http://www.ngs.ac.uk

http://www.grid-support.ac.uk

NGS Update

Stephen Pickles
<stephen.pickles@manchester.ac.uk>
Technical Director, GOSC
GridPP Deployment Board, Dublin,
14th September 2005

(based on Neil Geddes' presentation to e-Science centre directors)

National Grid Service

Outline

- Highlights since last meeting
- Users
  - registrations, usage, helpdesk queries
  - analysis of current users
- Other Developments
  - Partners and reviews
- GOSC Roadmap
- Summary

"Highlights"

- Steady growth
  - number of users continues to grow
  - utilization of resources (now 50(data)-80(compute) %)
- Training Course
  - Oxford (July 2005), NeSC (September 2005)
    - well attended (16), very positive feedback
  - NeSC course, Sept 29/30
  - thanks as ever to Mike Mineter & NeSC Training Team
- Expansion

Expansion

Partnership programme

- Lancaster ratified as partner by GOSC Board in July
  - (Bristol and Cardiff ratified previously)

Next:

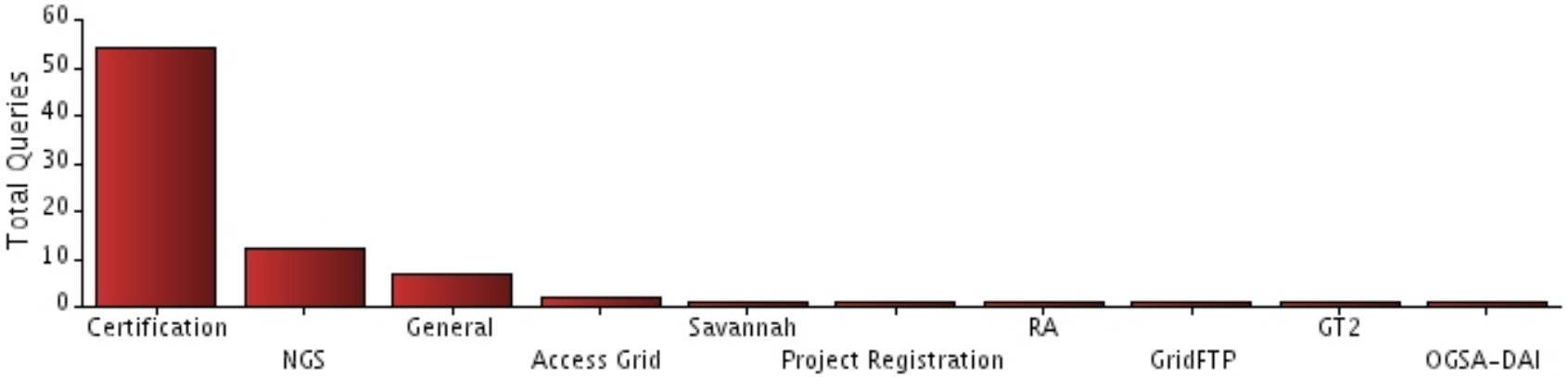
- NeSC (initially as affiliate)
  - work in progress

In discussions with:

- Belfast

Heterogeneity continues to increase

GOSC Helpdesk



Count of Queue:

