LCG Storage Workshop

"Service Challenge 2 Review"

James Casey, IT-GD, CERN

CERN, 5th April 2005

Overview

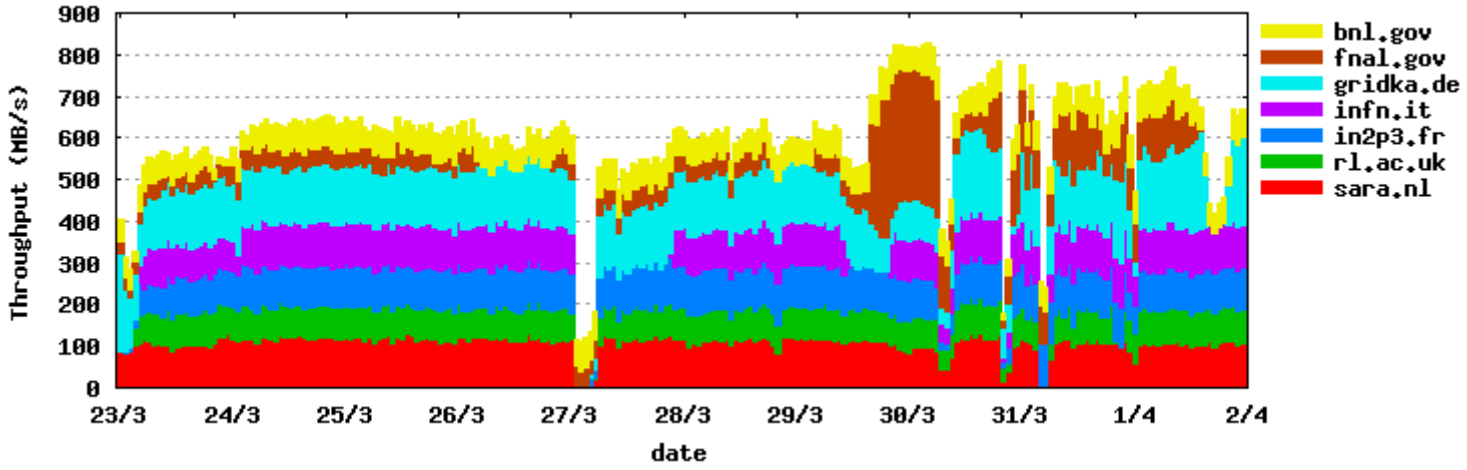
- Reminder of
  - Targets for the Service Challenge
  - Plan and timescales
- CERN Tier-0 Configuration
- Tier-1 Configuration
- Throughput phase of Service Challenge 2
  - Transfer Software
  - Monitoring
- Outages and problems encountered
- SC2 single site tests

Service Challenge Summary

- "Service Challenge 2"
  - Throughput test from Tier-0 to Tier-1 sites
  - Started 14th May
- Set up Infrastructure to 7 Sites
  - NL, IN2P3, FNAL, BNL, FZK, INFN, RAL
- 100MB/s to each site
  - 500MB/s combined to all sites at same time
  - 500MB/s to a few sites individually
- Goal : by end March, sustained 500 MB/s at CERN

SC2 met its throughput targets

- >600MB/s daily average for 10 days was achieved - Midday 23rd March to Midday 2nd April
  - Not without outages, but system showed it could recover rate again from outages
  - Load reasonable evenly divided over sites (give network bandwidth constraints of Tier-1 sites)



Division of Data between sites

Site	Average throughput (MB/s)	Data Moved (TB)
BNL	61	51
FNAL	61	51
GridKA	133	109
IN2P3	91	75
INFN	81	67
RAL	72	58
SARA	106	88
TOTAL	600	500

