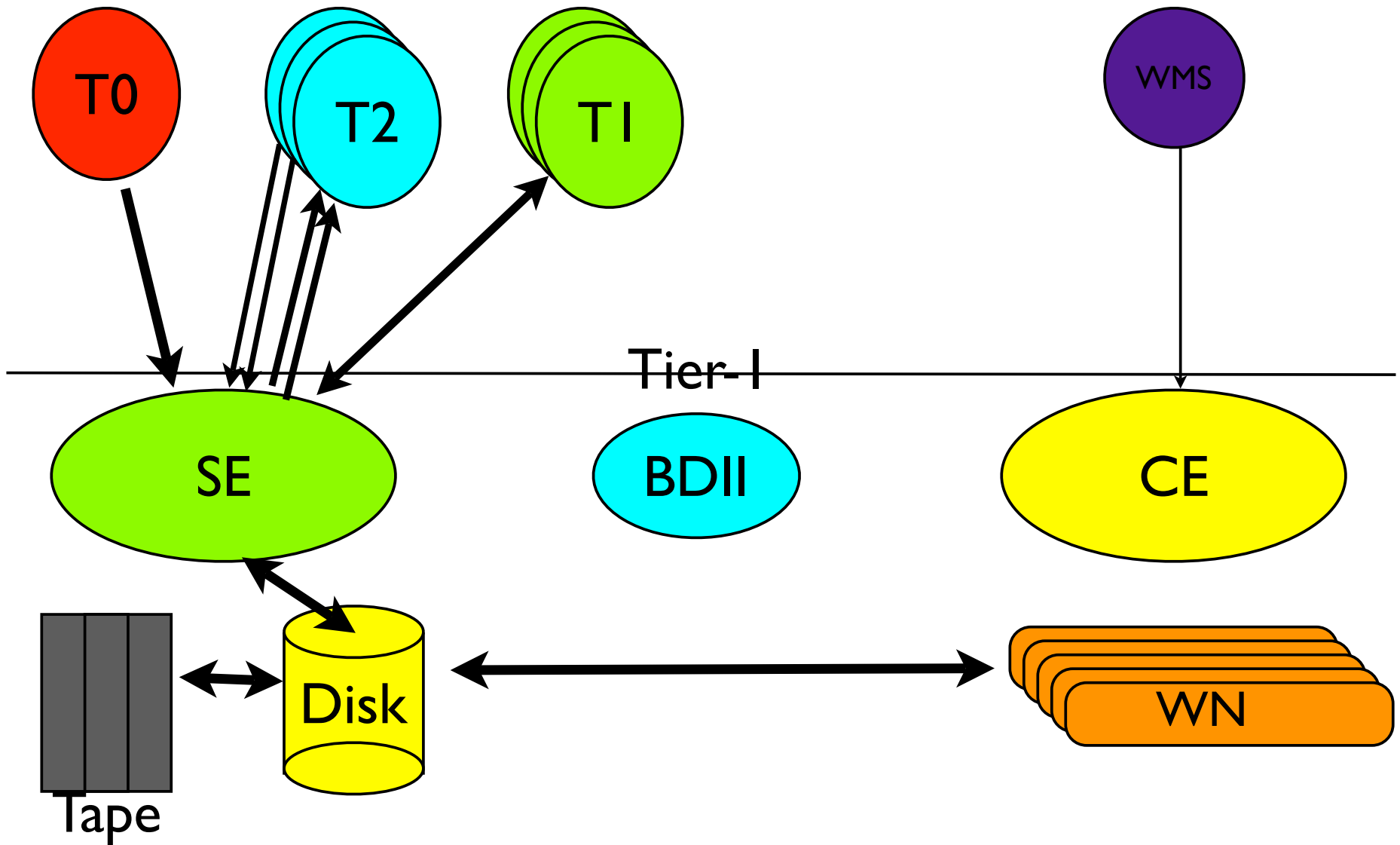


Tier-1 Issues

Ian Fisk
November 13, 2008

Tier-1



Feedback



There was not a great response to the request for feedback for this talk

- ➔ I believe it was less than 50%
- Some issues were common and some were site specific

The responses were generally from sites that run dCache

- ➔ Many of the issues will be about storage

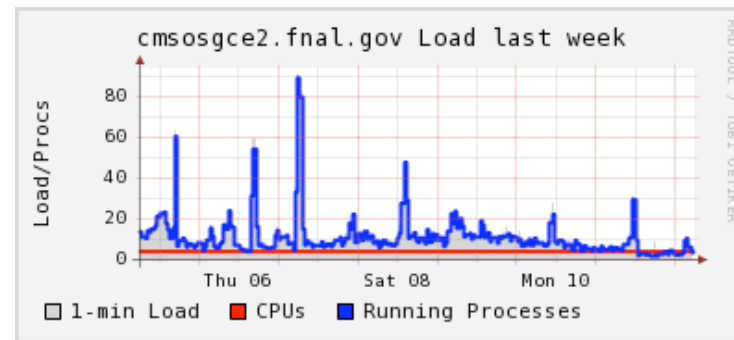
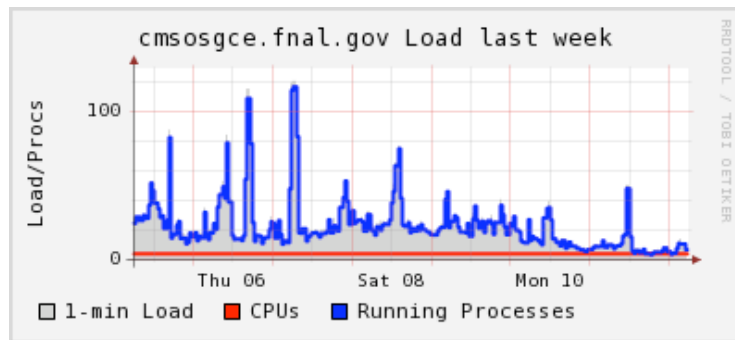
There will be time for other issues to be brought up