Certification 54

NGS 12

General 7

Access Grid 2

Savannah 1

Project Registration 1

RA 1

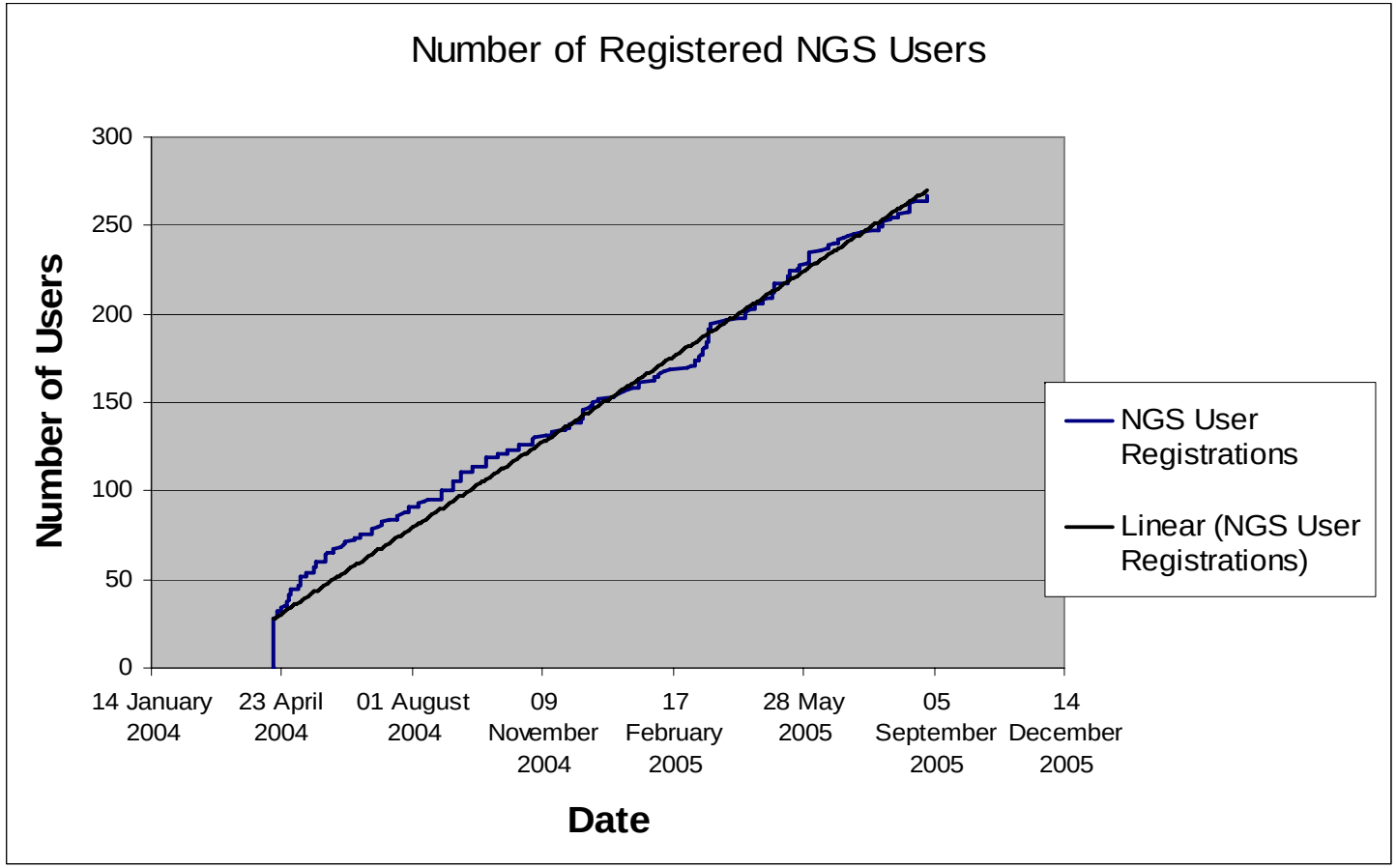
GridFTP 1

GT2 1

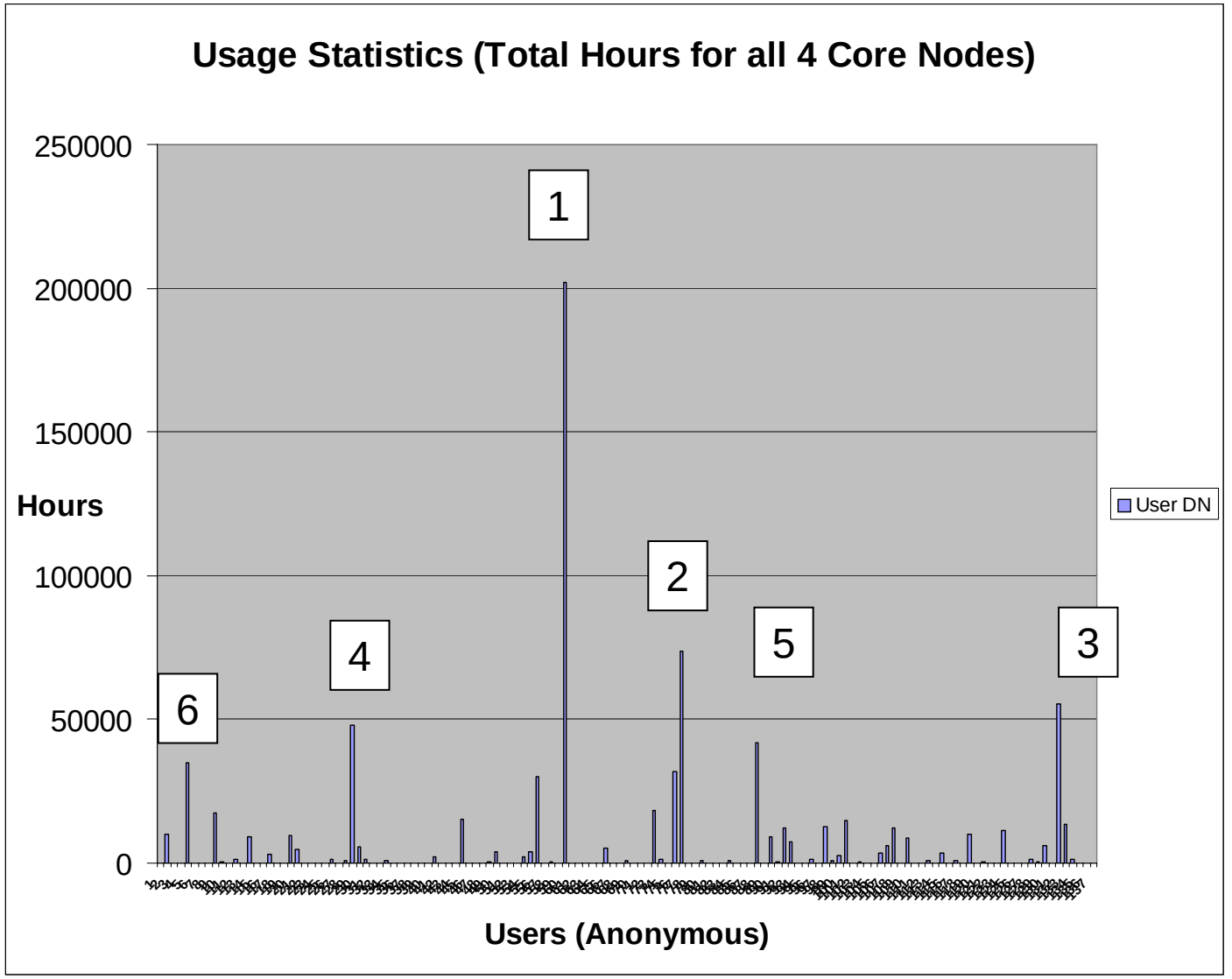
OGSA-DAI 1

National Grid Service

NGS Users



National Grid Service



National Grid Service

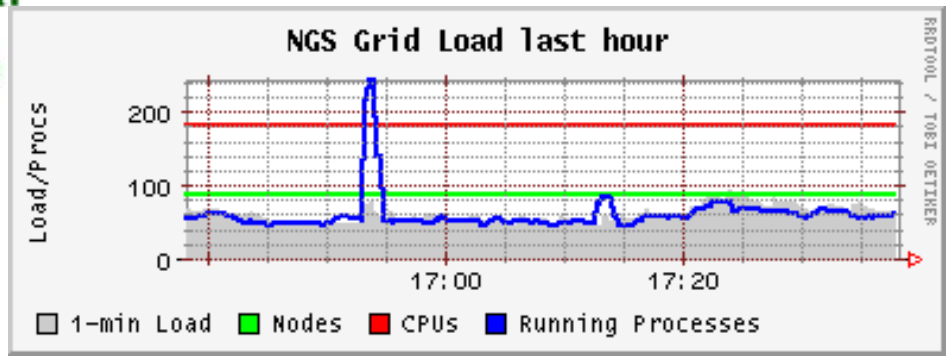
CPU Usage across the 4 core NGS sites

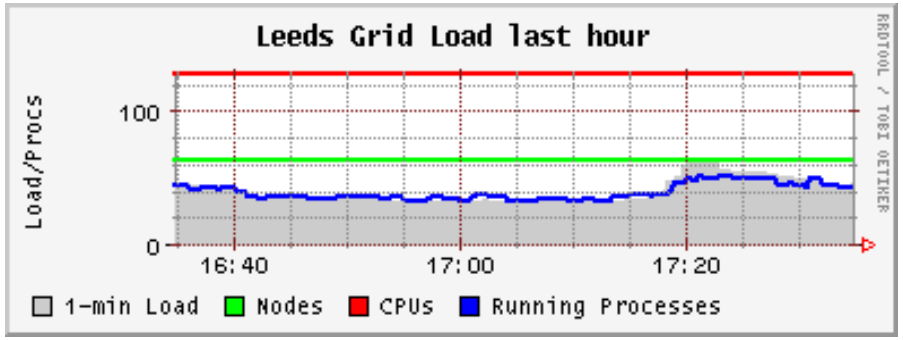
Usage Information	Hours Used	Hours Applied for	Hours Available Per yr @ 4 sites
TOTAL (hrs):	782731	2917280	2557920
TOTAL (yrs): (Proportion of total)	***0.306***	***1.14***	***1 (where 1 = total capacity above)***

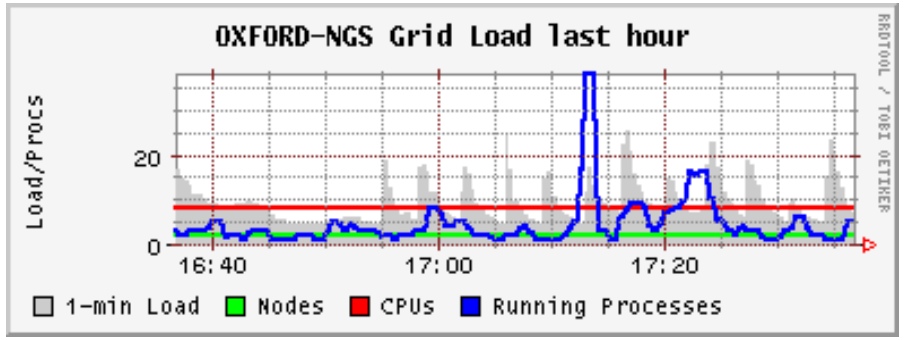
Taken from PBS account records at 4 core NGS sites

