



GridPP
UK Computing for Particle Physics

UK Status for SC3

Jeremy Coles

GridPP Production Manager: j.coles@rl.ac.uk

Service Challenge Meeting

Taipei 26th April 2005



GridPP
UK Computing for Particle Physics

GridPP sites

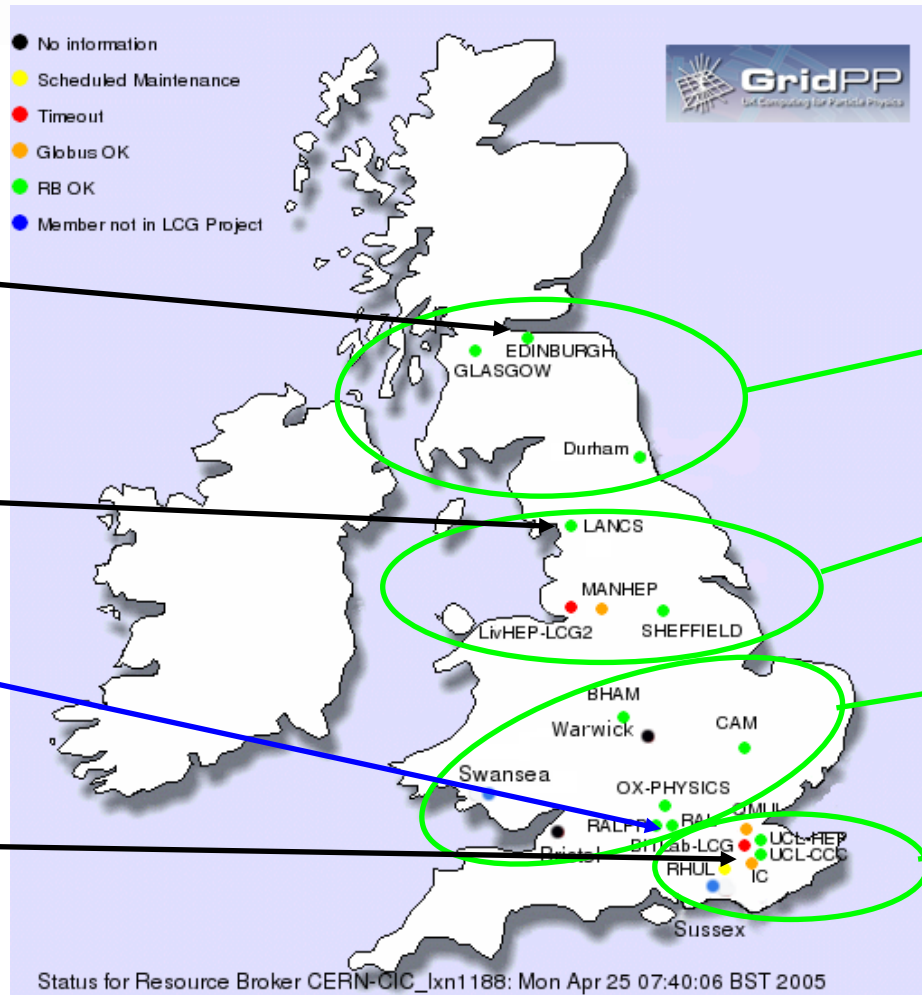
SC3 sites

LHCb

ATLAS

Tier-1

CMS



Tier-2s

ScotGrid

NorthGrid

SouthGrid

London Tier-2



- Mailing list established GRIDPP-SC@jiscmail.ac.uk
- Reviewed dCache usage and issues at a recent GridPP storage workshop:
<http://agenda.cern.ch/fullAgenda.php?ida=a052046>
- SC3 tutorial at Imperial college planned for early May (13th?)
- Small semi-related work on putting SE gridftp logs and dCache billing information into R-GMA. Some preliminary work is here:- <http://goc.grid-support.ac.uk/gridsite/monitoring/StorageElement/SEView.php?ServerHost=lcgse01.gridpp.rl.ac.uk>
- Tier-2s generally well engaged with experiments at respective sites
- SRM status reviewed at biweekly update calls
- DPM being evaluated by GridPP Storage Group who already support dCache deployment at Tier-2s
- SC3 weekly call to start in mid-May