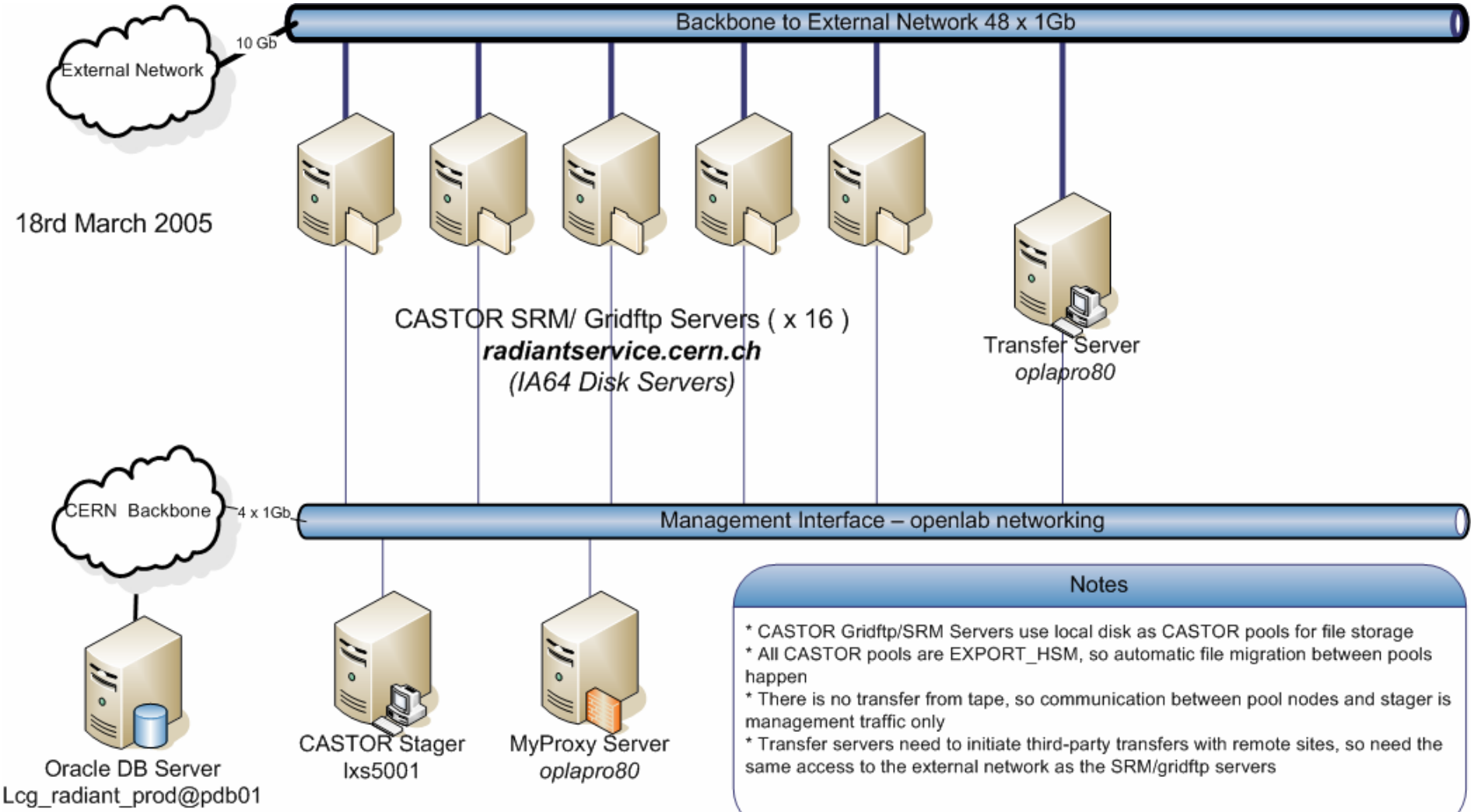
CERN Tier-0 Configuration (1/3)

- Service Challenge 1 was on an experimental hardware setup
  - 10 HP IA64 nodes within the opencluster
  - Not a standard service
    - No LEMON monitoring
    - No operator coverage
    - No standard configuration
- Plan was to move to a "standard" configuration
  - IA32 "worker nodes" running CASTOR SRM/gridftp
  - Still serving data from local disks
- Also to move to "production" networking
  - The opencluster is a test and research facility
  - Again, non-monitored network hardware

CERN Tier-0 Configuration (2/3)