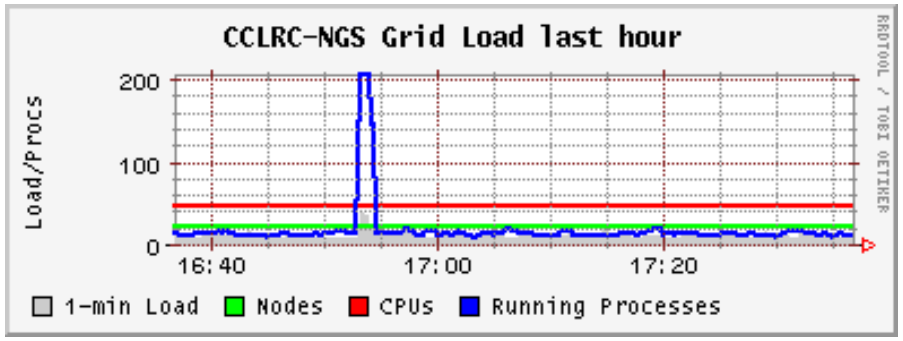
National Grid Service

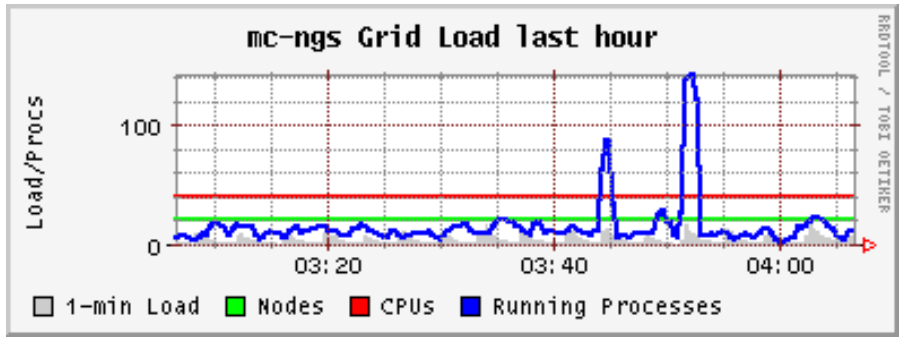
Example usage over 1hr period 31/08/05







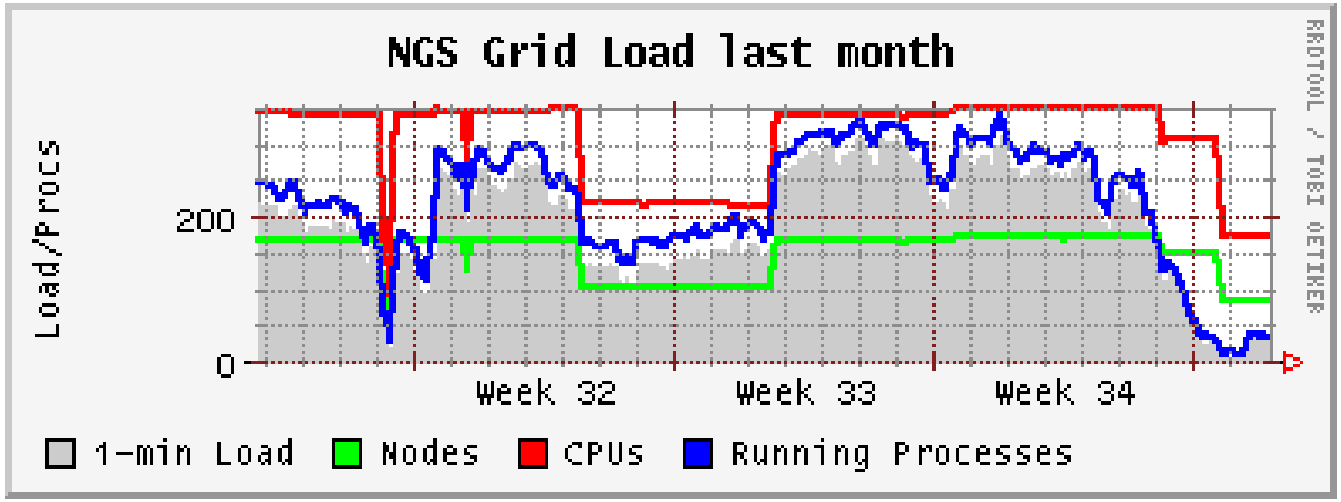


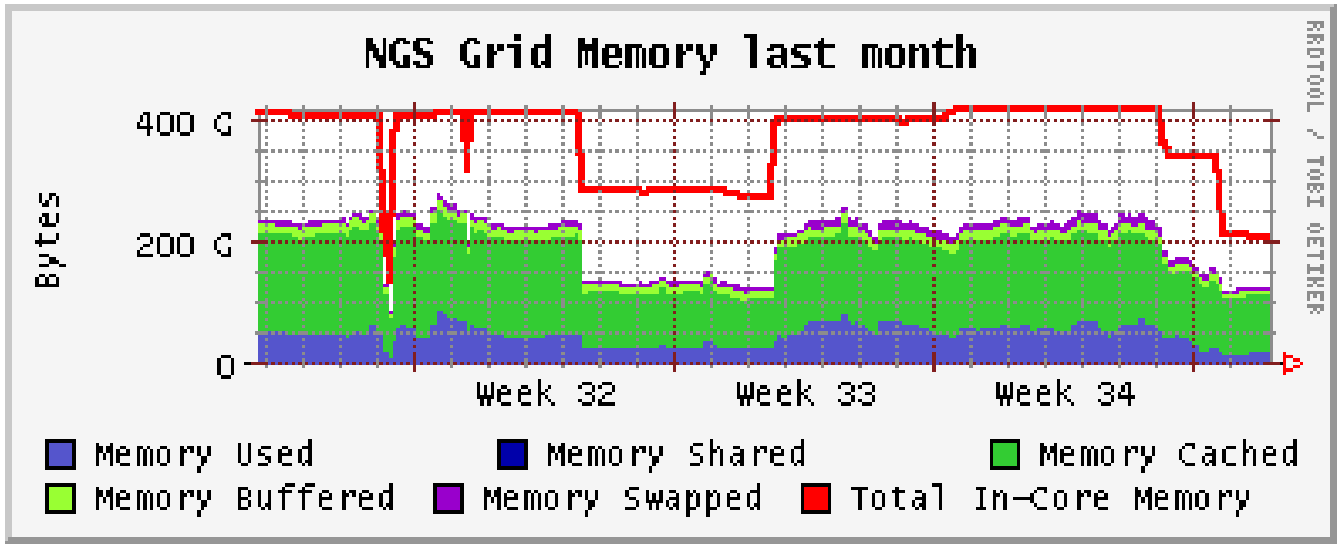


Example load over 1 hour. Low load due to Holiday weekend and system problems

National Grid Service

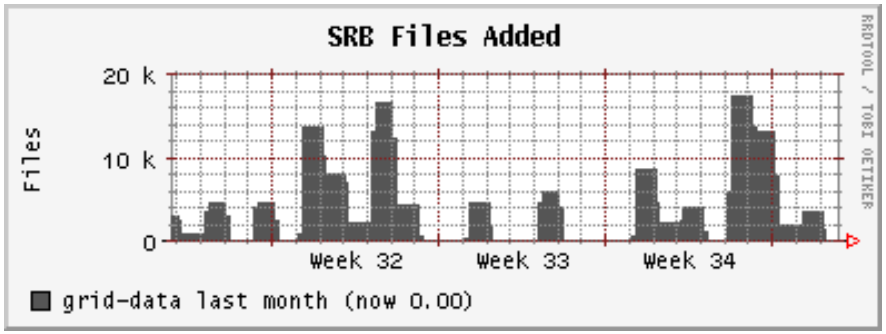
NGS Usage last month. Month prior to 31/08/05

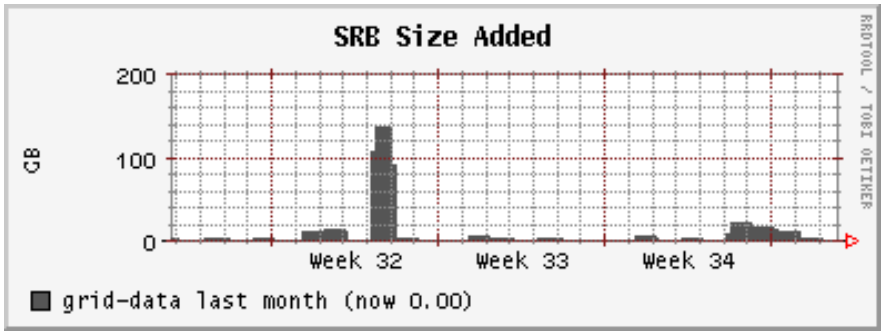


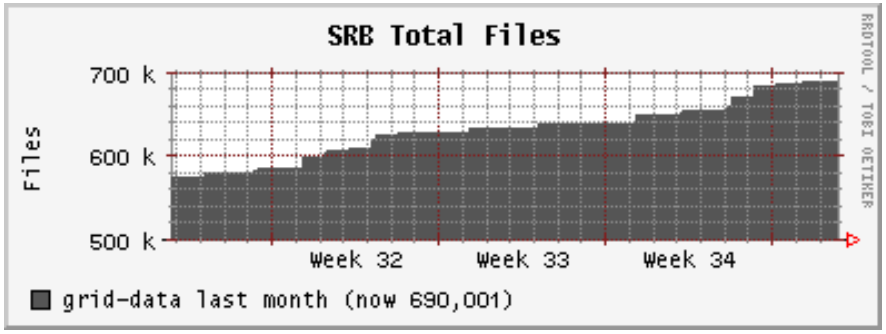


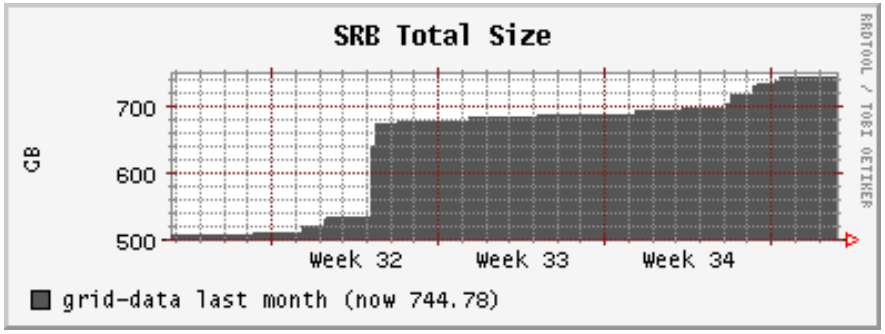
National Grid Service

SRB Storage history for month prior to 31/08/05