GridPP
UK Computing for Particle Physics

Tier-1A

Andrew Sansum



SRM Status

- Already run a production dCache SRM (CMS, ATLAS - LHCb soon). Deployed second SRM for SC2 but only tested gridftp (Radiant did not support SRM)
- During SC3, intend to use production SRM with existing hardware - phase out special SC SRM.
 - Hope to dual attach, **but need fix from DESY to support multiple interfaces on gridftp service.**
 - FTP protocol returned i.p. for data channel on wrong interface. (Can we bind the gridFTP doors?). Pool nodes tried to contact external interfaces of GridFTP Nodes.
 - Developing a “plan B” to inject SC UKLIGHT traffic into production network.
- Tape service tested during SC2, but stability problems on Radiant hampered measurements. Tape SRM working fine, but bandwidth to tape still too low (expected). New disk cache for tape service ordered, **expected in production by early June.** Expect to meet target rate.
- During recent tests the ADS was accepting data over the network at about 100Mbytes/sec. There was a mix of disks on the servers, but over the 4 servers they were doing about 300MBytes/sec (split between reads and writes). The tape drives were going at between 5 and 20MB/sec depending on which disk was in use.



- Existing UKLIGHT T0-T1 network (2*1Gbit) expected to remain for SC3. Some tuning problems during SC2. Will complete a “health check” before end May.
- Will need to introduce light-weight “firewall” during May.
- May re-engineer RAL network to avoid dual attachment problems
- **Have little idea what middleware requirements there are or what experiments will do. Awaiting information from SC planning (this meeting?).**
- T1 managed 80MB/s on 2 *1Gb/s links. We continue to suspect problem is down to the link.
- Likely to be delivering from 5-8 servers at CERN to 8-16 end points at RAL. At the RAL end, selection of the endpoint hosts will be based on local CPU load (not load balance on the aggregate).

Current planning assumptions:

- Throughput challenge starts 1st July 2005 - 1 Month.
- Target 150MB/s (real data rate) T0->T1. Network load is higher than this - say 1.5Gb/s+headroom.
- T1-T2 rates - need to move/upload data via RAL-Lancaster, RAL-IC and possibly RAL-Edinburgh.
- **Need more planning data on T1-T2 rates. It is understood that LCG would accept any rate. For what scale (of MC production) are the experiments aiming?**
- Tier-2s will likely connect with T1 via existing production network over SJ4



GridPP

UK Computing for Particle Physics

Edinburgh (LHCb)

Phil Clark



GridPP
UK Computing for Particle Physics

Edinburgh (ScotGrid)

HARDWARE

- GridPP frontend machines deployed and 1 TB storage currently attached to a classic SE.
- Prototype dCache SRM installed and tested. Waiting for staff replacement to continue work this week.
- **Limited CPU available so waiting for more details about requirements from LHCb**
- Disk server: IBM xSeries 440 with eight 1.9 GHz Xeon processors, 32Gb RAM
- GB copper ethernet to GB 3com switch.