- This didn't all go to plan…
  - First set of 20 IA32 worker nodes weren't connected to right network segments
    - They were installed, configured and tested before this was discovered
  - Replacement nodes (20 IA32 disk servers) were supplied too late to get installed, configured and tested before SC2
- Had to fallback to openlab IA64 nodes for SC2
  - These weren't connected to the production network on their inward facing interface (towards CERN LAN)
  - Added 10 extra IA64 nodes to system – 16 in total were used for data transfer in SC2

CERN Tier-0 Configuration (3/3)



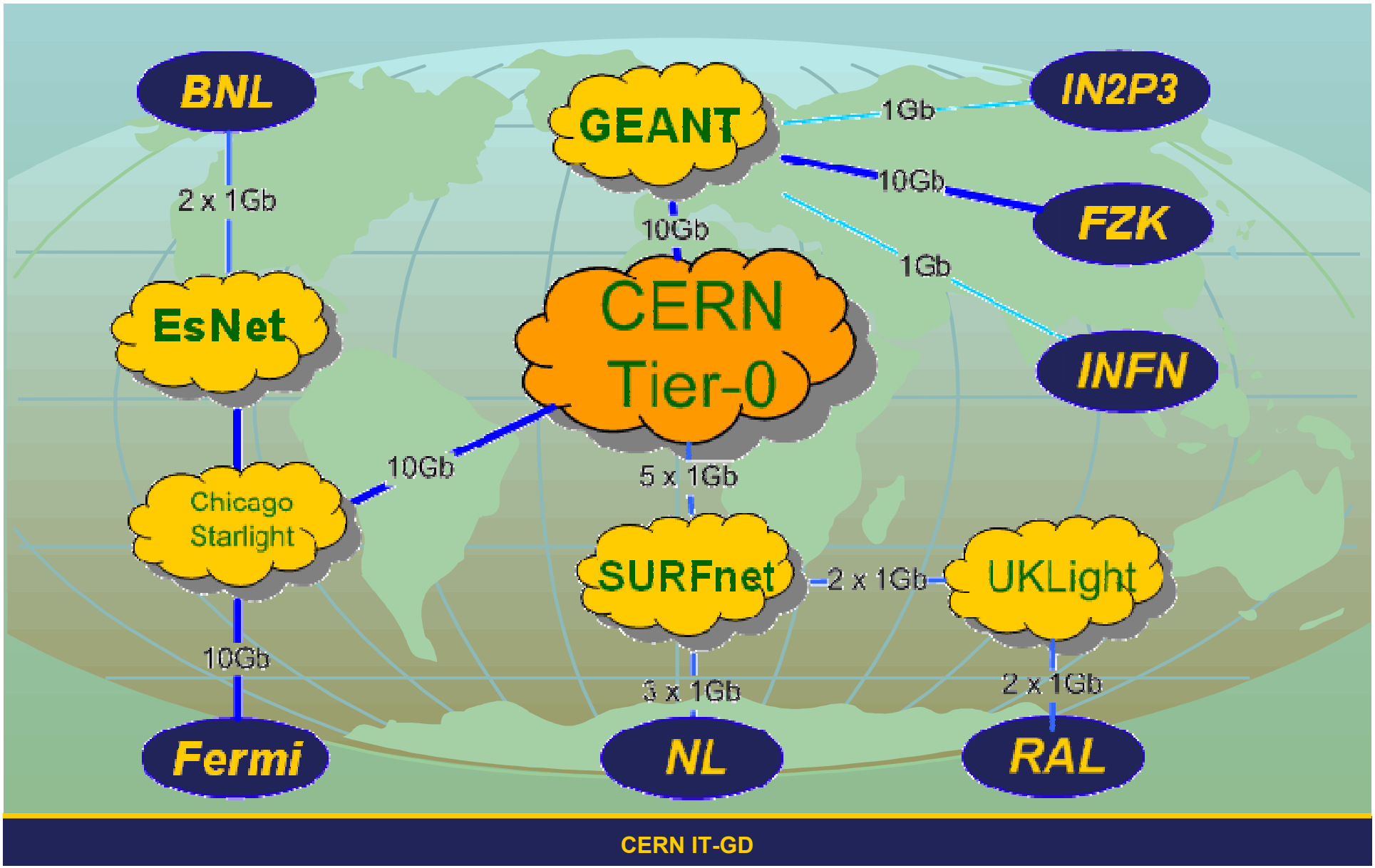
Tier-1 Storage Configuration

- Small set of storage configurations
  - Most sites ran "vanilla" Globus gridftp servers
    - SARA, INFN, IN2P3, FZK
  - The rest of the sites ran dCache
    - FNAL, BNL, RAL
- Most sites used local or system-attached disk
  - FZK used SAN via GPFS
  - FNAL used production CMS dCache, including tape
- Load-balancing for most plain gridftp sites was done at the RADIANT layer
  - INFN deployed "n-1" DNS alias – highest loaded machine was replaced in alias every 10 minutes
  - Alleviated problems see with pure round-robin on gridftp servers

Tier-1 Network Connectivity

- Sites are in the middle of their upgrade path to LCG networking
  - 10Gb will be the norm – now it's more like 1Gb
  - All this is heavily tied to work done in the LCG T0-T1 networking group
- Most sites were able to provided dedicated network links
  - Or at least links where we could get most of the supplied bandwidth
  - IN2P3, BNL still were on shared links with a bit more congestion
    - Needed to be dealt with differently
    - Upped the number of concurrent TCP streams per file transfer

