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

http://www.ngs.ac.uk

UK e-Infrastructure

Mike Mineter
mjm@nesc.ac.uk

National e-Science Centre
http://www.nesc.ac.uk/

JISC

PPARC
http://www.pparc.ac.uk/

e-Science

eGee
http://www.eu-egee.org/

Acknowledgements

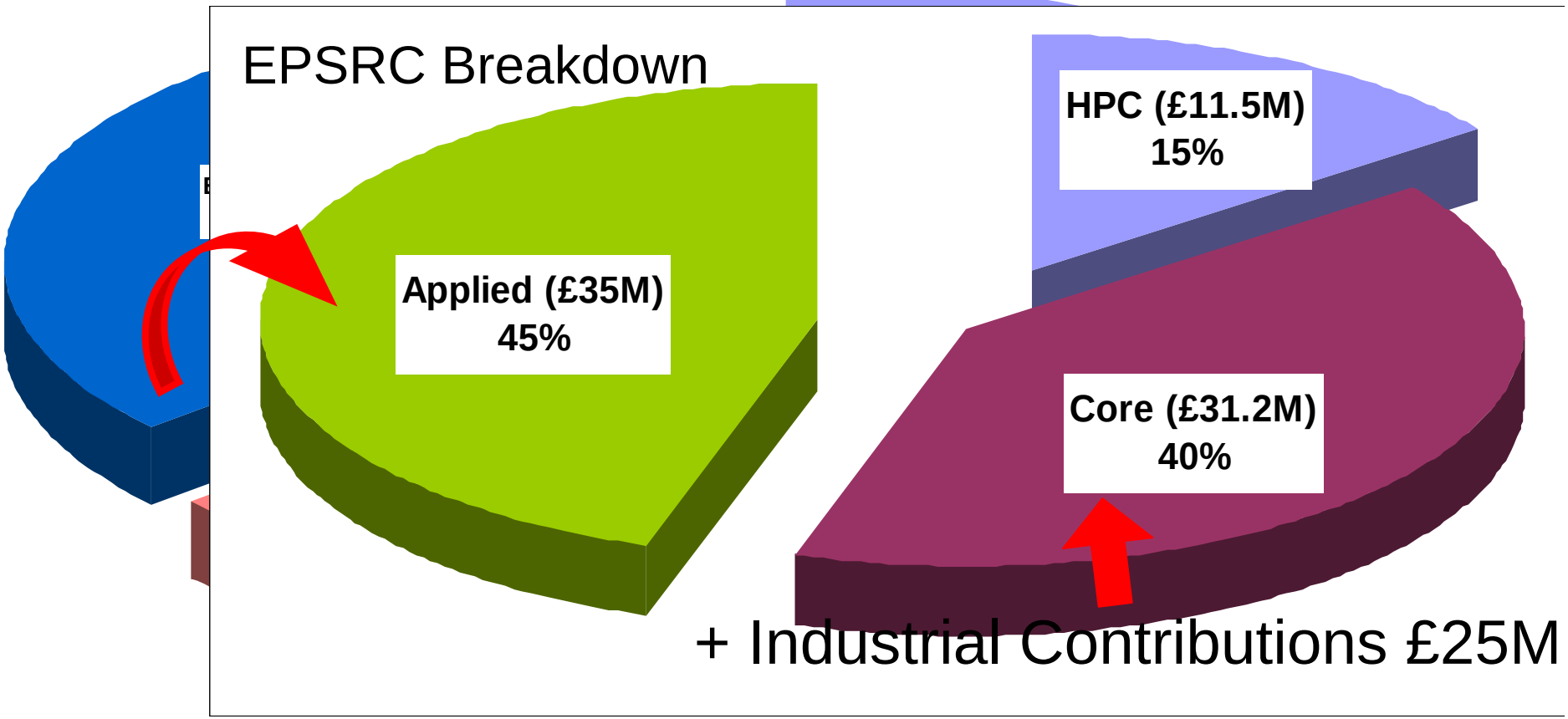
- For slides and information:
- NGS and GOSC - Stephen Pickles, Technical Director of GOSC
- OMII – Steven Newhouse
- JISC – Ann Borda, Sarah Porter, Sara Hassan, Shirley Wood
- Data centres – Peter Burnhill
- Integrative Biology

