



Tier-1/Tier-2 Workshop
"LCG Tier-0 Summary"

James Casey, IT-GD, CERN
Karlsruhe, 19th October 2005





Storage (1/3)



- During the throughput phase of SC3, CASTOR2 ran stably serving files from disk
- For the service phase, there have been major instabilities and downtime
- Several things changed for service phase
 - Internal pools were added for experiments to give more space
 - It was reconfigured to use internal pools for staging from tape and replication to the wan pools for transfer
 - The tape part was exercised at a much higher rate



Storage (2/3)



- Problems seen in CASTOR interaction with the database
 - DB server crashed several time due to running out of resources during GC procedure
 - Query to lookup location of file are consuming huge amount of resources and taking very long
- FTS had not been exercised doing pre-staging and it's interaction with the SRM had not been expected
 - cancelling the request did not cause the query for the tape request to be cancelled
- A problem with iptables modules in the kernel stopping traffic on loopback interface
 - Led to 451 TIMEOUT errors in gridftp
- From the use perspective this led to lots of transfers piling up on the SRM/gridftp causing timeouts and failed transfers



Storage (3/3)



- Problems are being fixed as they are seen
 - This has led to a lot of interventions to upgrade software
- Database server node is quite a low-spec machine
 - Replacing today with a 2-CPU node to try and improve performance
- We hope that the service can now stabilize and we can get reliable transfers running again



Network



- The first lot of new networking hardware has arrived at CERN
 - Force10 10G bit routers and HP switches
- We will start to build up the infrastructure for the OPN now based on this
 - This should relieve many of the routing issues we've seen at CERN during the last few months
 - Non-SC hosts will route to SC nodes over GEANT and through the CERN campus network
- We plan to deploy this the weekend of 29th October
 - This would be a good time to reconnect high-bandwidth links that have been removed due to problems e.g. SARA



FTS and LFC



- We are testing the latest version of FTS
 - part of gLite 1.4
- This provides
 - srm copy functionality
 - Better BDII integration
- We would like to roll it out at the end of October
 - Clients are now backwards compatible between 1.3 and 1.4 installs, so upgrade of all sites should be a lot smoother



Summary



- Major instability in CASTOR2 seen during Service Phase
 - Many bugs found in all areas of the system
 - Being systematically patched as found
 - Upgrade of some hardware components to try and alleviate problems
- Network upgrade will put in place start of the final LHC networking infrastructure
- New version of FTS supporting srm copy is ready
- An intervention before second phase of SC3 will put these components in place





Service Analysis



- Effort has started to examine services in order to provide MOU levels of availability
 - Decide on criticality of services
 - Based on scope and impact of unavailability
 - Recommendations for
 - Hardware
 - Fail-over and load-balancing software
 - Recovery
- that are appropriate for the criticality level of the service
- Work being tracked on SC Wiki
 - Hopefully it can be reused by other sites looking to deploy services in a high-availability mode