Tier-1 Network Topology



Transfer Control Software (1/3)

- LCG RADIANT Software used to control transfers
  - Prototype software interoperable with the gLite FTS
  - Plan is to move to gLite FTS for SC3
    - Initial promising results presented at Lyon SC Meeting in Feb.
    - More details in LCG Storage Workshop tomorrow
  - Run on a single node at CERN, controlling all Tier-0 to Tier-1 "channels"
  - Transfers done via 3rd-party gridftp
- 'radiant-load-generator' was used to generate transfers
  - Configured for each channel to load-balance where appropriate
    - Specifies number of concurrent streams for a file transfer
    - This was normally =1 for a dedicated link
  - Ran from cron to keep transfer queues full

Transfer Control Software (2/3)

- Control of load on a channel via number of concurrent transfers
- Final Configuration:

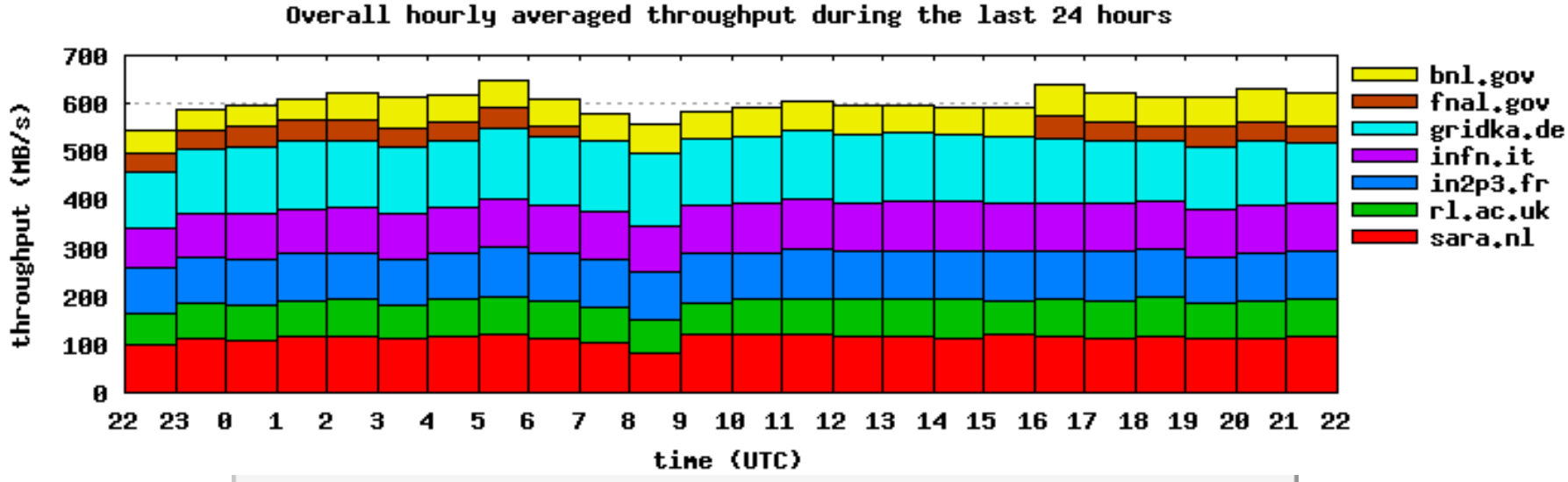
```
#Chan : State : Last Active       :Bwidth: Files: From    : To
FZK   :Active :05/03/29 15:23:47 :10240 :5     :cern.ch :gridka.de
IN2P35:Active :05/03/29 15:16:32 :204   :1     :cern.ch :ccxfer05.in2p3.fr
IN2P33:Active :05/03/29 15:20:30 :204   :1     :cern.ch :ccxfer03.in2p3.fr
INFN  :Active :05/03/29 15:23:07 :1024  :8     :cern.ch :cr.cnaf.infn.it
IN2P32:Active :05/03/29 15:21:46 :204   :1     :cern.ch :ccxfer02.in2p3.fr
FNAL  :Inactiv:Unknown           :10240 :0     :cern.ch :fnal.gov
IN2P3 :Active :05/03/29 15:23:40 :204   :1     :cern.ch :ccxfer01.in2p3.fr
IN2P34:Active :05/03/29 15:18:00 :204   :1     :cern.ch :ccxfer04.in2p3.fr
BNL   :Inactiv:Unknown           :622   :24    :cern.ch :bnl.gov
NL    :Active :05/03/29 15:22:54 :3072  :10    :cern.ch :tier1.sara.nl
RAL   :Active :05/03/29 15:23:09 :2048  :12    :cern.ch :gridpp.rl.ac.uk
```

