



NGS

National Grid Service

101010001000000100100
101010001000000100100

?



101010001000000

United Kingdom National Grid Service

Introduction to Data Management

JISC e-Research

30 Jan 2009

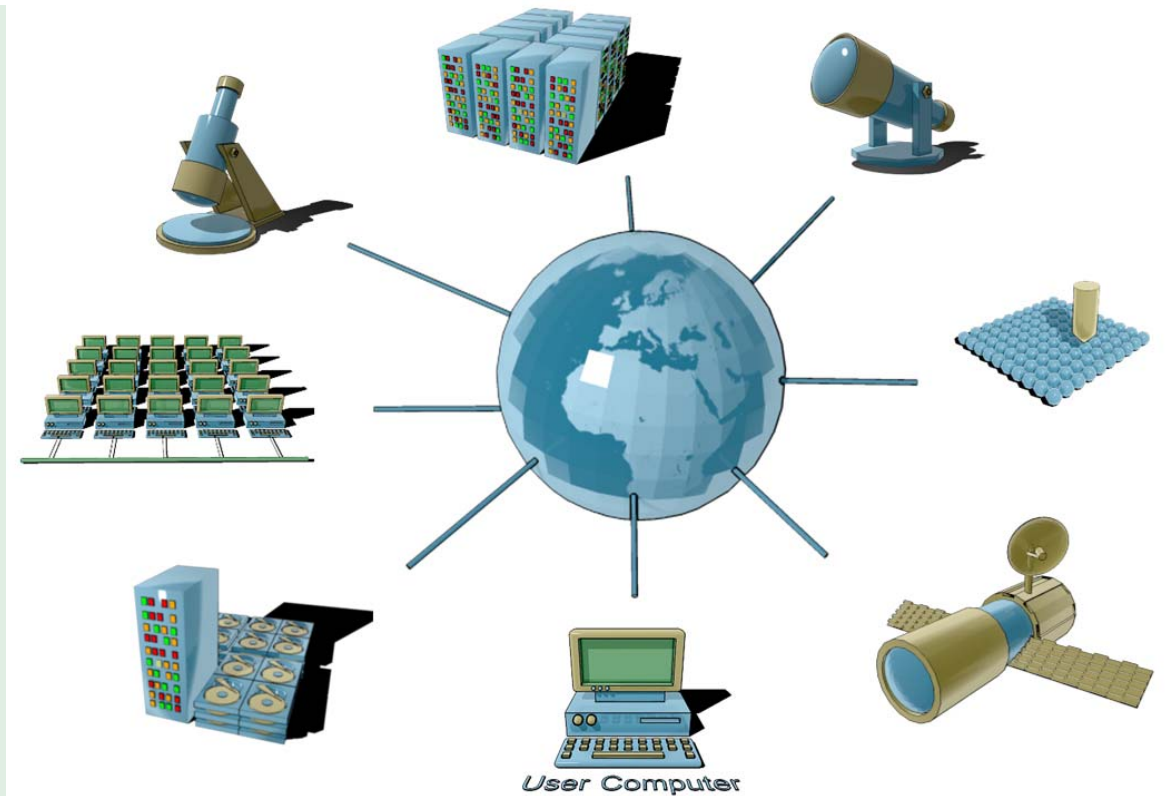
J Jensen, STFC

< jens . jensen @ stfc . ac . uk >

What is the “Grid”

- Distributed
- Heterogeneous
- High performance
- Resources brokered
- Security
- Interoperable (usually)
- Accounting
- Monitoring

What is the Grid?



Why the Grid?

- High performance/throughput computing
- Data storage
- Interoperability
- Security

What is the NGS?

- 4 core sites, funded by JISC and EPSRC
 - Running VDT Globus on RHEL
- 7 partner sites
- 12 affiliated sites
- Integrated support, monitoring, accounting
- EGEE resource broker
- EGEE resources: APEL, GOCDB,

Why the NGS?

- High throughput computing
 - “Many small problems”
- State-of-the-art middleware
 - Proven robust and usable!
- Useful applications and support libraries
 - Many different user communities
- Free at the point of use

PRACTICAL STUFF

How to get access

- Writing applications
 - The proposal
 - And the software ones
- Authentication
 - Traditionally with identity certificates
 - Now with Shibboleth access
- Access
 - Accounts set up, Virtual Organisations (VOs)

Accessing NGS

Command line

- Powerful
- X.509 certificates
- X.509 certificates
- Learning curve

Portal

- Friendly (friendlier)
- Single Sign-on
- Some familiarisation
- Generic and special

What's in there

- Bioinformatics
- Engineering
- Analysis/Statistics packages
- Biomed
- Chemistry
- Astrophysics
- Image analysis
- ...and there are other portals with more

DATA TRANSFER TOOLS

GridFTP

- Data transfer workhorse
- Used by almost every single grid in the world;
- Very high performance transfers
- Security features

DATA STORAGE

Oracle database

- It's, er, a database implementation
- It can store data
 - Some datatypes more suitable for databases
 - Than others
- Helpful experts to help you set up schemas etc

Storage Resource Broker

- It's a “data grid” implementation
- Widely used to store lots of different things
- Basic metadata capabilities
- iRODS deployment – future
 - ASPiS project: Shib-to-iRODS with provenance metadata management

OGSA-DAI

- Gridified XML database
- Standardised in OGF

Disk...

- It's, er, disk
- Different types of disk
 - Plain ol' scratch space
 - Raided high performance disk
 - Backed up to tapestore disk

Other Aspects

- Standards
 - Whenever possible, using standards-based schemas and protocols
 - Preferably open standards
- Interoperation
 - Active in many interoperation activities
 - Collaborating with other Grids
 - GridPP, TeraGrid, EGEE,

Other Aspects

- “External” Applications
 - Matlab
 - Sciency things
- Bespoke applications development
 - E.g. Visualisation

Conclusion

- ... at last
- Come talk to us
 - User forum
 - A&H+NGS Workshops
 - Experts available
- Feel free to contact me: