North West Grid Overview

John Brooke[1], R.J. Allan[2]
[1] The University of Manchester
[2] CCLRC Daresbury Laboratory

A world-class Grid infrastructure for the North West.

A Unique North-West Project:
advancing Grid Technologies and Applications

Technology "tuned to the needs of practicing scientists".

Hooks to other Grid consortia: NGS, WRG

Top end: HPCx and CSAR

Applications and industry

Pharma, meds, bio, social, env, CCPs

Portals, client toolkits, active overlays

Mid range: NW-GRID and local clusters

User interfaces

Sensor networks and experimental facilities

Advanced Network technology

Desktop pools: Condor etc.

Rob Allan/John Brooke

The NW-GRID Project

- Funded by NWDA
- A Grid resource for the North West
- The project will be structured around three major testbed milestones:
  - Testbed 1 (2006): High-performance and parametric applications
  - Testbed 2 (2007): Safe, bookable and accountable Grid
  - Testbed 3 (2008): Collaborative real-time Grid
- Academic users
- Commercial users
- Partners

Statement

- The NW-GRID project, a collaboration between CCLRC Daresbury Laboratory and the Universities at Lancaster, Liverpool and Manchester, will establish a computational Grid comprising large-scale commodity computing systems coupled by a high-speed network. It will establish, for the region, a world-class activity in the deployment and exploitation of Grid middleware technologies (the software that glues together the various data and computing resources) and demonstrate the capabilities of the Grid in leading edge computational science and engineering applications.

- The Grid was a key component of the proposed NW science strategy and the project resonates strongly with the key elements of the NWDA's regional strategy in particular in working with targeted emerging sectors in the environment, biotechnology and pharmaceutical and complex materials areas, establishing the North-West as a global player in Grid technologies and in embedding e-competencies across the region's business, academic and industrial base.

Key Applications and Research

- chemistry
- life sciences
- nano-technology
- human anatomy
- population modelling and epidemiology
- earth science
- oceanographic modelling
- environmental monitoring
- disaster response
- astronomy
- bio-informatics
- materials modelling
- social science and statistics
- management science
- complex fluids
- protein crystallography
- next generation middleware
- user interfaces
- enhancing production grids

Web site: http://www.nw-grid.ac.uk

NW-GRID Core

Four compute clusters funded by NWDA are accessible using Grid middleware and connected using a dedicated fibre network:

- Daresbury –
  - dl1.nw-grid.ac.uk - 96 nodes (some locally funded)
- Lancaster –
  - lancs1.nw-grid.ac.uk - 192 nodes (96 dedicated to local users)
- Liverpool –
  - lv1.nw-grid.ac.uk - 44 nodes
- Manchester –
  - man2.nw-grid.ac.uk – 27 nodes

Pictures

Anatomy of a Cluster

- Compute clusters from Streamline Computing
- 2x Sun x4200 head nodes
  - dual-processor single-core AMD Opteron 2.6GHz (64-bit)
  - 16GB memory
- Panasas file store
  - around 3-10TB per cluster (Manchester client only)
- Sun x4100 nodes
  - Dual-processor dual-core AMD Opteron 2.4GHz (64-bit)
  - 2-4GB memory pre core
  - 2x 73GB disks per node
  - Gbit/s ethernet between nodes
- Sun Grid Engine and Streamline cluster management
- NGS compatible Grid Middleware - Globus 2.4