PROBLEMS/ISSUES

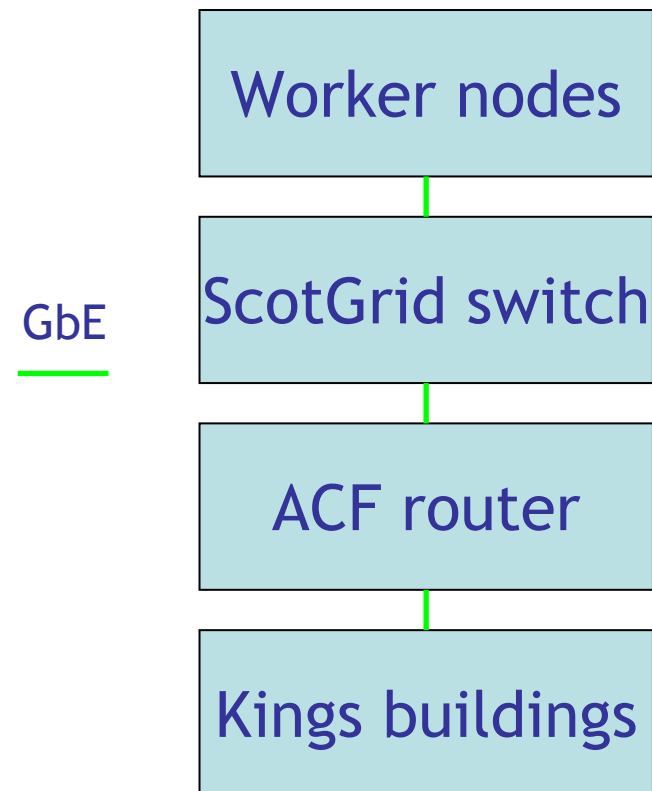
- 1) **LHCb hardware requirements not yet known**
- 2) **LHCb software installations required for SC3 not clear**
- 3) **Unlikely to be connected to UKLIGHT for SC3**



Current configuration

SE-> GBit ethernet (Copper) ->
ScotGrid Switch -> GB ethernet ->
SRIF switch -> 2x10GB/s -> 2nd SRIF
switch -> dedicated East Man
connection to Janet.

- Expected bottleneck is the IO.
- Unlikely to use UKLIGHT





GridPP

UK Computing for Particle Physics

Imperial College (CMS)

Dave Colling



Setup

- London MAN (LMN) is connected to Janet backbone.
- LMN connectivity is at 2.5Gb/s.
- There are 3 Points of Presence (PoPs) on the LMN, of which Imperial is one. Each PoP is connected at 1Gb/s.
- The traffic from the Imperial campus goes through 2 firewalls (with the campus divided up geographically so that traffic is split between them). These have never really been stressed but are unlikely to give a sustained throughput greater than ~400Mb/s
- IC has Gb/s fibre and switches from the firewall to machine room.
- The disk servers all have Gb/s connections (mixture of fibre and copper)
- Network information is available here:
<http://www.lmn.net.uk/events/29oct2002/map.pdf>



SOFTWARE

- Site upgrading to the latest LCG release
- dCache deployment starts next week
- Good UK support for experiment software installations required - but not much information yet on what this will be!

PROBLEMS/ISSUES

- Unlikely that IC ready to use UKLIGHT so production network expected to be used
- Production network likely to be limited by firewall
- **CMS requirements unclear - how much disk space should be dedicated to this challenge? Any additional equipment will need to be ordered this month!**
- Site asking about any assurances of security of dCache code - they want to view the source code!