Transfer Control Software (3/3)

- RADIANT controlled transfers were pure gridftp only
  - SRM components did not get developed in time
  - FNAL and BNL did SRM transfers by pulling from their end
    - FNAL used PhEDEx to pull data
    - BNL used simple driver scripts around srmCopy command
    - BNL started some radiant-controlled gridftp transfers too
  - Used the exact transfer nodes as the rest of the sites (radiantservice.cern.ch)
- SRM interactions helped debug issues with dCache SRM and DNS load-balanced SRM/gridftp servers
  - Fixed java DNS alias problems seen at FNAL
  - Did not work smoothly for most of the challenge with FNAL mode of operation
    - Main problem solved and fix in place by end of throughtput phase
    - How to do srmCopy's with 1000s of files is still not well understood

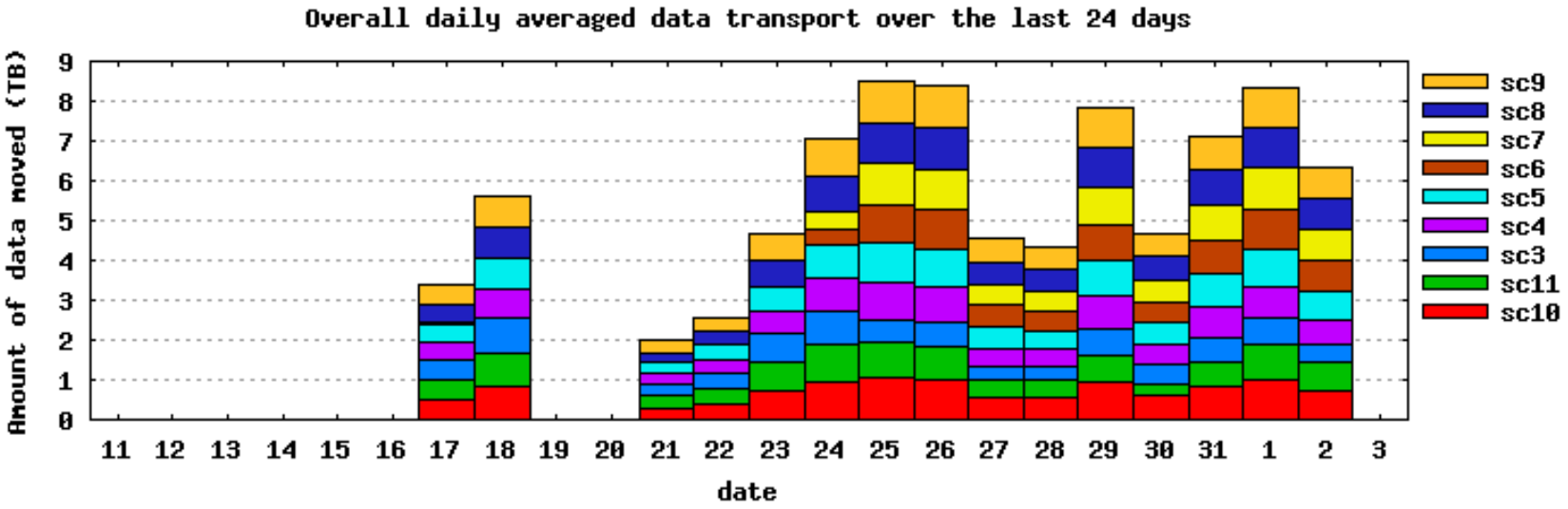
Monitoring @ CERN

- MRTG Graphs of network usage out of cluster
- LEMON monitoring of cluster
  - CPU usage
  - Disk usage
  - Network usage
- Gridftp logfile monitoring



Tier-1 Monitoring

- Gridftp Monitoring (http://challenge.matrix.sara.nl/SC/)
- Main monitoring tool used during the SC2