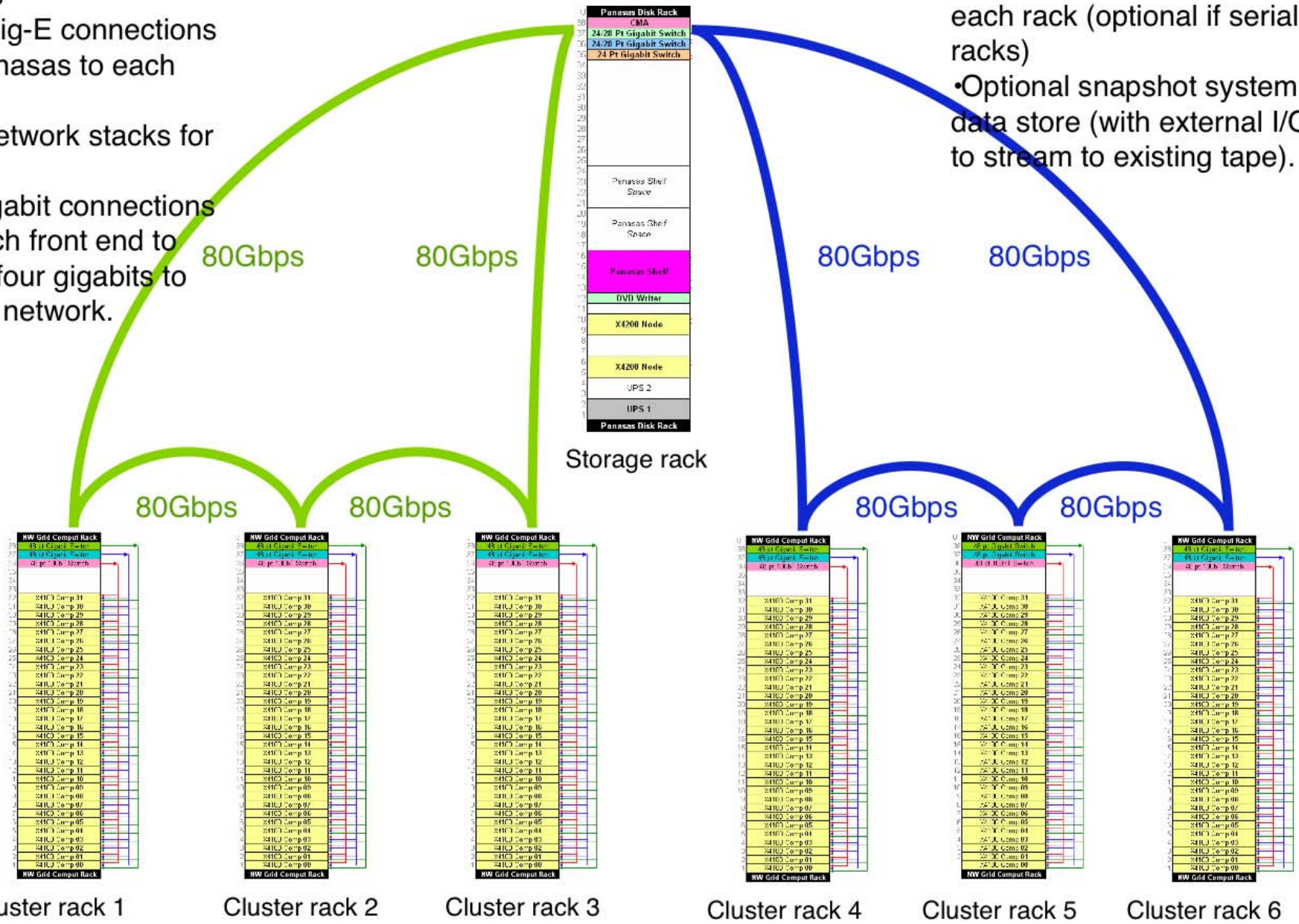
Connecting the Nodes

- within a “module” and between modules
- dual Gbit-E network
  - inter-node communications via stacked Nortel Gbit switch
  - command, control and file access via 2nd stacked Nortel Gbit switch
- Ethernet network
  - diagnostics via standard ethernet switch
- Link to Panasas if installed

Lancaster

- six compute node racks
- dedicated MPI network in each rack (optional if serial racks)
- Optional snapshot system for data store (with external I/O to stream to existing tape).

- Gig-E connections from each front end to both switches
- Dual Gig-E connections from Panasas to each switch
- Dual network stacks for storage
- Two gigabit connections from each front end to provide four gigabits to external network.



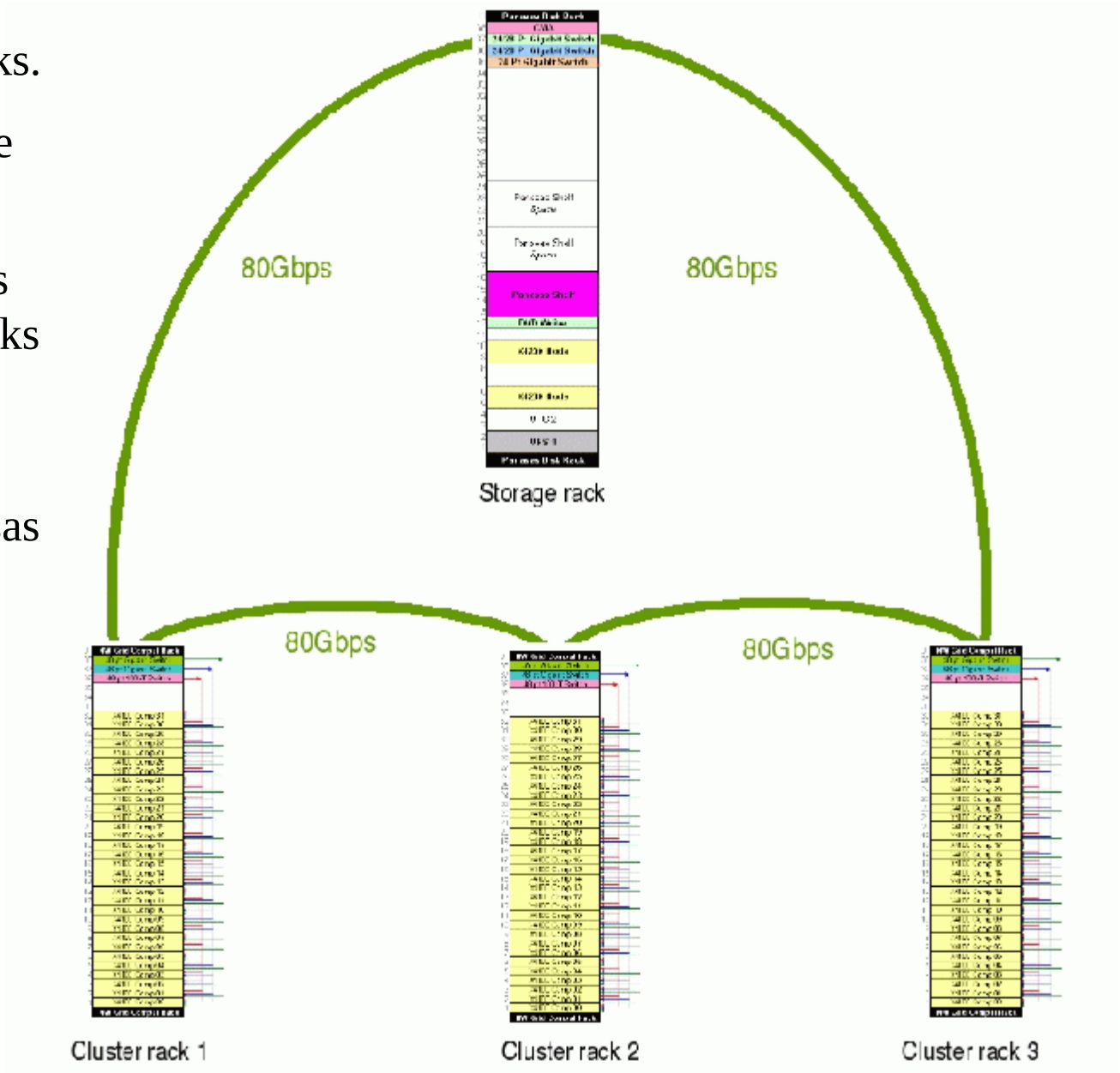
DL Cluster Detail

DL Cluster has 4 racks.

32 nodes per compute rack plus 3 switches.

Gbit-E (80Gb/s) links between compute racks and into head node shown in green.

2x head nodes, Panasas storage and external connection in central rack.