GridPP

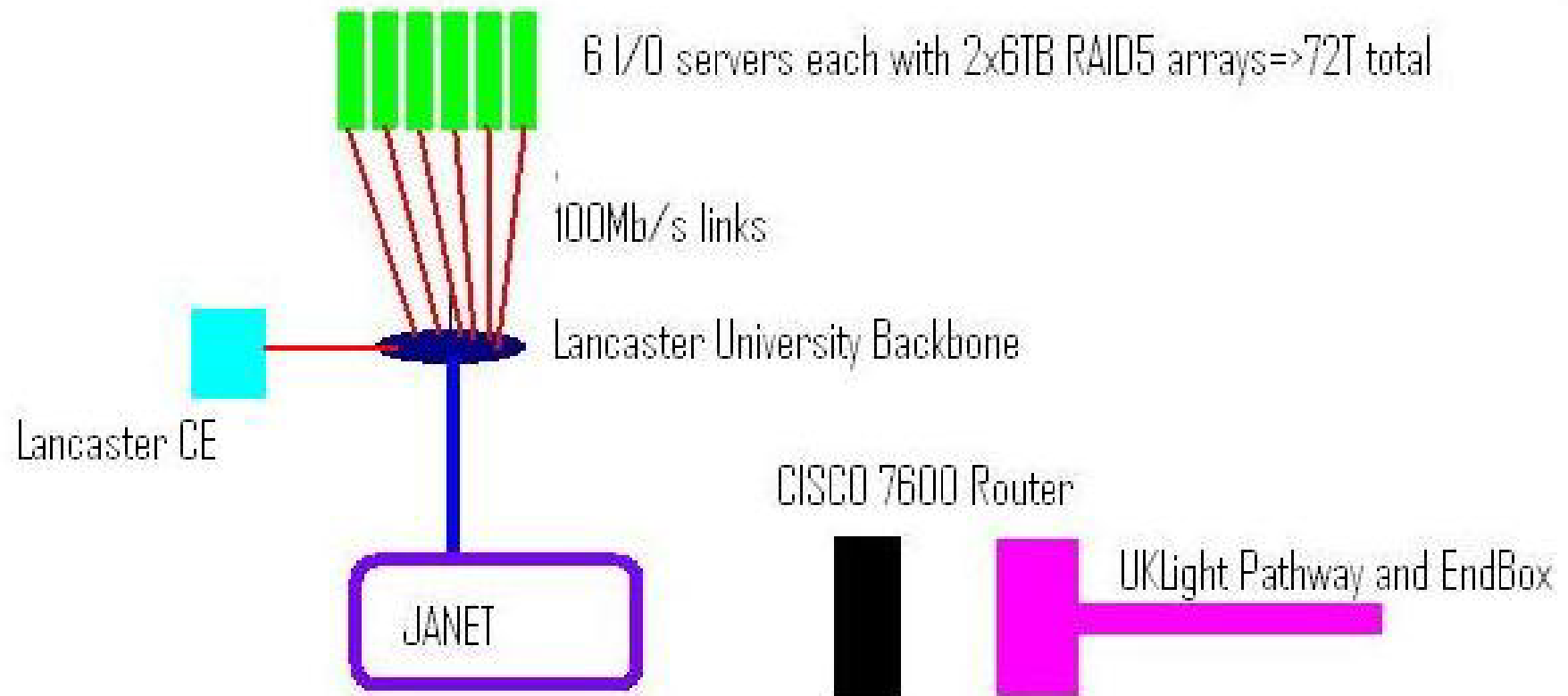
UK Computing for Particle Physics

Lancaster (ATLAS)

Brian Davies

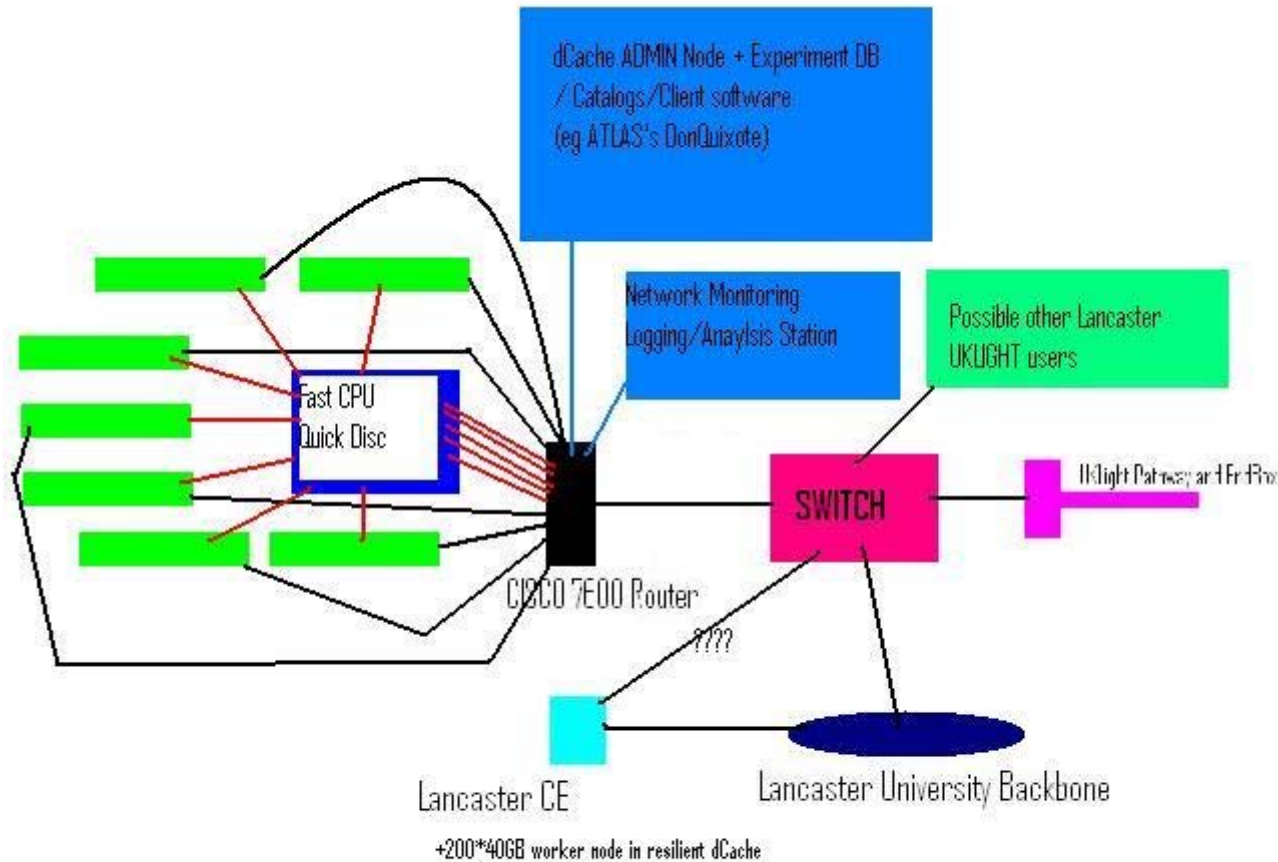


Current hardware status





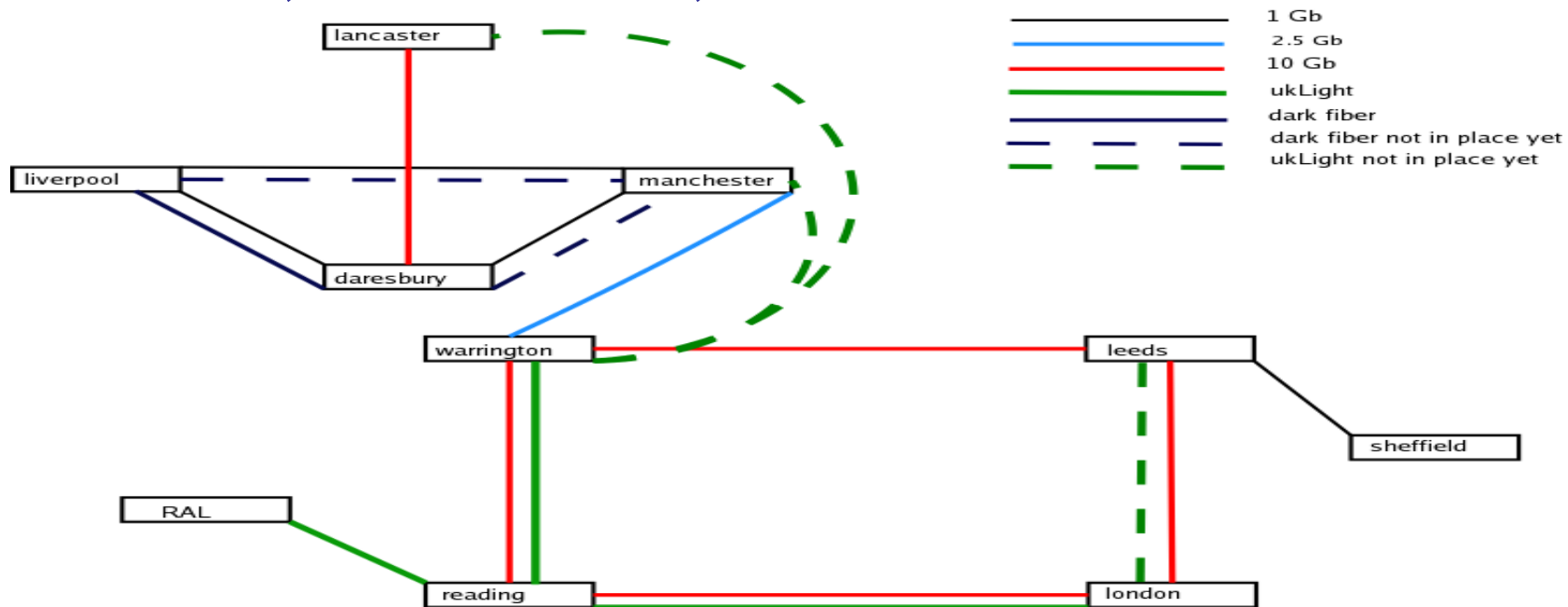
Planned site setup



1. LighPath and terminal Endbox installed.
2. Still require some hardware for our internal network topology.
3. Increase in Storage to ~84TB to possible ~92TB with working resilient dCache from CE



1. Currently data storage element connected to university backbone via 100Mb/s link
2. No Plan for transfers via university network but planning with assumption of production network as a possible backup
3. Plan traffic for UKLight connection
4. T1=RAL; T2=UK NorthGrid; T2 site=Lancaster





- LCG 2_4_0 deployed.
- dCache and SRM testbed on two nodes
 - Will deploy onto production SE when SE is fully installed.
 - Need to optimise dCache installation onto SE dCache.
- **Currently evaluating what software and services needed overall** and their deployment for site at T0/T1/ but in particular T2.
- T2 Designing system so that no changes needed for transition from throughput phase to service phase.
 - Except possible IPerf and/or transfers of “fake data” as backup of connection and bandwidth tests in throughput phase
- T1 transitions may require T2 changes.
