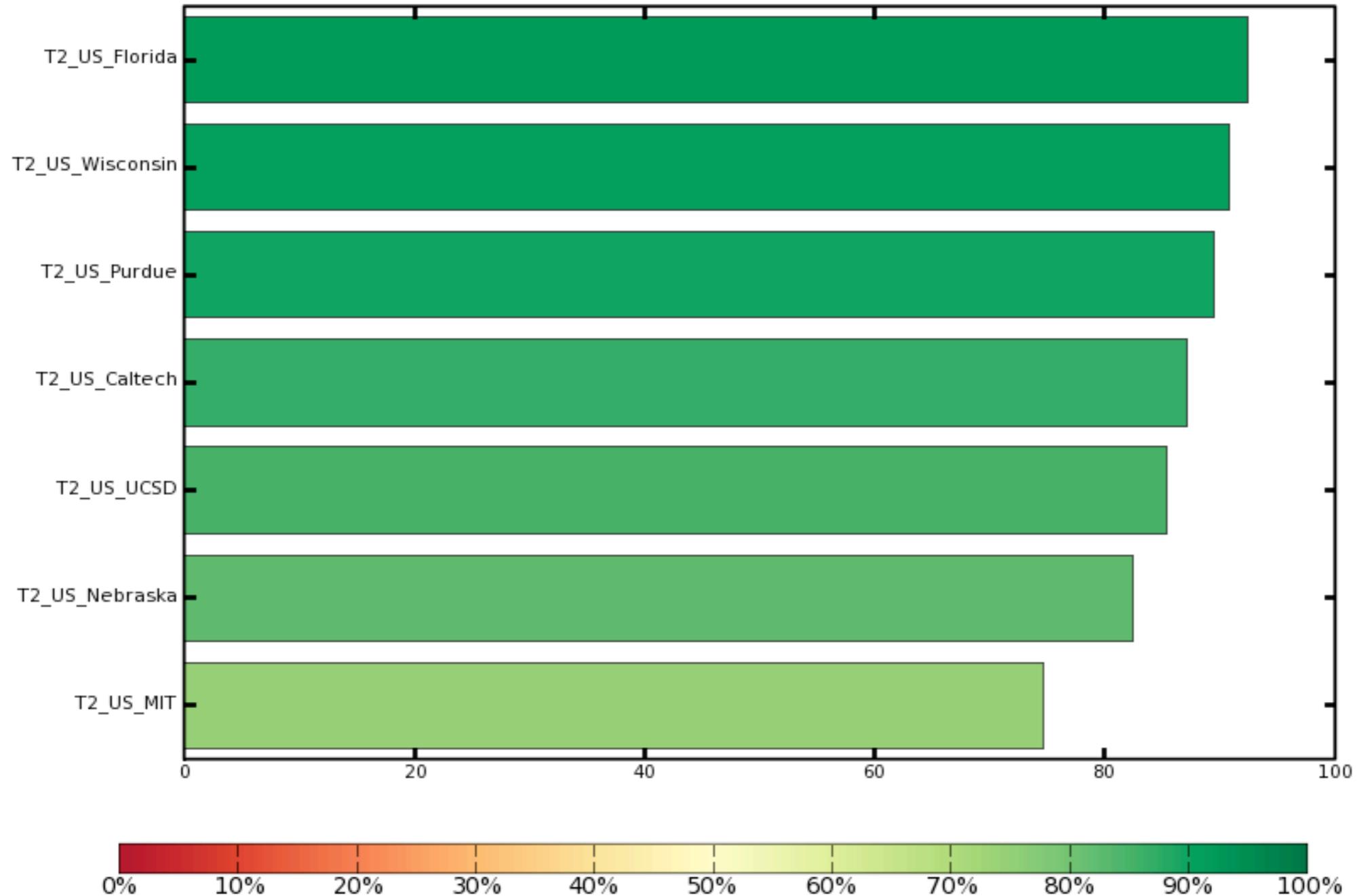
## Site Availability

181 Days from Week 35 of 2008 to Week 09 of 2009



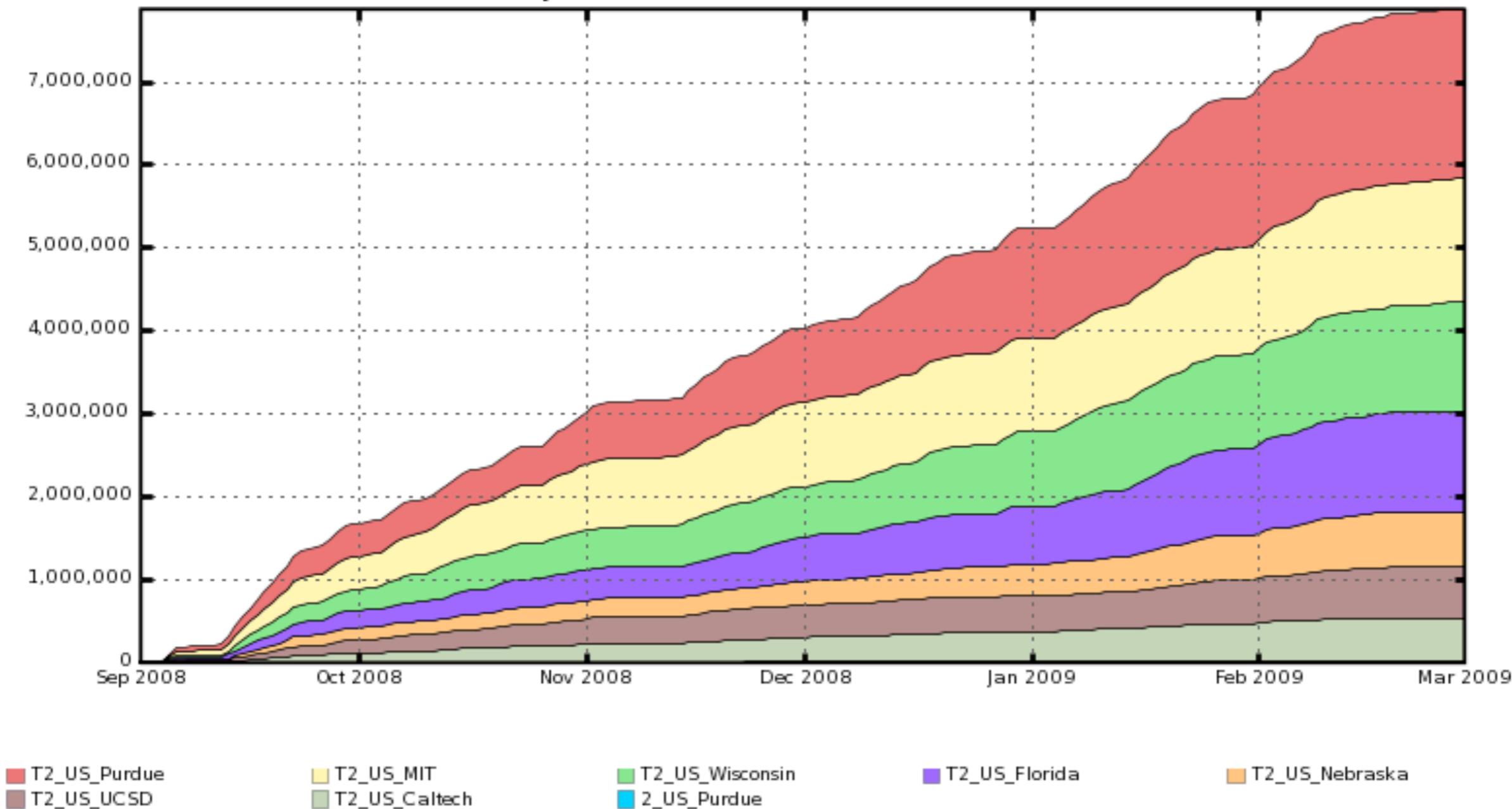
>90% is more than acceptable over the long term (especially when 2/7 of these days are weekends). What can best sites teach non-best sites?

**Site Availability, 2008-09-01 - 2009-03-01**



Does amount of time on production simply scale with site resources, or are there operational issues that can be addressed?