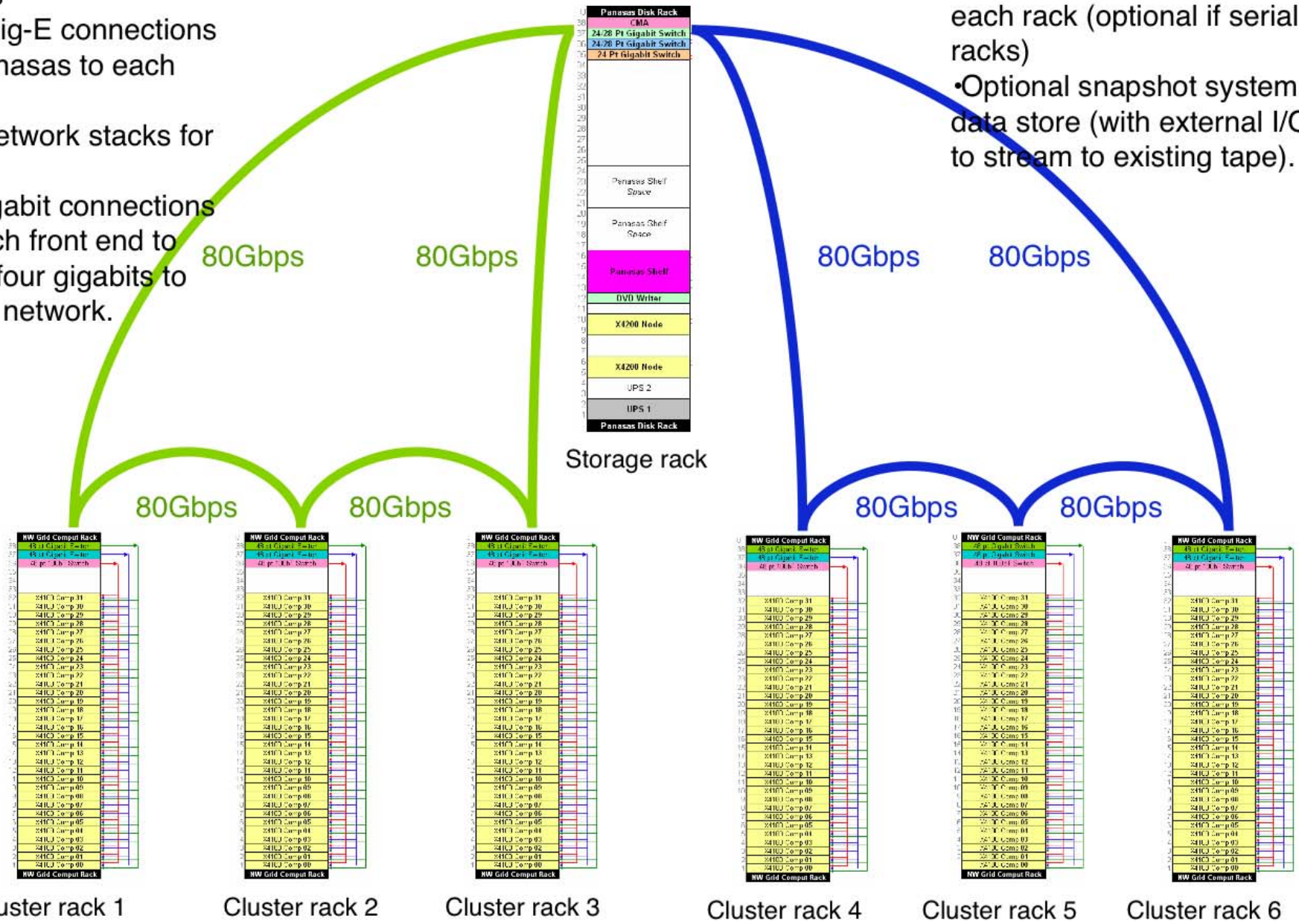


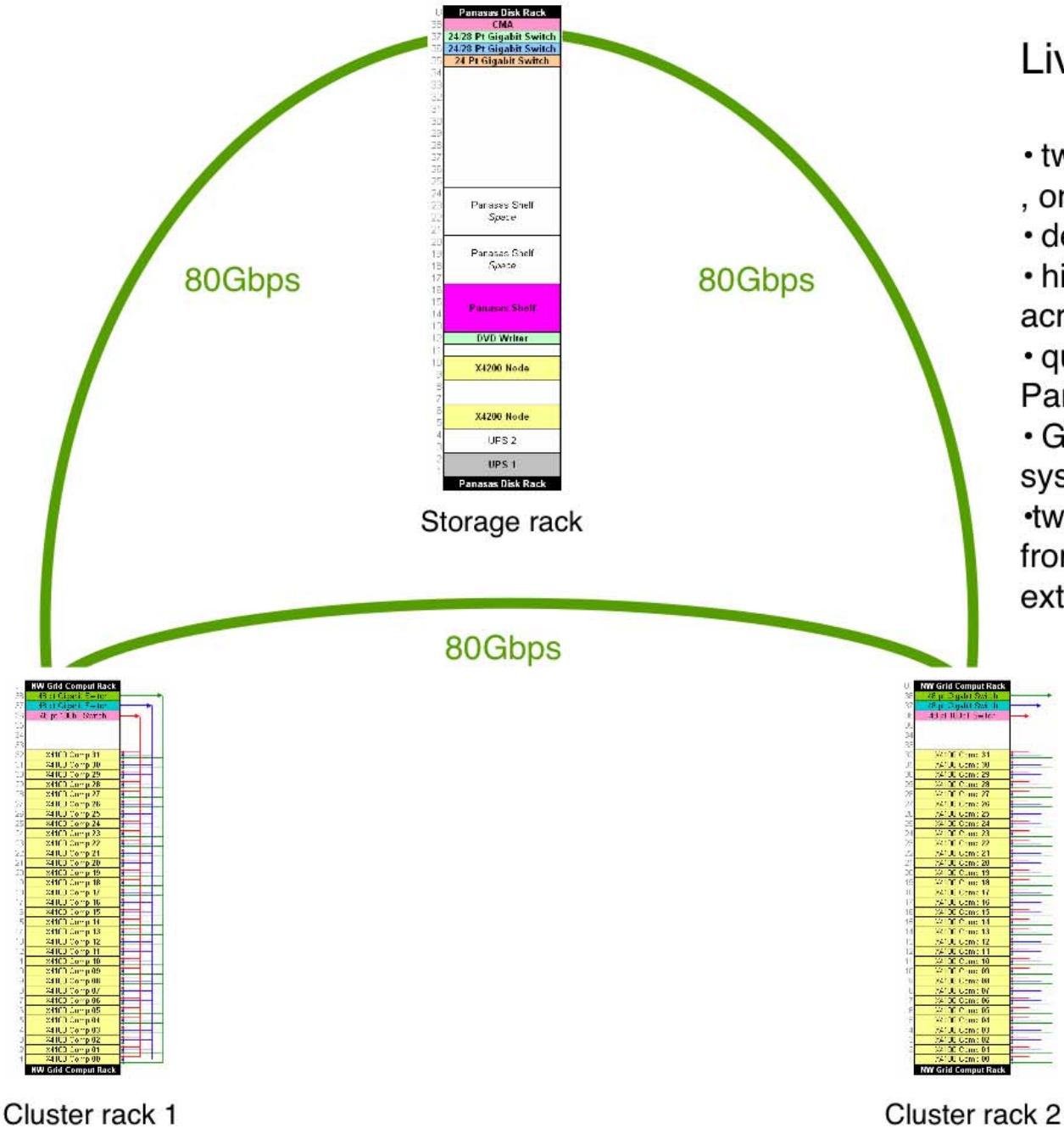
Rob Allan/John Brooke

Lancaster

- six compute node racks
- dedicated MPI network in each rack (optional if serial racks)
- Optional snapshot system for data store (with external I/O to stream to existing tape).