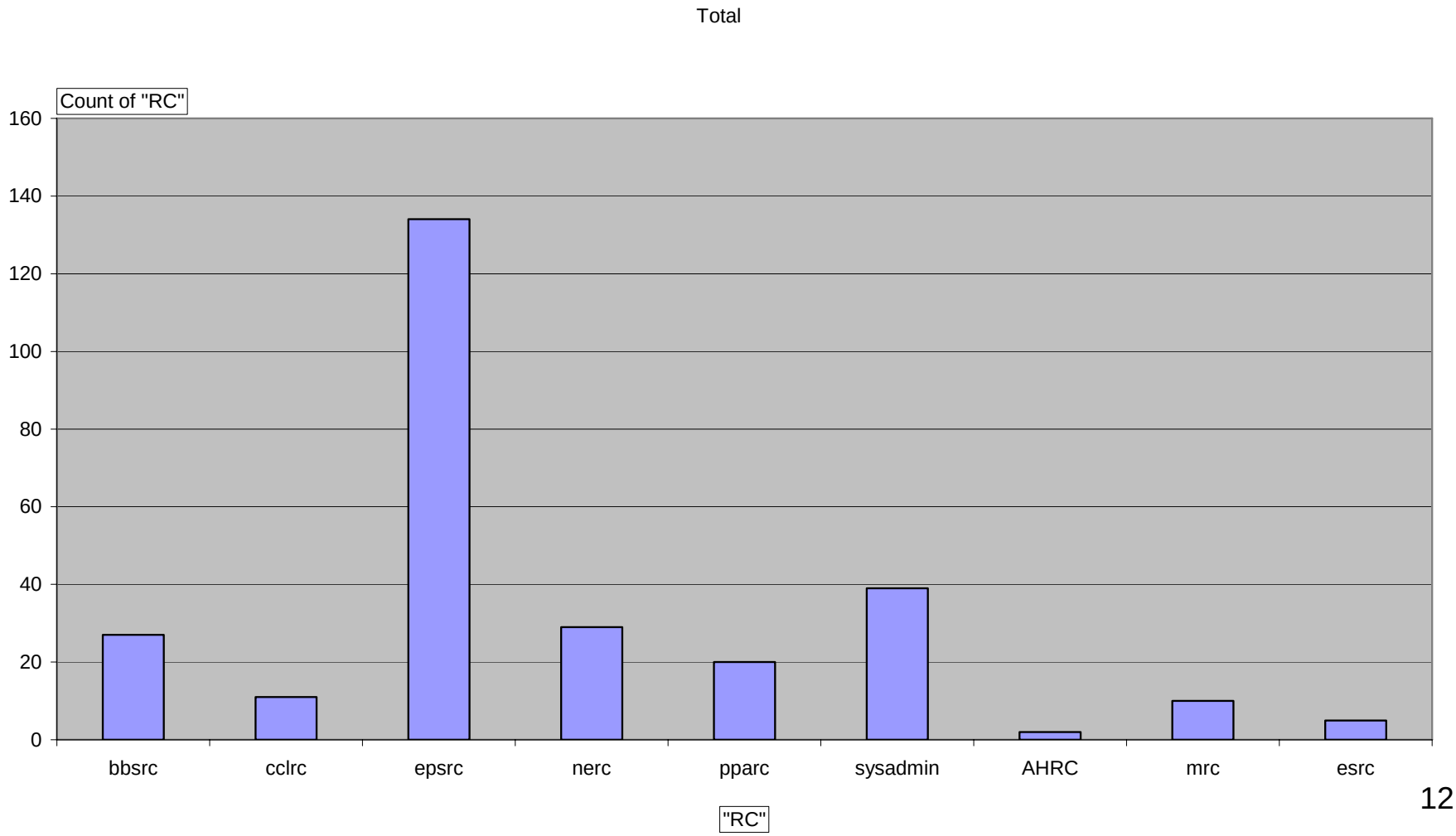




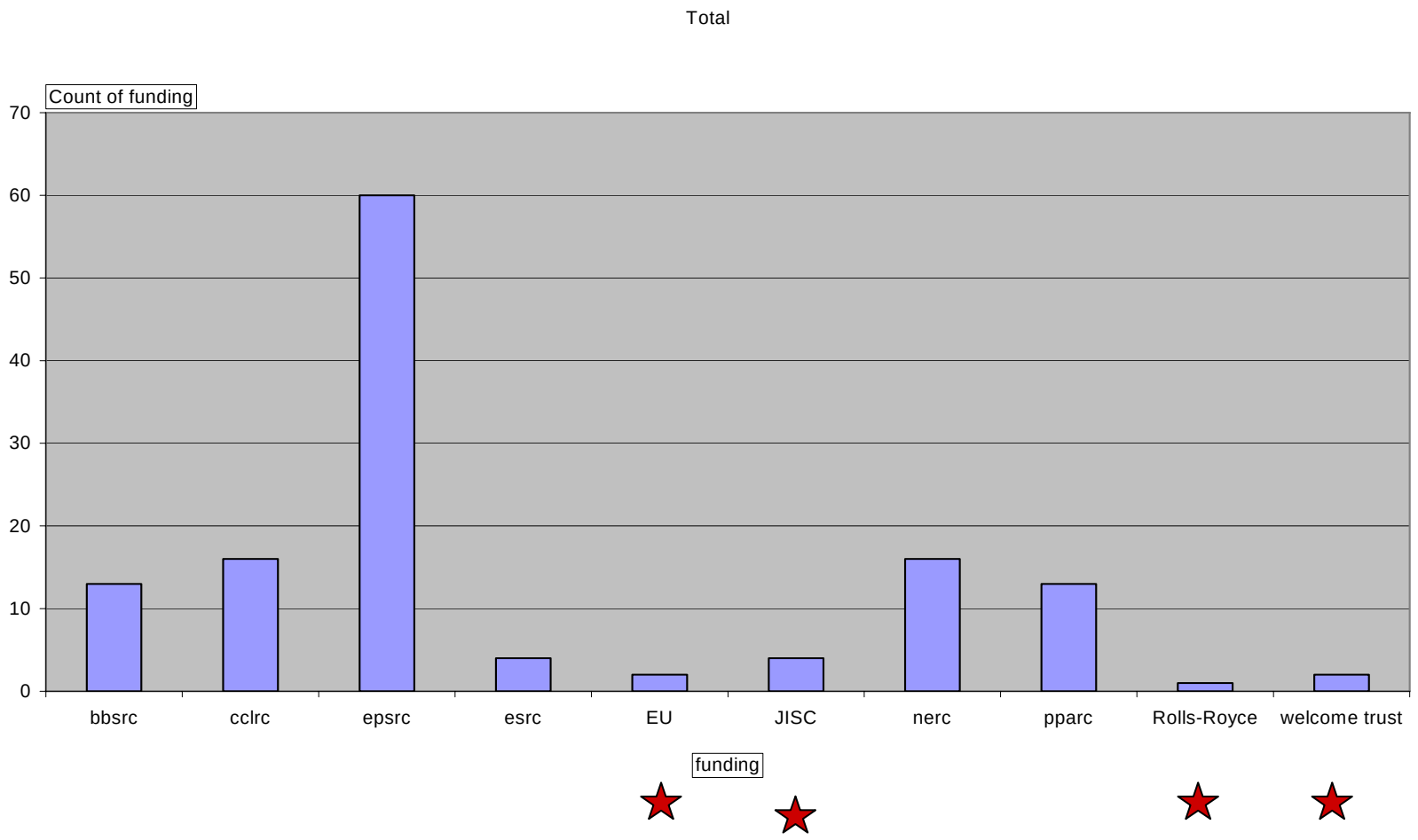


Detailed information -> https://www.ngs.ac.uk/ops/gits/srb/srbreport.txt

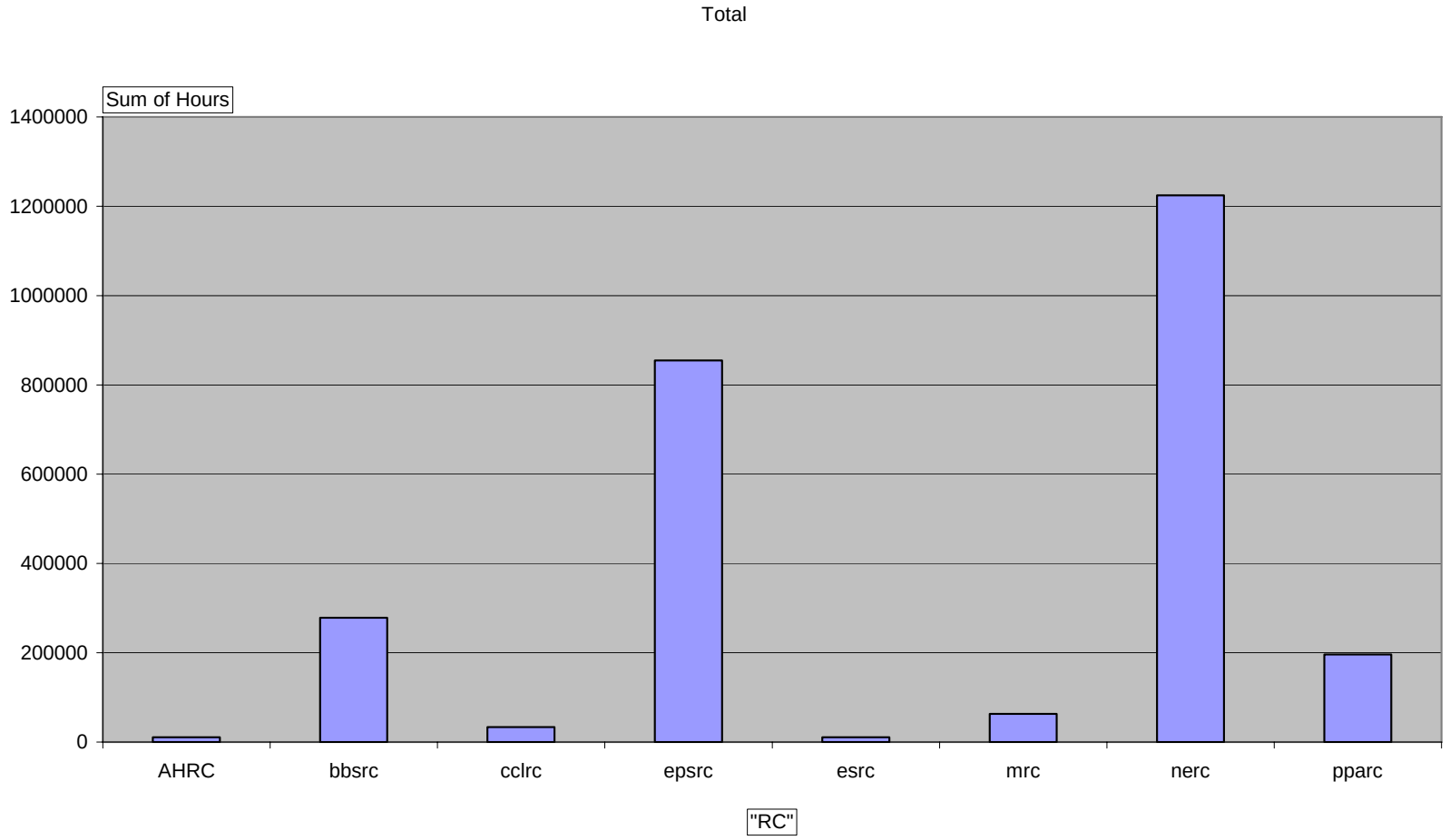
Users by “Research Council”