- **Problems mainly consist of hardware software procurement/implementation!!**
 - dCache deployment
 - Experiment software/services/interfaces



- Plan organised through end of SC3 into SC4

Task	Start Date	End Date	Resource
Demonstrate sustained data transfer (T0-T1-T2)			
Integrate Don Quixote tools into SC infrastructure	Mon 27/06/05	Fri 16/09/05	BD, ATLAS
Provision of end-to-end conn (T0-T1)			
Test basic data movement (CERN-UK)	Mon 04/07/05	Fri 29/07/05	BD, MD, ATLAS, RAL
Review of bottlenecks and required actions	Mon 01/08/05	Fri 16/09/05	BD, ATLAS, RAL
ATLAS Service Challenge 3 (Service Phase)	Mon 19/09/05	Fri 18/11/05	BD, ATLAS, RAL
Review of bottlenecks and required actions	Mon 21/11/05	Fri 27/01/06	BD, ATLAS, RAL
Optimisation of network	Mon 30/01/06	Fri 31/03/06	BD, MD, BG, NP, RAL
Test of data transfer rates (CERN-UK)	Mon 03/04/06	Fri 28/04/06	BD, MD, ATLAS, RAL
Review of bottlenecks and required actions	Mon 01/05/06	Fri 26/05/06	BD, ATLAS, RAL
Optimisation of network	Mon 29/05/06	Fri 30/06/06	BD, MD, BG, NP, RAL
Test of data transfer rates (CERN-UK)	Mon 03/07/06	Fri 28/07/06	BD, MD, ATLAS, RAL
Review of bottlenecks and required actions	Mon 31/07/06	Fri 15/09/06	BD, ATLAS, RAL