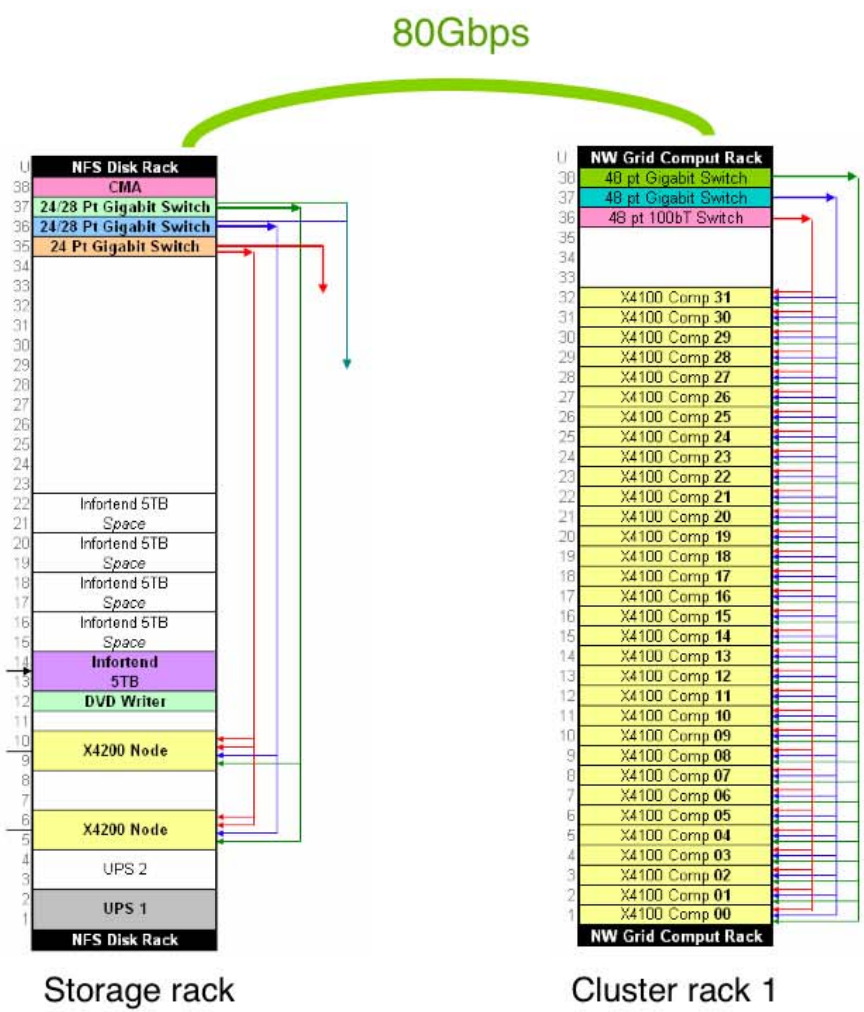
- Gig-E connections from each front end to both switches
- Dual Gig-E connections from Panasas to each switch
- Dual network stacks for storage
- Two gigabit connections from each front end to provide four gigabits to external network.





Liverpool

- two compute node racks One full , one part full.
- dedicated MPI network per rack
- high-bandwidth storage network across all racks
- quad Gig-E connection from Panasas to storage network
- Gig-E connection from front end systems to storage network
- two gigabit connections from each front end to provide four gigabits to external network.