Time spent on successful jobs  
181 Days from Week 35 of 2008 to Week 09 of 2009

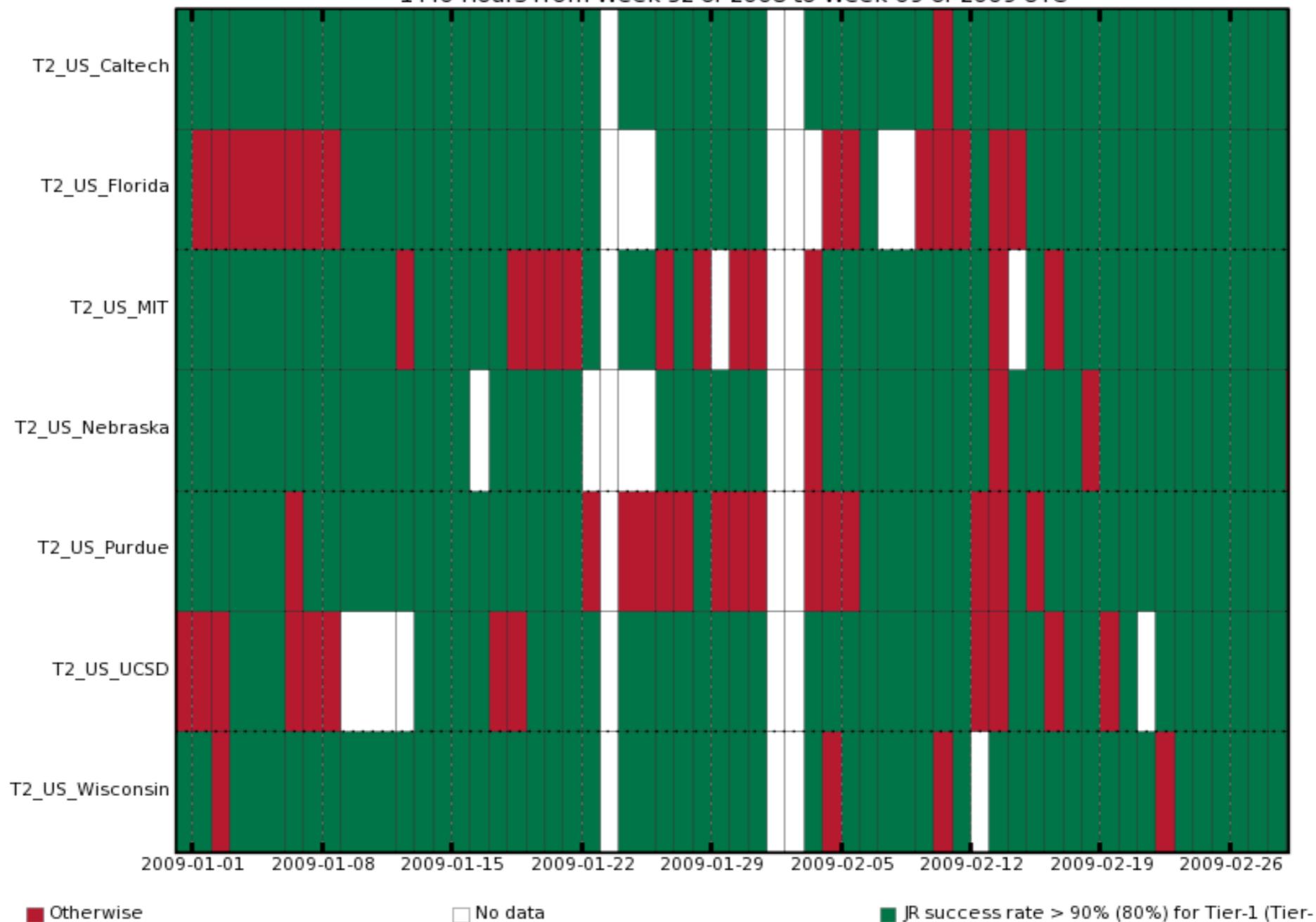


Total: 8,389 , Average Rate: 0.00 /s

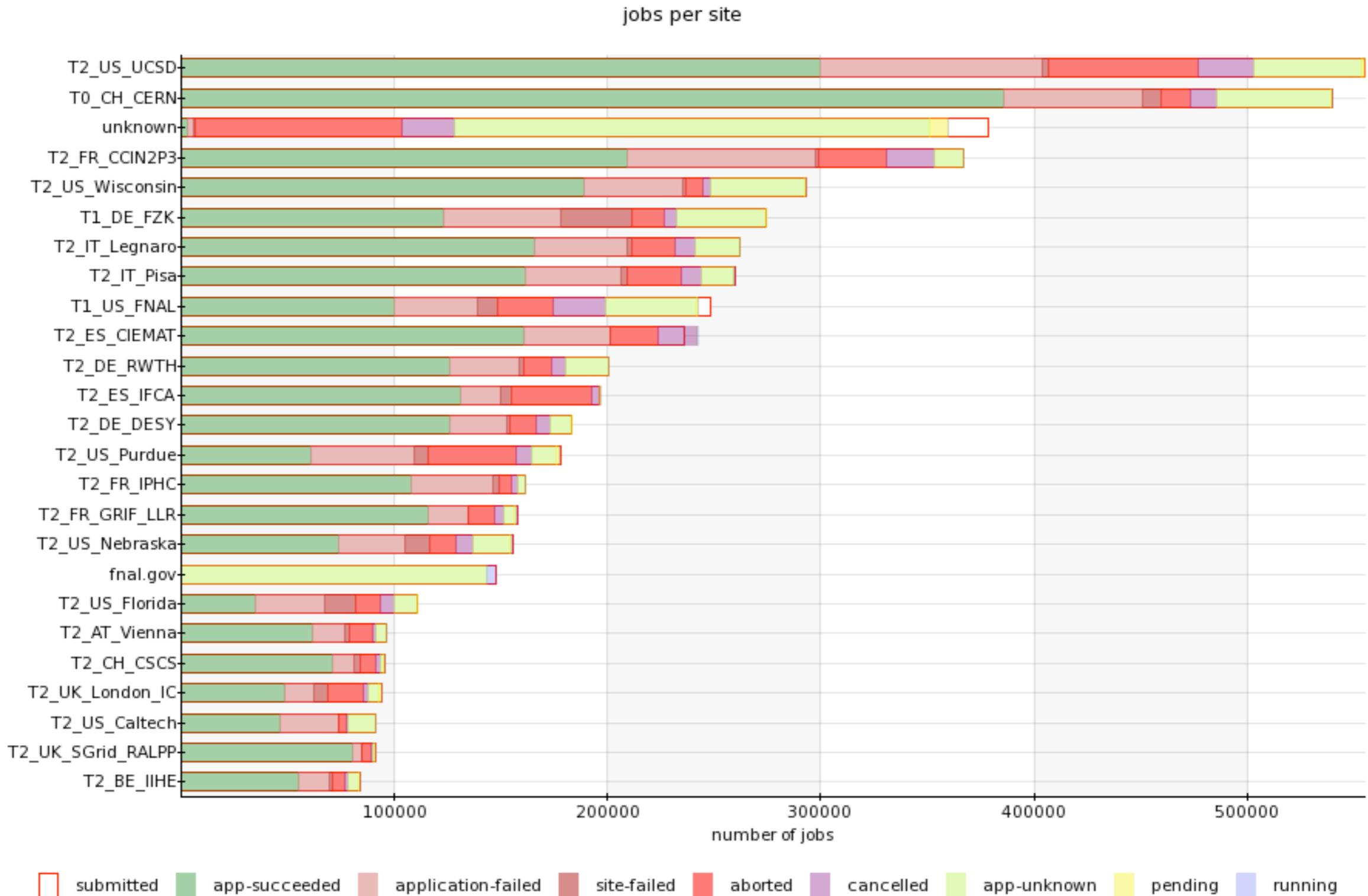
The job robot sends the same job on the same dataset, over and over. Why are some sites more successful than others in handling it? (Does this correlate with SAM etc.? I wish I knew...)