Computing Resources



So far scaling and stability of the current CE is reasonable

- ➔ Scalability is handled by adding computing elements
 - Possible to get reasonable balance across CEs through the WMS
 - At FNAL we currently have 4 CEs with 6500 batch slots
 - Average levels are OK, peaks are a little concerning



- ➔ Good Availability of CE resources with multiple instances
- ➔ Experiments appear to be able to utilize the resources on the sites
- ➔ Pilots may be helping

Tier-I Concerns



Concern from the Tier-I sites about the lack of clear targets

- ➔ Unclear how close we target job rates
 - TDRs generally have average job rates, but more specific per site expectations would be useful including elements like average and expected peak
 - Not clear if we're at 50% of scale or 100% of scale

Concern in periods throughout the year about the lack of utilization of Tier-I resources

- ➔ Not in stable operations yet

Execution of proper workflows at Tier-I's and restricting others

- ➔ Varying efforts to protect the Tier-I's

Given the length of time it's taken to reach stable service the schedule for the deployment and operations of CREAM should be carefully assessed.

Information Systems



The information systems appear stable

- ➔ Information inconsistencies or sites dropping out appear to be rare

General issues about the dynamic reports of priority and utilization

- ➔ May have been partially alleviated by the large scale adoption of pilot jobs by 3 experiments.

Storage



Most of the comments came regarding storage.

Concern about the scalability of the SRM implementation for dCache

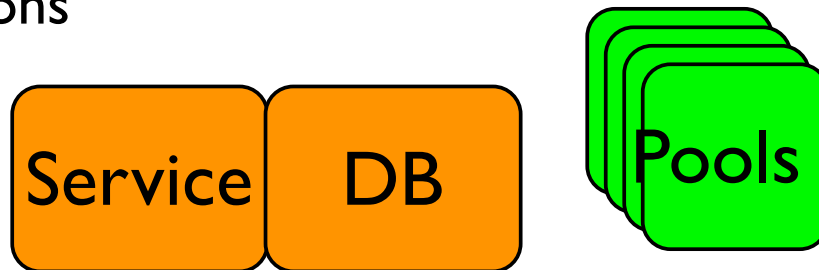
- ➔ Concern how SRM impacts pnfs performance
 - Many sites deploy a single dCache instance for all experiments. Heavy load on one can negatively impact another VO
 - Situation improved by the upgrade of Postgres, but concern about the scale of the final system.
 - Questions whether the situation would be improved by the migration to CHIMERA
 - Ability to roll back if needed

Storage



Concern regarding the overall scalability of the dCache SRM

- ➔ Currently we have a system that can stably handle about 1 Hz of SRM transactions



- One of the limitations in the authorization in the service
 - Efforts to share the load over several hosts haven't panned out yet

The current scale supports incoming and outgoing transfers from the Tier-1

- ➔ Provided the file size remains reasonable
- ➔ Doesn't support using the SRM interface to write data products from the local nodes
- ➔ There is a richness in the SRM functionality and it's not clear if it was widely used if the interface would scale

Storage Manageability



There are still core services in the storage system that connect be configured to fail over without human intervention

- ➔ Cannot be configured without a few

Regulation of SRM

- ➔ Common to all implementations
- ➔ FTS identifies the concept of fair share, but direct connects to SRM with lcg or osg tools will go into a queue with the same priority

Concern about the usability of the SRM logs to diagnose why a problem was occurring

- ➔ May be improved in dCache 1.9

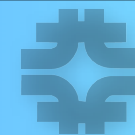
Storage Targets



This issue of sites not knowing how far they are for the target metric for storage as well

- ➔ How many SRM transactions is the goal?
- ➔ Transfer rate from CERN to Tier-1 is reasonably well specified
 - Tier-1 to Tier-1 and Tier-1 to Tier-2 can be driven by burst to the peak and average is less understood.

Usage of Archival Storage



A general issue across storage technologies and experiments is the effective demonstration are efficient restoration of files from tape

- ➔ Caching percentages vary experiment to experiment, but for large scale reprocessing some recovery from tape may be needed.
- ➔ Need to work on aggregating files onto tapes.
 - Creation of file families dynamically with reasonable granularity
 - Prediction of what groups of data need to be accessed together.
- ➔ Triggering prestaging on a large scale
 - Reasonable methods for ensuring data isn't immediately staged out

System Recommendations



There was a request for configuration recommendations designed to get to a particular scale for services

- ➔ For FTS, LFC, CE, SRM/dCache, documentation specifying the recommended configuration for reaching the availability/reliability targets set by WLCG

Some sites report off-hour support is best effort

- ➔ Can't bring the whole team nor can the off-hour support work continuously

Communication with Tier-1s



Consistency of communication tools could be improved between experiments and Tier-1s

- ➔ Some combination of mailing lists, mails, Savannah and GGUS

Due to the security incident in one of the CMS machines, CMS locked down the agenda servers

- ➔ Looking at restricting TWIKI access
- ➔ Site administrators don't always have CERN accounts