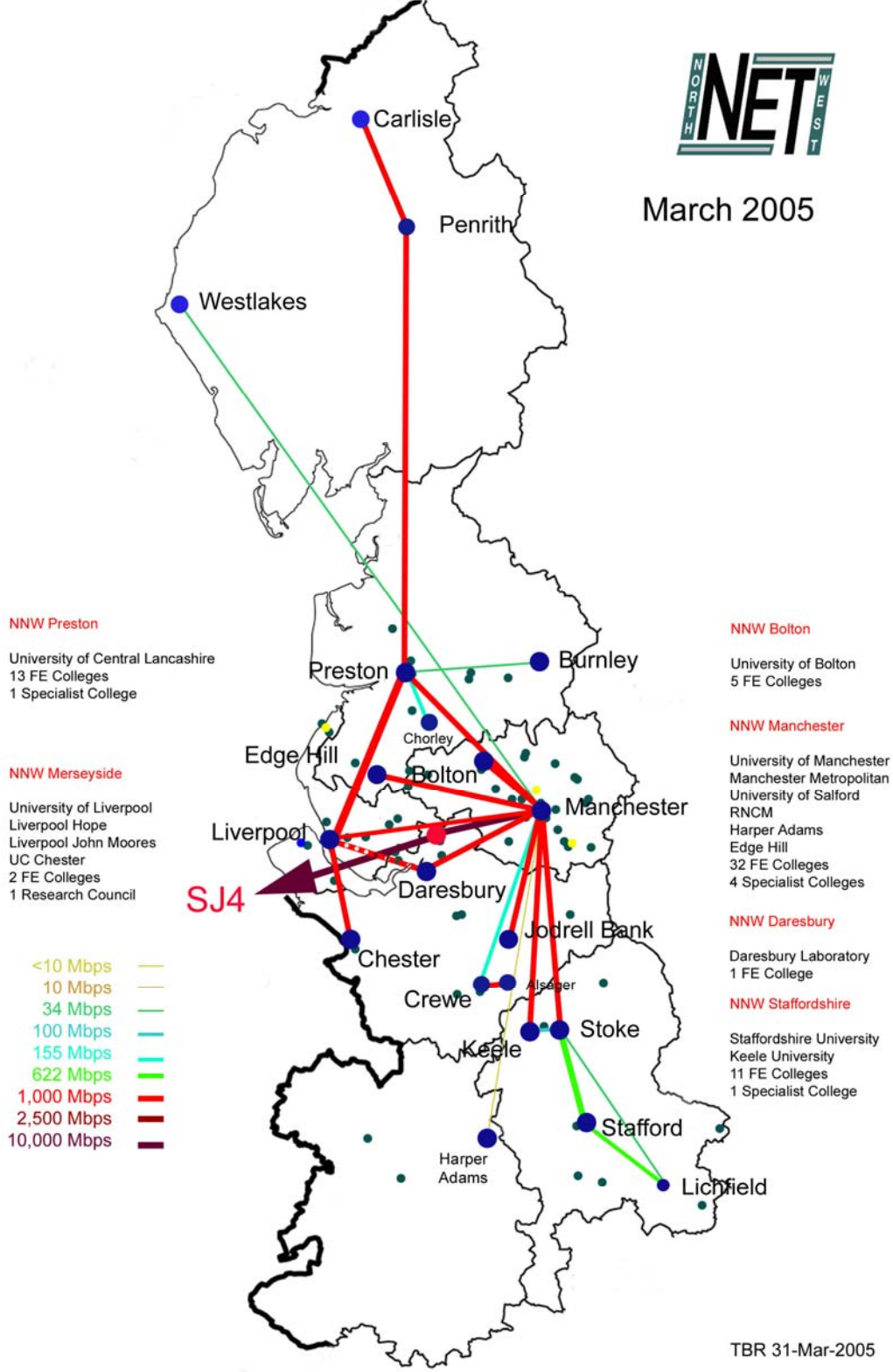


Manchester

- one part full compute node rack with dedicated MPI network
- high-bandwidth network to storage rack
- dual Gig-E connection from NFS server to storage network
- Gig-E connection from front end systems to storage network
- Two gigabit connections from each front end to provide four gigabits to external network.

Connecting the Clusters

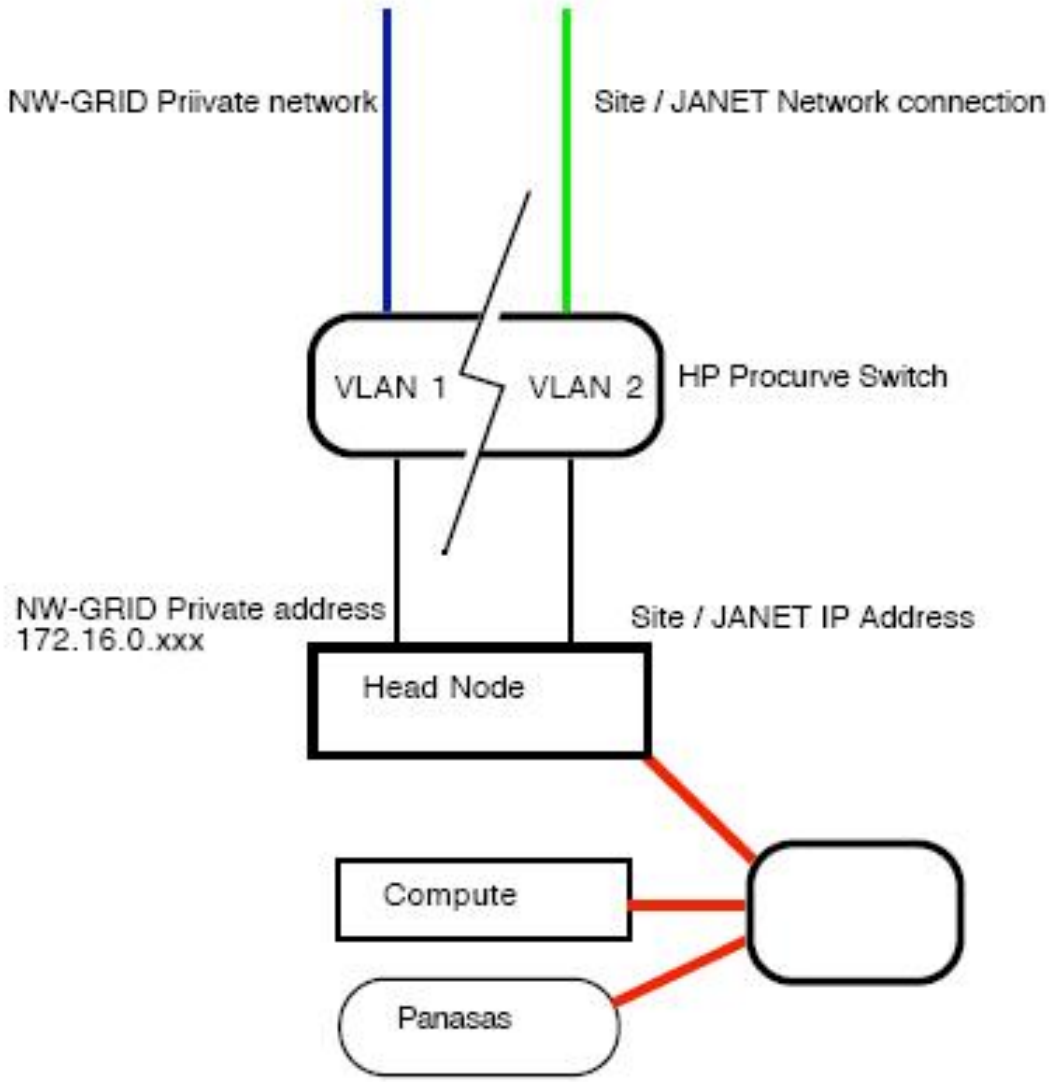
- Map of the NW – on next slide
- HP Procurve/ 3Com switch is part of each control module
- External connection
  - Private Fibre network
  - SuperJANET
- Commercial access is possible by VPN and will use the fibre network


NORTH NET WEST
March 2005
Carlisle
Penrith
Westlakes
Preston
Burnley
Chorley
Edge Hill
Bolton
Manchester
Liverpool
Daresbury
SJ4
Jodrell Bank
Chester
Crewe
Alsager
Stoke
Keele
Stafford
Harper Adams
Lichfield
NNW Preston
University of Central Lancashire
13 FE Colleges
1 Specialist College
NNW Merseyside
University of Liverpool
Liverpool Hope
Liverpool John Moores
UC Chester
2 FE Colleges
1 Research Council
<10 Mbps
10 Mbps
34 Mbps
100 Mbps
155 Mbps
622 Mbps
1,000 Mbps
2,500 Mbps
10,000 Mbps
NNW Bolton
University of Bolton
5 FE Colleges
NNW Manchester
University of Manchester
Manchester Metropolitan
University of Salford
RNCM
Harper Adams
Edge Hill
32 FE Colleges
4 Specialist Colleges
NNW Daresbury
Daresbury Laboratory
1 FE College
NNW Staffordshire
Staffordshire University
Keele University
11 FE Colleges
1 Specialist College
TBR 31-Mar-2005