  - Hourly throughput per site
  - Hourly throughput per host
  - Daily throughput per site
  - Daily throughput per host



Service Outages (1/2)

- Progress page kept in SC wiki of
  - All tunings made to the system
  - All outages noted during the running
  - Any actions needed to cleanup and recover service

    <http://service-radiant.web.cern.ch/service-radiant/wiki/ow.asp?ChallengeSC2Progress>
- No real 24x7 service in place
  - Manual monitoring of monitoring webpages
  - Best-effort restart of service
  - Also at Tier-1 sites – problems communicated to service challenge teams, but this was not a 24x7 coverage

Service Outages (2/2)

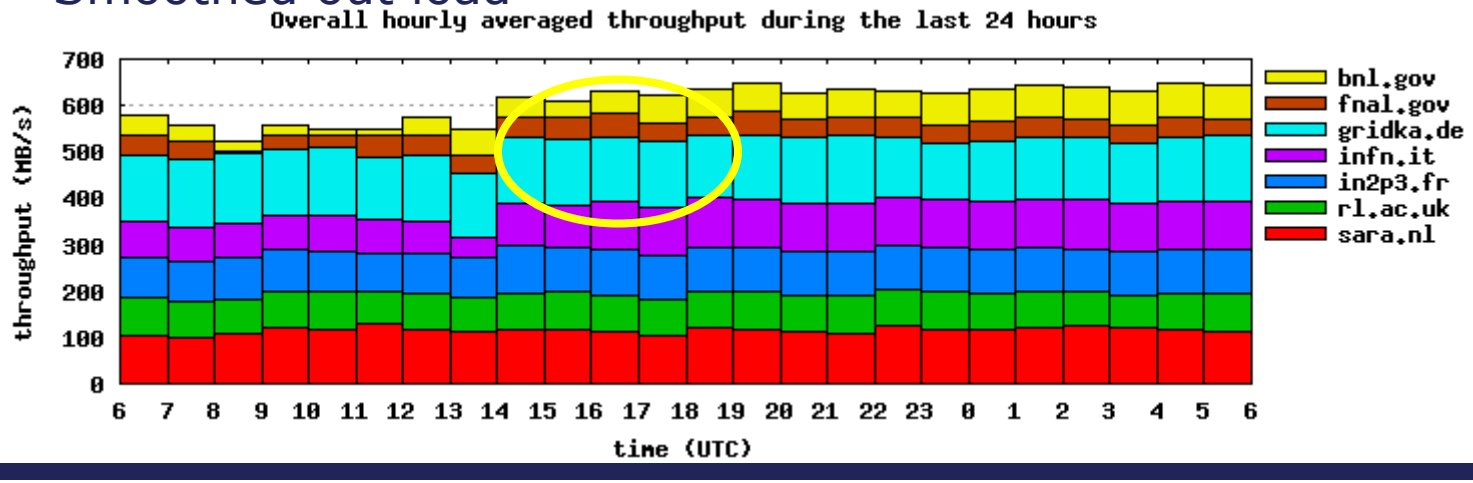
- Capacity in the cluster meant that we could recover from one site not being active
  - Other sites would up their load a bit automatically due to gridftp stream rates increasing
  - Only thing that killed transfers were CERN outages
- We did not do any scheduled outages during the SC
  - No procedures for starting a schedule outage
  - If we had done one to move to managed network infrastructure, it would have removed some of the scheduled ones