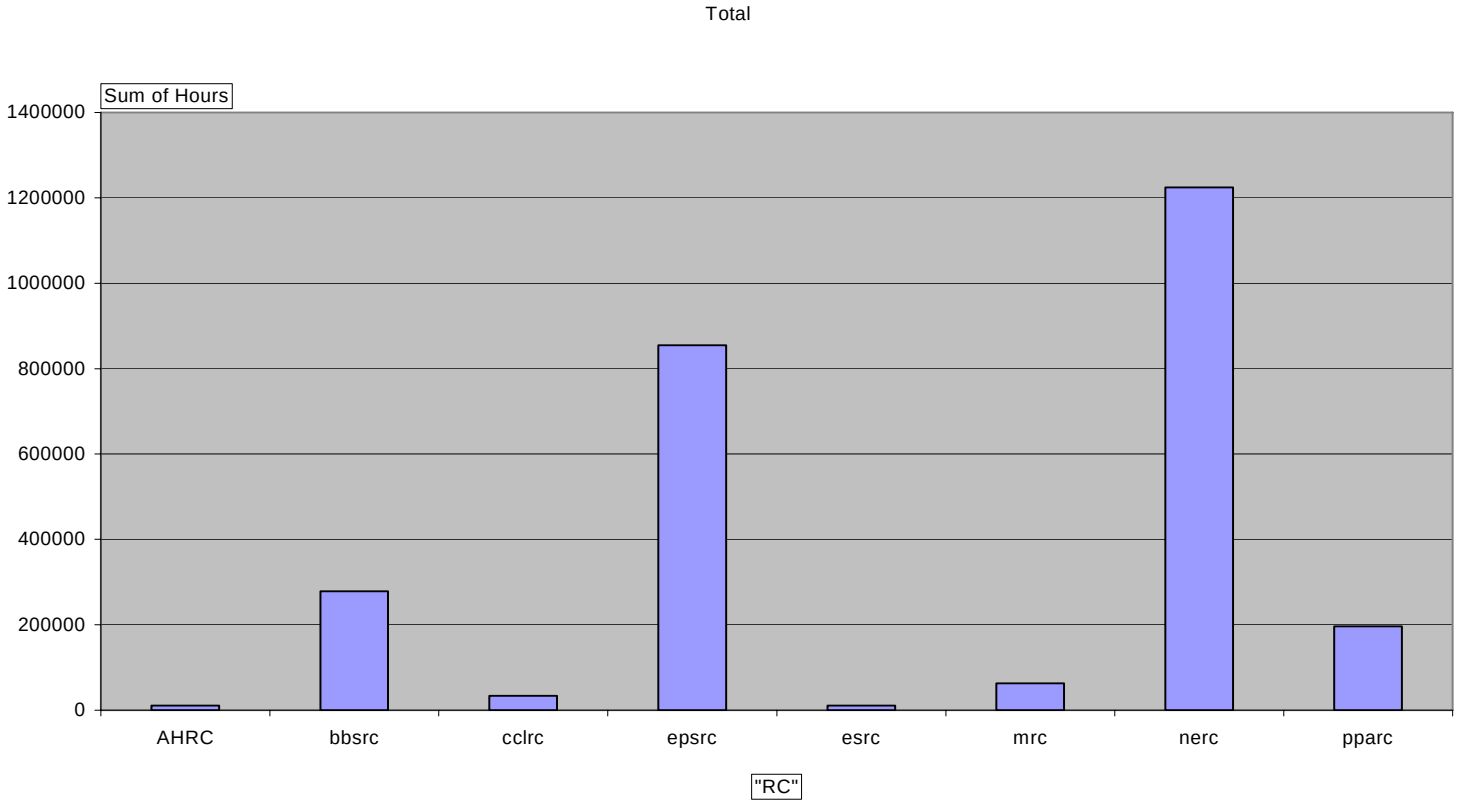
Users known funding



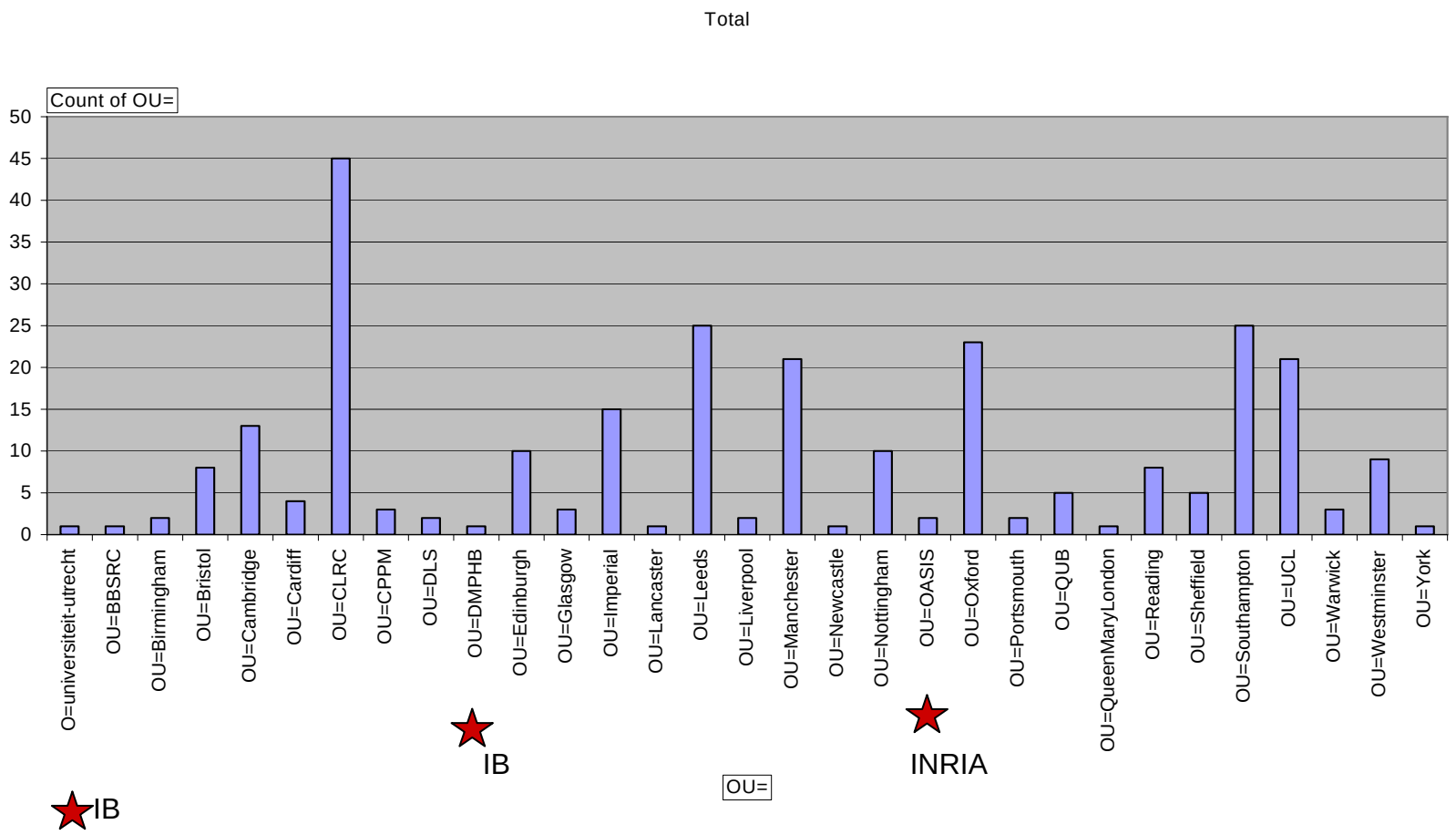
CPU hours requested, by Research Council



Storage requested, by research council



Users by institution



"New" Requirements

All are still hard to satisfy:

- "interactivity"
  - more user control than submit job to batch queues
    - e.g. reservation, co-allocation
  - working with users on ad-hoc (telephone) basis
    - new web form for reservations at CSAR
  - need projects and staff

- Visualization facility
  - some evaluation ongoing in the ETF
  - looking for a "partner" to work with

- metacomputing
  - specific request for PACX-MPI
  - interest in MPICH-G2
    - CSAR & HPCx deploying for joint TeraGrid (NSF) – NGS (EPSRC) projects

Site Reviews

- First Reviews of NGS services
  - Largely internal in this first round
    - but external chair
  - Done Oxford, CCLRC, Manchester, CSAR
    - accepted as very useful by all parties
    - highlighted:
      - documentation
      - monitoring
      - GOSC SLD/SLA
      - data services
  - Report and define future process
  - extend to all partners every 6 monthly

VOMS deployment

- Joint work between GridPP and NGS at Manchester
- Thinking of cross-registering some VOs
  - e.g. QCDGrid

Other Recent Developments

- Improved Monitoring Framework
  - INCA from Teragrid
    - Steve Pickering (Leeds)
    - framework for our own tests
  - TeraGrid looking at integration of GITS tests into framework

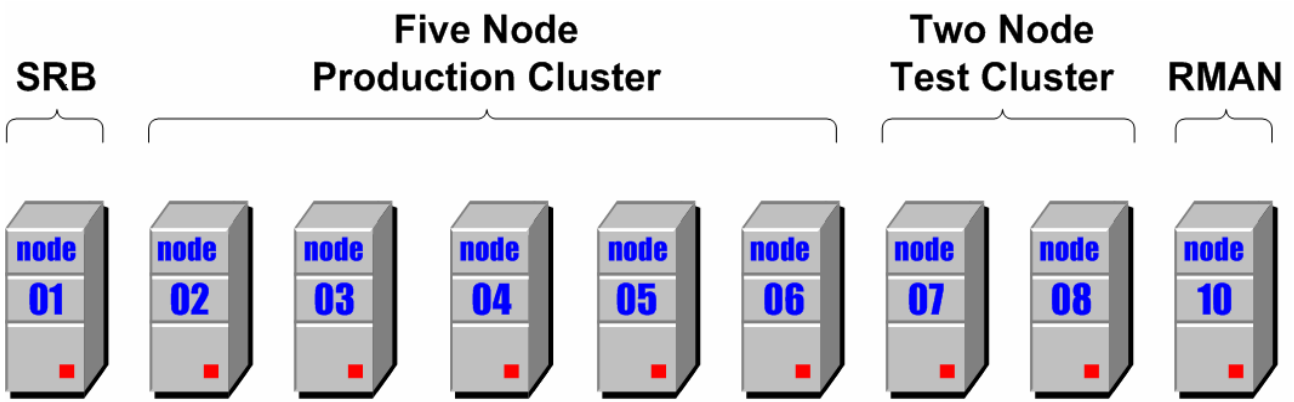
- Accounting
  - Central PBS accounting across all core nodes
  - Finalizing Usage Records feeds to RUS
  - RUS spec now in GGF editor pipeline