NW-GRID Fibre Network

- 1Gb/s between sites – latency?
  - DL-Lancs trunk - 10x 1Gb/s links
  - DL-Manchester – single link
  - DL-Liverpool – single link
- What can it enable us to do?
  - Scheduling, job migration, visualisation etc.
  - inter-site MPI, MPI-G
  - resource co-scheduling

External Connections



HARC

- Highly Available Resource Co-scheduler.
- John MacLaren's implementation of Paxos.
- Acceptors at different sites for fault tolerance.
- They issue prepare instructions to the RMS (e.g. SGE) then run the consensus algorithm before commit.
- Will be used in the GENIUS (Grid-Enabled Neurosurgical Imaging Using Simulation) project to link NGS, NW-GRID and TeraGrid

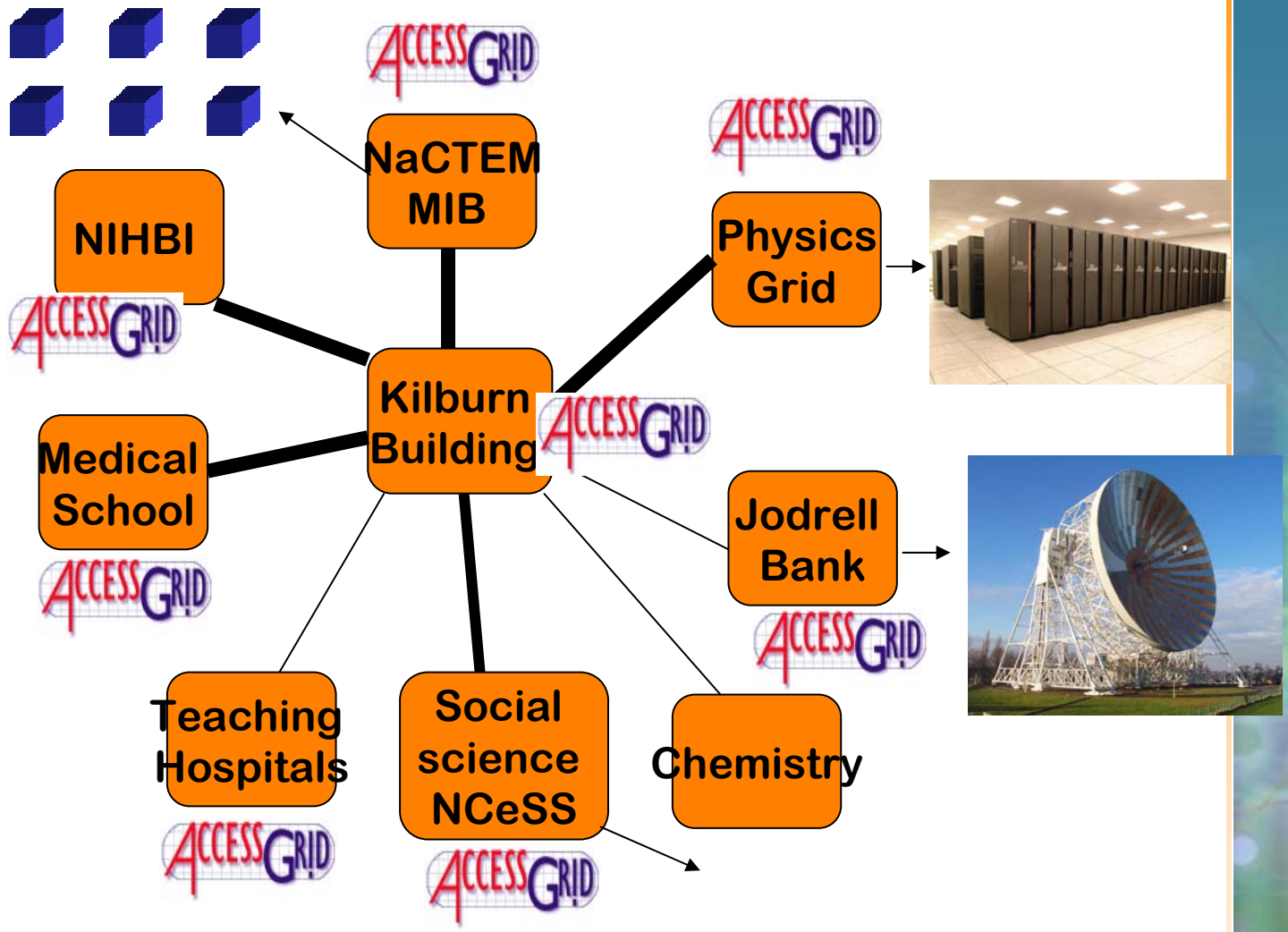
Other Resources

- Manchester – see http://man4.nw-grid.ac.uk
  - man1.nw-grid.ac.uk - Weyl cluster, It is a 44 node system plus login and master nodes. The address of the login node is weyl.mc.man.ac.uk and IP address is 130.88.200.240. Compute nodes 01 - 40 are dual processor AMD Opteron 248 2GHz, 78GB of disk, 2GB memory. Nodes 41 - 43 are fat nodes with the same spec as above except the memory is 8GB. The user, scratch and software disks are RAID, 2.5TB, NFS mounted.
  - man3.nw-grid.ac.uk - Escher is a Silicon Graphics Prism (IA64) with 8 Itanium 2 processors, 4 FireGL 3 graphics pipelines running SUSE Linux Enterprise Server 9 with Silicon Graphics ProPack 4
  - man4.nw-grid.ac.uk – Biruni is a Web server with Ganglia monitoring results.

Other Resources

- Daresbury
  - BlueGene L – 1000 processors, for software development, available soon.
- Liverpool
  - POL Cluster – 192x Xeon (32 bit), likely to become part of NW-GRID in April.
- Preston
  - 8-node SGI Altix 3700. Discussions underway with Dept. of Theoretical Astrophysics, U. Central Lancashire.