## Status of SiteComm JR

1440 Hours from Week 52 of 2008 to Week 09 of 2009 UTC



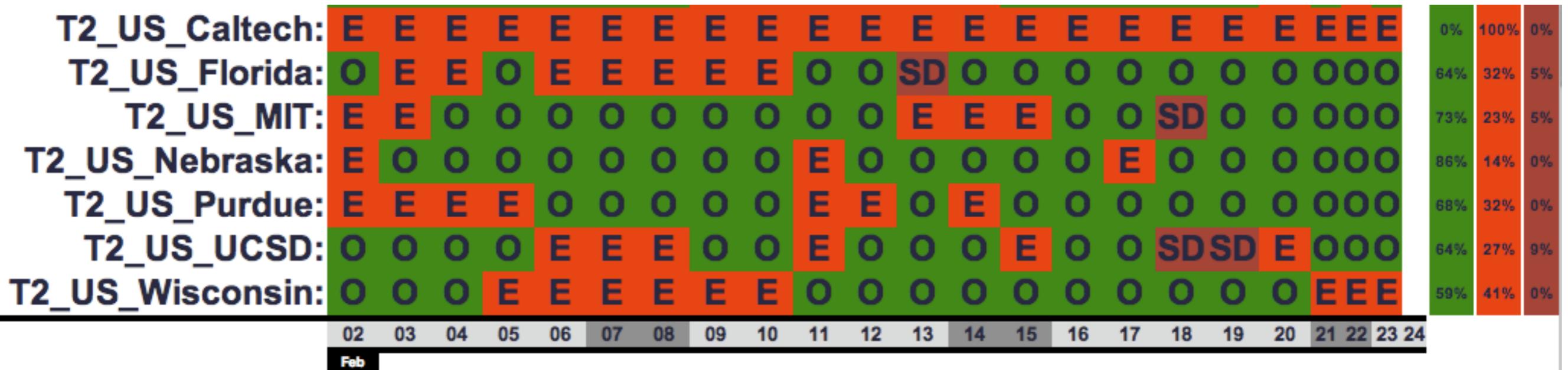
OK, there are many free parameters that drive this plot...



# Site readiness flag



This is by and large an assemblage of other metrics (SAM, robot, commissioned links). But we need to be attentive to it -- it is getting greater focus as an uber-metric.



This is not meant to take people to task -- it's consciousness raising. The real question: how can we help each other do better on all of these metrics?

When we started in on this in 2005, dCache/SRM was the only reasonable option as a storage system, in terms of scalability and supportability. Four years later, this may no longer be the case:

- ➔ Some sites already exploring alternative filesystems (Hadoop, Lustre)
- ➔ Bestman has emerged as an SRM alternative
- ➔ dCache presents its own set of challenges

Today, a discussion of the viability of various storage technologies.

- ➔ A change could give us new capabilities...but could also bring new challenges
- ➔ Moving away from dCache means cutting ties to the FNAL mother ship



- ➔ Considered installing the network-monitoring tools (e.g. perfSONAR) advocated by Rich Carlson et al.?
- ➔ Given any thought to contributing to a repository of shared monitoring code? (Oops, that's one I've done a poor job of follow-up on....)
- ➔ Deployed opportunistic storage for CDF or D0?
  - They and OSG appreciate it!
- ➔ Gotten gLexec working?
- ➔ Played with ERT to attract jobs to your site?
- ➔ Tried to install the Catalin/Subir Web-based job monitor?

Please check the BlockDownloadVerify results at your site:

- ➔ Tons of errors showing up, but it is speculated that these are configuration problems, not data corruption!
- ➔ Implementing dcls2 as discussed two weeks ago might help.
- ➔ Please update the twiki that James made as you learn things.
- ➔ Don't be shy about asking for help on/constructively criticizing the PhEDEx agent.

This is coupled to the issue of maintaining integrity of files produced by ProdAgent before they are archived to FNAL.

- ➔ Still under discussion; Metson/Evans considering adding a validation step at the end of ProdAgent

## Thanks in advance to:

- ➔ Our hosts at LLO 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 Louisiana -- nice weather, but not exactly centrally located for us
- ➔ Everyone involved with the T2 project -- all these computers don't do a darned thing without you!