

Grid Storage Monitoring



6th February 2007

Flavia Donno
IT/GD, CERN



A few thoughts

- "I have been struggling in ATLAS with disks at the sites being filled up with Monte Carlo data at a higher rate than expected...One control I didn't have and still don't have is the usage of the disk space from within ATLAS. I just could not understand how the 12 TB of disk pool for ATLAS at NIKHEF could be 100% full."
- "...We could look in the DPM database and list the owner_uid / SUM(filesize) and see who was actually using the disk. It turned out that 25% of all disk space is used by ATLAS users and not ATLAS Production."
- "...I would like to know, say per VO group what the actual usage of the disk space is at a site ..."

Kors Boss, 2 Feb 2007

- Furthermore:
 - How can experiments find out about the storage resources they have allocated at a site (is the promise respected ?).
 - Are site configured as they were told ?
 - What is the status of the storage resources ?
 - ...



Can we do the following ?

- Use implementation specific tools to collect this information
 - *Sensors ?*
- The information can be made available through SAM or GridView
- The experiment dashboard can be feed by GridView or SAM DBs.
- Probably not all information available at the moment.
 - *We can start to investigate what information is available and write GridView or SAM "feeders".*
- Good targets at the moment seem to be dCache and DPM
- What do you think ?