Manchester Campus Grid



Other sites, e.g. Liverpool are also using Condor.

Lancaster is an NGS partner.

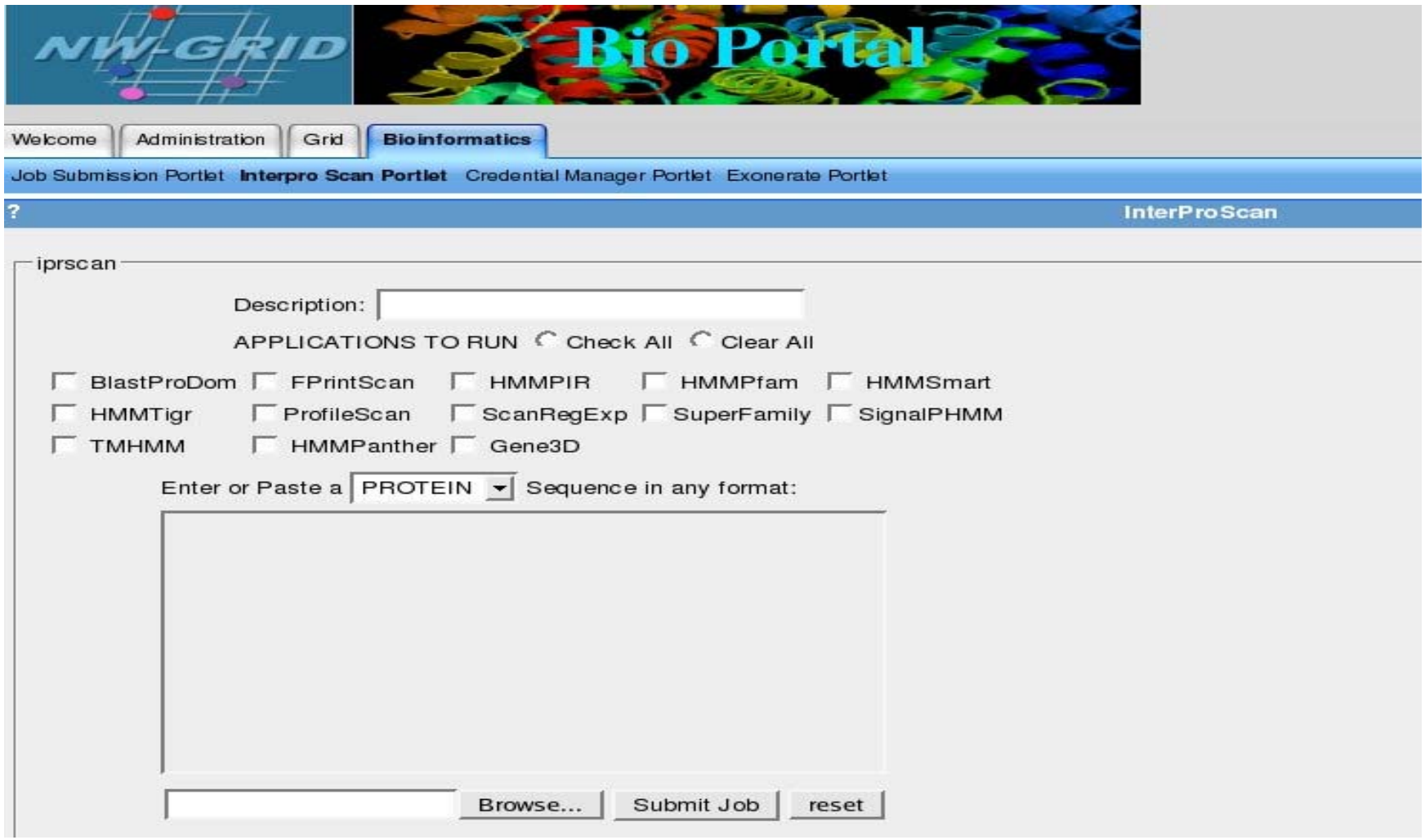
HPCx at Daresbury is part of the NGS.

Rob Allan/John Brooke

Bio Portal on NW-GRID

- User Interface to a number of applications
  - InterProScan
    - An integrated framework for a number of sequence searching applications
  - Exonerate
    - A generic tool for pairwise sequence comparison
- Job monitoring Interface
- Results can be requested/received via email

InterProScan Portal



Job Monitoring

Job Submission Portlet

Refresh List | Delete Jobs

	Job Id	Description	Resource	Status	Date Submitted
	https://man2.nw-grid.ac.uk:64024/20740/1174395060/	test 2	man2.nw-grid.ac.uk	Job completed with message success	Mar 20, 2007 12:49:08 PM
	https://man2.nw-grid.ac.uk:64024/15205/1174388242/	test mail	man2.nw-grid.ac.uk	Job completed with message success	Mar 20, 2007 10:55:30 AM
	https://man2.nw-grid.ac.uk:64024/30689/1174386802/	test mail	man2.nw-grid.ac.uk	Job completed with message success	Mar 20, 2007 10:31:29 AM
	https://man2.nw-grid.ac.uk:64024/17181/1174385709/	test	man2.nw-grid.ac.uk	Job completed with message success	Mar 20, 2007 10:13:17 AM
	https://man2.nw-grid.ac.uk:64024/31605/1174384179/	test mail	man2.nw-grid.ac.uk	Job completed with message success	Mar 20, 2007 9:47:47 AM
	https://man2.nw-grid.ac.uk:64026/23425/1174383640/	test	man2.nw-grid.ac.uk	Job completed with message success	Mar 20, 2007 9:38:47 AM
	https://man2.nw-grid.ac.uk:64024/19369/1174383562/	test	man2.nw-grid.ac.uk	Job completed with message success	Mar 20, 2007 9:37:29 AM
	https://man2.nw-grid.ac.uk:64028/7258/1173968912/	test9	man2.nw-grid.ac.uk	Job completed with message success	Mar 15, 2007 2:26:38 PM
	https://man2.nw-grid.ac.uk:64024/3345/1173968833/	test 8	man2.nw-grid.ac.uk	Job completed with message success	Mar 15, 2007 2:25:20 PM
	https://man2.nw-grid.ac.uk:64026/25488/1173968464/	test5	man2.nw-grid.ac.uk	Job failed with error code 131; the user proxy expired (job is still running);	Mar 15, 2007 2:19:10 PM
	https://man2.nw-grid.ac.uk:64024/21594/1173968385/	test	man2.nw-grid.ac.uk	Job completed with message success	Mar 15, 2007 2:17:51 PM
	https://man2.nw-grid.ac.uk:64028/16511/1173968226/	test 3	man2.nw-grid.ac.uk	Job completed with message success	Mar 15, 2007 2:15:13 PM
	https://man2.nw-grid.ac.uk:64026/12629/1173968144/	test 2	man2.nw-grid.ac.uk	Job completed with message success	Mar 15, 2007 2:13:51 PM
	https://man2.nw-grid.ac.uk:64024/8748/1173968066/	testcred	man2.nw-grid.ac.uk	Job completed with message success	Mar 15, 2007 2:12:33 PM
	https://man2.nw-grid.ac.uk:64027/28417/1173960398/	testcred	man2.nw-grid.ac.uk	Job completed with message success	Mar 15, 2007 12:04:44 PM
	https://man2.nw-grid.ac.uk:64027/16713/1173959574/	test	man2.nw-grid.ac.uk	Job completed with message success	Mar 15, 2007 11:51:00 AM
	https://man2.nw-grid.ac.uk:64024/24686/1173885265/	test all	man2.nw-grid.ac.uk	Job completed with message success	Mar 14, 2007 3:12:31 PM