Task	Start Date	End Date	Resource
Optimisation of network	Mon 18/09/06	Fri 13/10/06	BD,MD,BG,NP, RAL
Test of data transfer rates (CERN-UK)	Mon 16/10/06	Fri 01/12/06	BD,MD, ATLAS, RAL
Provision of end-to-end conn. (T1-T2)			
Integrate Don Quixote tools into SC infrastructure at LAN	Mon 19/09/05	Fri 30/09/05	BD
Provision of memory-to-memory conn. (RAL-LAN)	Tue 29/03/05	Fri 13/05/05	UKERNA,BD,BG,NP,RA L
Provision and Commission of LAN h/w	Tue 29/03/05	Fri 10/06/05	BD,BG,NP
Installation of LAN dCache SRM	Mon 13/06/05	Fri 01/07/05	MD,BD
Test basic data movement (RAL-LAN)	Mon 04/07/05	Fri 29/07/05	BD,MD,ATLAS, RAL
Review of bottlenecks and required actions	Mon 01/08/05	Fri 16/09/05	BD
[SC3 – Service Phase]			
Review of bottlenecks and required actions	Mon 21/11/05	Fri 27/01/06	BD
Optimisation of network	Mon 30/01/06	Fri 31/03/06	BD,MD,BG,NP
Test of data transfer rates (RAL-LAN)	Mon 03/04/06	Fri 28/04/06	BD,MD,
Review of bottlenecks and required actions	Mon 01/05/06	Fri 26/05/06	BD



GridPP
UK Computing for Particle Physics

Lancs/ATLAS SC3 Plan

Task	Start Date	End Date	Resource
Optimisation of network	Mon 29/05/06	Fri 30/06/06	BD,MD,BG,NP
Test of data transfer rates (RAL-LAN)	Mon 03/07/06	Fri 28/07/06	BD,MD
Review of bottlenecks and required actions	Mon 31/07/06	Fri 15/09/06	BD
Optimisation of network	Mon 18/09/06	Fri 13/10/06	BD,MD,BG,NP
Test of data transfer rates (RAL-LAN)	Mon 16/10/06	Fri 01/12/06	BD,MD



Task	Start Date	End Date
Review dCache issues from SC2 and T2 test installations	13/04/05	14/04/05
Basic infrastructure deployment	15/04/05	27/05/05
Reconfirm experiment hardware and software requirements	25/04/05	29/04/05
Procure any additional hardware required	2/05/05	27/05/05
Training for additional site services	13/05/05	13/05/05
Deploy and test dCache/SRM at T2s for SC	2/05/05	18/05/05
Deploy and test experiment specific software for Phase 2?	ASAP?	??
Tune and test T1-T0 with SC hardware	09/05/05	10/06/05
Network tests T1-T2s (Mem-Mem, Disk-Disk)	16/05/05	18/06/05
Review of bottlenecks and other issues	18/06/05	29/06/05
SC3 throughput phase	01/07/05	31/08/05
SC3 service phase	01/09/05	...

Draft



- All UK participating sites have started work for SC3
- Lack of clear requirements is causing some problems in planning and thus deployment - what are the critical success factors for SC3? What service level is expected of Tier-2s?
“We don’t know what we want but we’ll know when we have it!”
- Lower T1-T2 data rates than expected is pushing GridPP towards using current production network across SJ4 (UKLight may be available for Lancaster)
- dCache deployment is progressing at a moderate pace. Still need to resolve T1 dual-node problem
- Some areas still unclear such as what extra services will be deployed at the participating sites. (Nb. next LCG release is on 1st July). This is becoming clearer but a list (with responsible group) would help
- Several open questions about how network monitoring will be performed during SC3 - what are the requirements on sites?