- #Jobs "project"

- 1117 nano-particles
- 772 protein folding
- 774 ab-initio protein structure prediction
- 582 lattice-boltmann simulations
- 555 radiation transport (radiotherapy)
- 255 Neutron data analysis
- 242 Godiva
- 228 IXI (medical imaging)
- 228 Biological membranes
- 171 micromagnetics
- 123 Integrative Biology

NGS Oracle Service

- Five node Oracle 9i Real Application Cluster database for NGS users
- Metadata Catalogue (MCAT) Database for Storage Resource Broker (SRB)
- Administration database for back-ups using Oracle Recovery Manager (RMAN), monitoring and statistics



NGS Oracle Users

- (SRB MCAT)
- Geodise, Southampton University
- Ematerials Crystals and Metadata, University College London
- Mircobase, Newcastle University
- NGS Portal
- Integrative Biology Metadata
- SAKAI Portal
- Structure Health Monitoring project, Southampton University

NGS Oracle Current Work and Plans

- Replication of SRB Database to Manchester using Oracle Data Guard for high availability
- Upgrade to Oracle 10g
- Additional users and projects
- Publication of key database statistics including transaction rates, logins, schema size and availability

OGSA-DAI plans

- Currently on OGSA-DAI 4.0
  - based on GT 3.2
- Announced intention to upgrade to OGSA-DAI WSRF 1.0 in September
  - i.e. GT4 flavour
  - some fixes
  - but some issues with multiple connections and streaming data
- Usage is low

ETF Evaluations

- OMII_1 complete
  - Does what it says on the tin
- GT4 complete
  - Service developers view: much better than GT3
  - Claims on interop of pre-WS components with GT2.4.3 hold up
  - Final report now available
- gLite evaluation suspended
  - dependencies on SRM
- CROWN
  - in progress
- United Devices
  - in progress

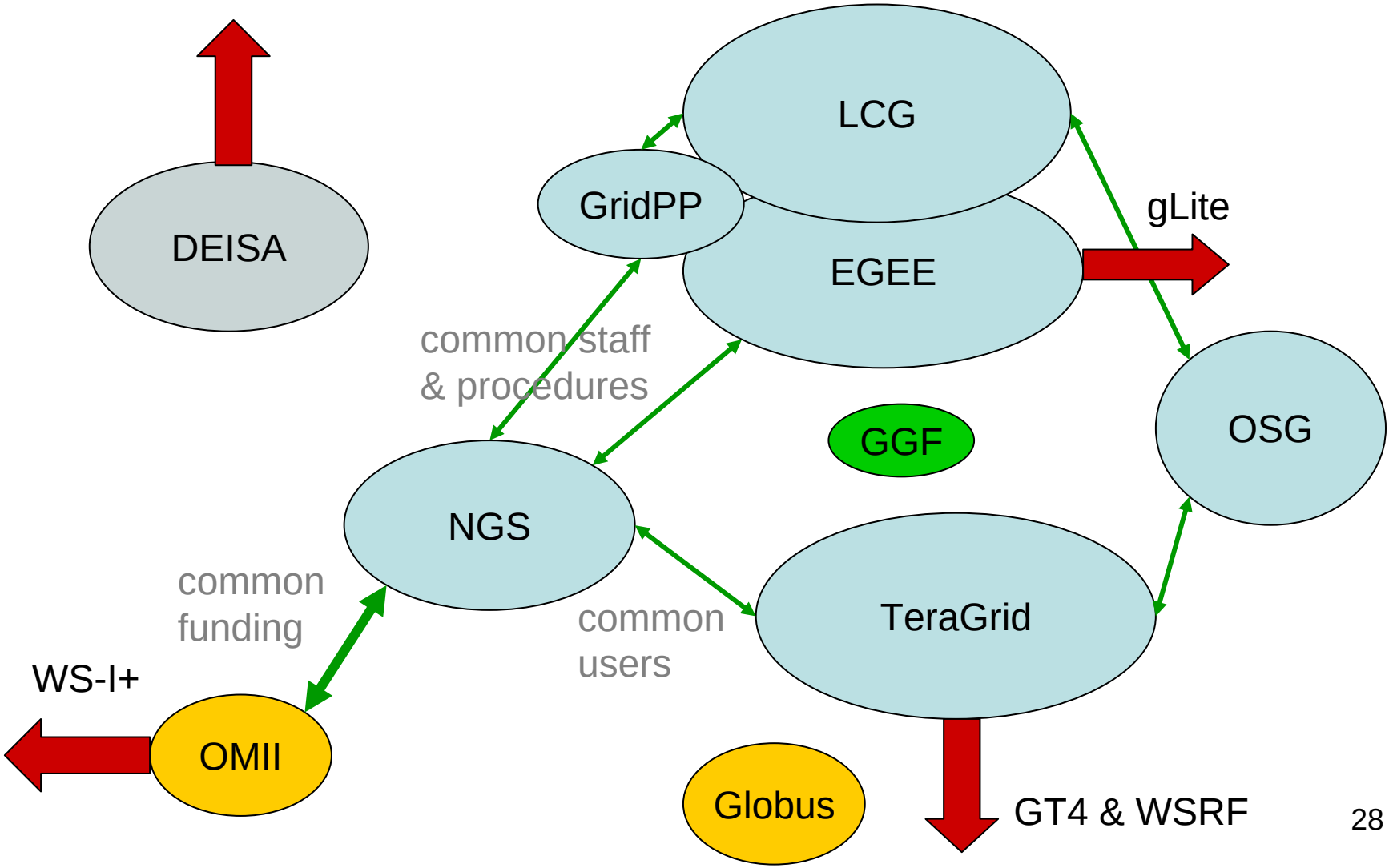
GOSC Plans

Agreed at July GOSC Board:

- GT4 software more reliable.
  - Sufficient compatibility between GT2 and GT4,
  - improvement in stability (cf GT3)
- Currently no plans to deploy a middleware alternative to VDT/GT2
  - Will work with users interested in trialling/testing GT4
- Will update again at end of year.

- Next meeting on 11 October will be face-to-face

Divergent Forces?



NGS usage policy

- Current policy is driven by need to
  - encourage new "grid" users
  - ensure that NGS resources are used effectively
- Balance between "initial trial" and "production Service"

- Approaching the point where the NGS is full.
  - Existing users vs new users ?
    - Now, there is no "elsewhere" of course so this raises a real issue of policy and strategy.
  - One way forward is to make a more severe peer review process happen.
    - Not actually resourced to do this at present – and very reluctant

"It is interesting how these grid tools have generated a large and active community so quickly – I suspect much more quickly and wide reaching than the national supercomputers – something I always thought would happen – there is much more science to be done by investing in large national clusters that are easy to use than big iron!"

Some Key Questions

- Should there be any distinction between the NGS and GOSC ?
- What should the relationship be between GOSC and NGS ?
- What reporting structure and/or bodies are required in future?
- Should directly funded core nodes automatically become members of the GOSC ?
- How should access to the NGS be regulated?
- Should the NGS or GOSC support a formal peer review process?
- How should reviewers for NGS access be appointed?
- Should the core nodes be upgraded or re-tendered ?
- How closely coupled should the core nodes be ?
- How prescriptive should the second phase be ?
- What should the balance between different NGS nodes be ?
- Should the NGS remain free
- Should the NGS allow commercial partners
- What services should be provided by the GOSC ?
- What development role/effort should the GOSC have?
- What is the scope of the GOSC, reflecting the goals of the NGS ?
- Should the GOSC provide centrally funded gateway or related machines ? and if so under what criteria ?
- How is supporting 100,000s of users to be provided?
- What (sort of) applications should be centrally provided?