Results Via Email

Subject: Results of job test
From: jobmailer@bio-portal.mc.manchester.ac.uk
Date: 15/03/07 11:54
To: Bharathi.Kattamuri@manchester.ac.uk

```
Results of job 'test' submitted to man2.nw-grid.ac.uk on 15 March 2007 11:51:00 o'clock GMT

Q9RHD9  D44DAE8C544CB7C1        267     HMMPfam PF00575 S1      1       55      38.7    T       15-Mar-2007
Q9RHD9  D44DAE8C544CB7C1        267     HMMPfam PF00575 S1      68      142     81.6    T       15-Mar-2007
Q9RHD9  D44DAE8C544CB7C1        267     HMMPfam PF00575 S1      155     228     82.8    T       15-Mar-2007
Q9RHD9  D44DAE8C544CB7C1        267     HMMSmart        SM00316 S1      3       55      38.0    T       15-Mar-2007
Q9RHD9  D44DAE8C544CB7C1        267     HMMSmart        SM00316 S1      70      142     81.0    T       15-Mar-2007
Q9RHD9  D44DAE8C544CB7C1        267     HMMSmart        SM00316 S1      157     228     86.8    T       15-Mar-2007
Q9RHD9  D44DAE8C544CB7C1        267     HMMPanther      PTHR23270       PTHR23270       72      239     38.5    T       15-Mar-2007
Q9RHD9  D44DAE8C544CB7C1        267     FPrintScan      PR00681 RIBOSOMALS1     6       27      4.6E-18 T       15-Mar-2007
Q9RHD9  D44DAE8C544CB7C1        267     FPrintScan      PR00681 RIBOSOMALS1     85      104     4.6E-18 T       15-Mar-2007
Q9RHD9  D44DAE8C544CB7C1        267     FPrintScan      PR00681 RIBOSOMALS1     125     143     4.6E-18 T       15-Mar-2007
Q9RHD9  D44DAE8C544CB7C1        267     Gene3D  G3DSA:2.40.50.140       NA-bind_OB_sub  3       71      1.5E-15 T       15-Mar-2007
Q9RHD9  D44DAE8C544CB7C1        267     Gene3D  G3DSA:2.40.50.140       NA-bind_OB_sub  72      154     4.6E-22 T       15-Mar-2007
Q9RHD9  D44DAE8C544CB7C1        267     Gene3D  G3DSA:2.40.50.140       NA-bind_OB_sub  159     249     2.8E-22 T       15-Mar-2007
Q9RHD9  D44DAE8C544CB7C1        267     superfamily     SSF50249        Nucleic_acid_OB 3       60      1.4E-7  T       15-Mar-2007
Q9RHD9  D44DAE8C544CB7C1        267     superfamily     SSF50249        Nucleic_acid_OB 61      205     6.3999999999999995E-24  T
15-Mar-2007
Q9RHD9  D44DAE8C544CB7C1        267     ProfileScan     PS50126 S1      1       55      14.869  T       15-Mar-2007
Q9RHD9  D44DAE8C544CB7C1        267     ProfileScan     PS50126 S1      72      142     20.809  T       15-Mar-2007
Q9RHD9  D44DAE8C544CB7C1        267     ProfileScan     PS50126 S1      159     228     22.541  T       15-Mar-2007
RS16_ECOLI      F94D07049A6D489D        82      Gene3D  G3DSA:3.30.70.690       no description  1       79      2.9e-29 T       15-Mar-2007
Y902_MYCTU      CD84A335CCFFE6D7        446     Gene3D  G3DSA:1.10.287.240      no description  220     292     3.3e-10 T       15-Mar-2007
Y902_MYCTU      CD84A335CCFFE6D7        446     Gene3D  G3DSA:3.30.565.10       no description  300     443     1e-30   T       15-Mar-2007
Y902_MYCTU      CD84A335CCFFE6D7        446     HMMSmart        SM00304 no description  170     222     3.1e-06 T       15-Mar-2007
Y902_MYCTU      CD84A335CCFFE6D7        446     HMMSmart        SM00388 no description  230     296     2.4e-12 T       15-Mar-2007
Y902_MYCTU      CD84A335CCFFE6D7        446     HMMSmart        SM00387 no description  338     446     5e-24   T       15-Mar-2007
RS16_ECOLI      F94D07049A6D489D        82      ScanRegExp      PS00732 RIBOSOMAL_S16   2       11      8e-5    T       15-Mar-2007
RS16_ECOLI      F94D07049A6D489D        82      BlastProDom     PD003791        RS16_BUCAI_P57474;      9       77      1e-19   T
15-Mar-2007
```

Future Plans

- NW-GRID will have a dual role as a production and as an experimental Grid.
- The first phase was designed to be compatible with NGS
- The second phase will introduce experimental capability not yet available on NGS - co-scheduling
- The third phase will integrate sensor nets using middleware based on overlays
- Eventually we would like to see fuller integration both with NGS and with campus Grids.

To find out more

- NW-GRID stand at UK e-Science Village here at OGF20
- Web site http://www.nw-grid.ac.uk
- ESNW representation on UK Engineering Task Force will enable software developed on NW-GRID to be evaluated for UK use.
- International collaborations, e.g. with TeraGrid