Outage Breakdown - CERN

- Mxproxy instability
  - Locks up under high load
    - Understood by developers of myproxy
    - Can be handled by watchdog job
  - Particular problem on restart after network outage
- CERN LAN Network outages
  - Had 2 long-ish outages (~12 hours)
  - Issue with being on non-managed network hardware
- Database quota limit hit
  - Tablespace was being monitored but not quota
  - Quota monitoring added
- Database load problems
  - Caused intermittent drops in throughput due to new jobs not being scheduled
  - In-memory hob queues in the transfer agents meant these we're a big problem

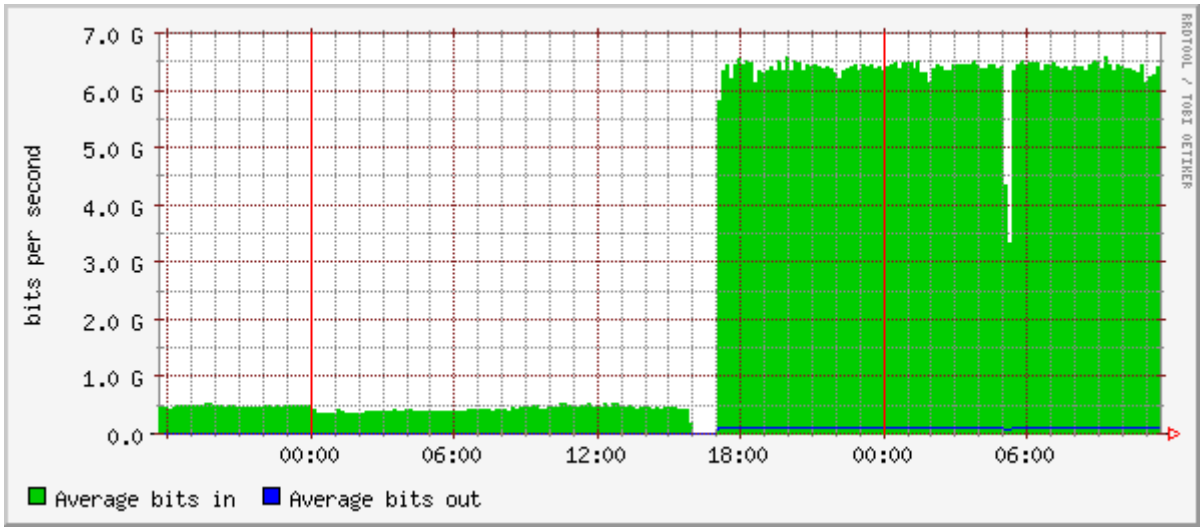
DNS load-balancing of gridftpd

- An issue raides at Lyon SC Meeting
- "load-balanced" gridftp servers
  - Not really load-balanced with round-robin – no awareness of state or behaviour of other nodes
  - Can end up with all transfers going to one host, with all others idle
- CNAF put in place "worst one out" load-balancing
  - Heaviest loaded node is taken out of alias every 10 minutes
  - Smoothed out load



Individual site tests

- Being scheduled right now
  - Sites can pick days in next two weeks when they have the capacity
  - 500MB/s to disk
  - 60MB/s to tape
- FNAL is running 500MB/s disk tests right now



Summary

- SC2 met it's throughput goals
  - An improvement from SC1
- We still don't have something we can call a service
  - But monitoring is better
  - We see outages when they happen, and we understood why they happen
    - First step towards operations guides
- Some advances in infrastructure and software will happen before SC3
  - gLite transfer software
  - SRM service more widely deployed
- We have to understand how to incorporate these elements