  - **There are doubtless many more**
  - **We can not do everything, and need to define and understand the boundary conditions to meet user needs**

eSI Theme Event - Spatiotemporal Databases for Geosciences, Biomedical sciences and Physical sc - Microsoft Internet Explorer p

File Edit View Favorites Tools Help

Back Search Favorites

Address http://www.nesc.ac.uk/esi/events/608/index2.cfm Go Links

01 November, 05 09:00 - 02 November, 05 16:30, e-Science Institute 15, South College Street, Edinburgh
Organiser: Neil Geddes

| e-SI events | theme information | registration |

There are many research communities in the UK using shared or published data which requires multi-dimensional properties. Examples include the geospatial or temporal information used in oceanography, climatology and geosciences, the nano-scale 3D data in biological molecular systems, novel material structures or nano-engineering and the meso-scale spatio-temporal structures of living organisms, organs and developmental biology.

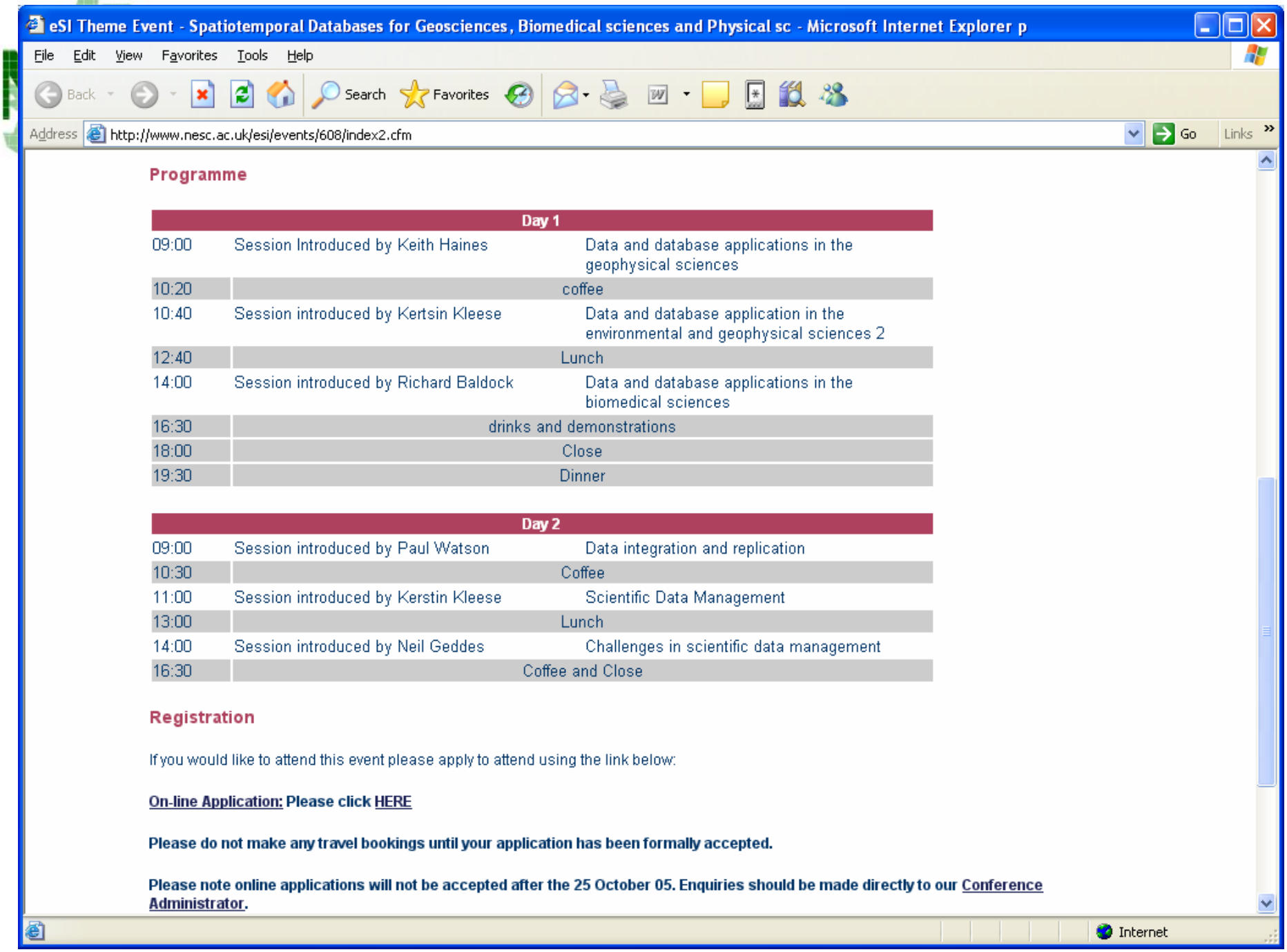
The meeting should attract those that have identified requirements, discovered techniques or are developing products or services that apply to this data. By bringing together the diverse user communities we hope to better understand the current state of the art and priority issues in spatio-temporal data management for scientific research. The commercial and academic technology providers present will share their current solutions and planned developments so that users will be better informed about alternative technical solutions and imminent developments. The related topics of metadata management, data replication services and distributed caching will also be addressed.

We hope that the meeting will provide an opportunity to form alliances and to plan collaborative strategies for further developing our capabilities in managing and exploiting complex scientific data.

Target Audience

Technical specialists from databse vendors and (e-)Scientists who already, could, or intend to use databases or other data management tools extensively to facilitate their research.

http://www.nesc.ac.uk/esi/themes/theme_02/index.htm Internet



Programme

Day 1		
09:00	Session Introduced by Keith Haines	Data and database applications in the geophysical sciences
10:20	coffee	
10:40	Session introduced by Kertsin Kleese	Data and database application in the environmental and geophysical sciences 2
12:40	Lunch	
14:00	Session introduced by Richard Baldock	Data and database applications in the biomedical sciences
16:30	drinks and demonstrations	
18:00	Close	
19:30	Dinner	

Day 2		
09:00	Session introduced by Paul Watson	Data integration and replication
10:30	Coffee	
11:00	Session introduced by Kerstin Kleese	Scientific Data Management
13:00	Lunch	
14:00	Session introduced by Neil Geddes	Challenges in scientific data management
16:30	Coffee and Close	

Registration

If you would like to attend this event please apply to attend using the link below:

On-line Application: Please click HERE

Please do not make any travel bookings until your application has been formally accepted.

Please note online applications will not be accepted after the 25 October 05. Enquiries should be made directly to our Conference Administrator.

eSI Theme Event - Oracle Corporation and the e-Science Institute Seminar - Temporal Database in - Microsoft Internet Explorer p

File Edit View Favorites Tools Help

Back Search Favorites

Address http://www.nesc.ac.uk/esi/events/601/ Go Links

Oracle Corporation and the e-Science Institute Seminar - Temporal Database in Depth: Time and the Data Warehouse

e-Science Institute

e-Science

In association with

ORACLE

03 November, 2005 08:45 - 17:00, e-Science Institute 15, South College Street, Edinburgh
Organiser: Anna Kenway

| e-SI events | theme information | registration |

Temporal Database in Depth: Time and the Data Warehouse

This technical seminar is based on the book "Temporal Data and the Relational Model" by Date, Darwen and Lorentzos.

KEYNOTE ADDRESS - CHRIS DATE Chris Date is one of the world's best-known specialists in relational database technology, and is making one of his rare visits to the UK.

RESPONSE - ROB SQUIRE - ORACLE CORPORATION Rob Squire is a UK based consultant actively involved in research for Oracle into techniques for managing time varying data within the relational model.

Synopsis

Chris Date will argue that the plummeting cost of storage and the widespread adoption of data warehouse technology have led to an increasing interest in temporal databases. The concept of maintaining and processing historical data has become a reality for many organisations. As a consequence, the ability to deal properly with the time dimension in databases has become an increasingly important practical problem.

At present DBMS products offer nothing to help in this area.

Internet

The End