Overview

- The UK e-science programme
- The National Grid Service
- GOSC – Grid Operations Support Centre
- OMII – Open Middleware Infrastructure Institute
- JISC - Joint Information Service Committee

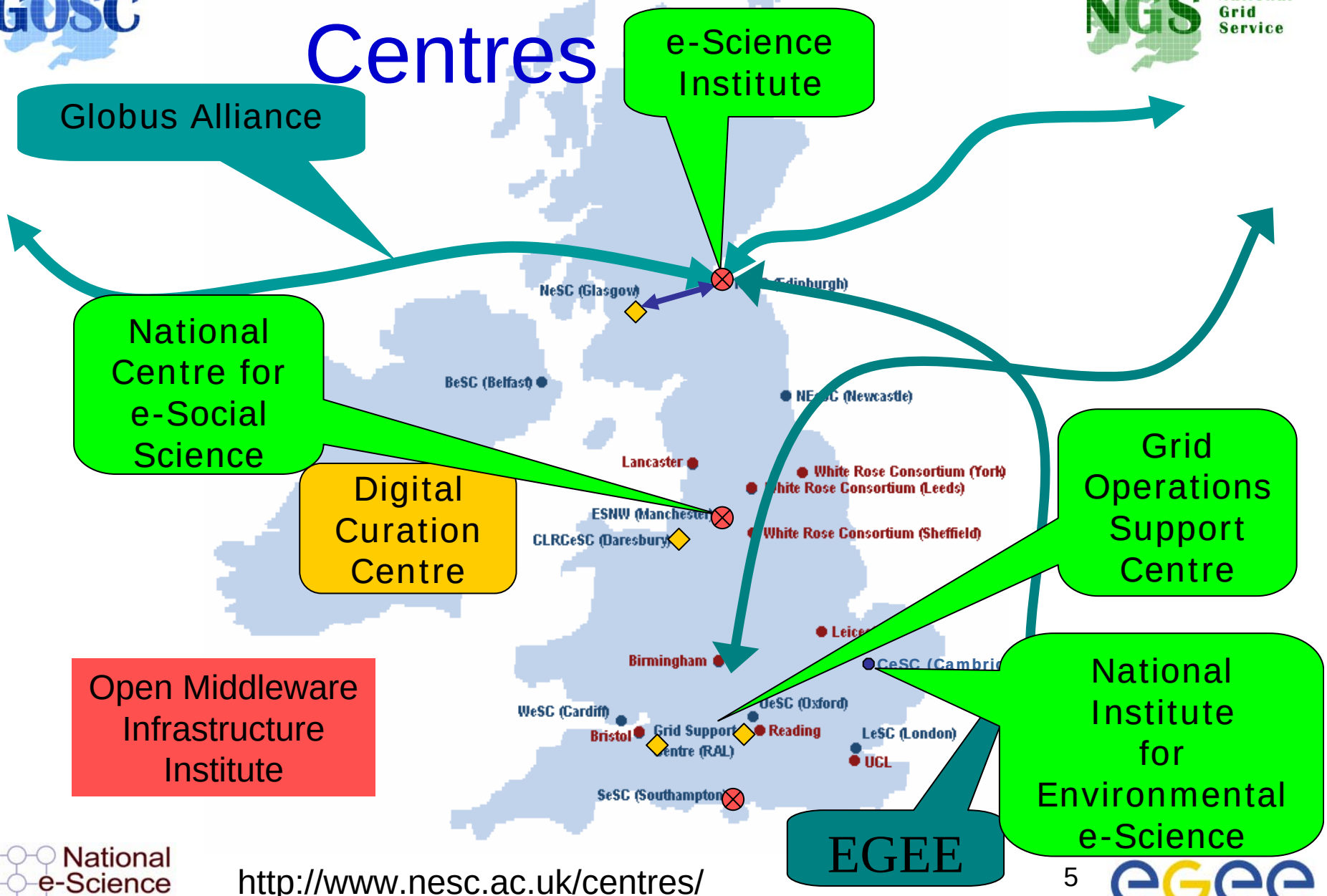
UK e-Science Budget (2001-2006)

Total: £213M + £100M via JISC



Source: Science Budget 2003/4 – 2005/6, DTI(OST)

The e-Science Centres



National e-Science Centre

http://www.nesc.ac.uk/centres/

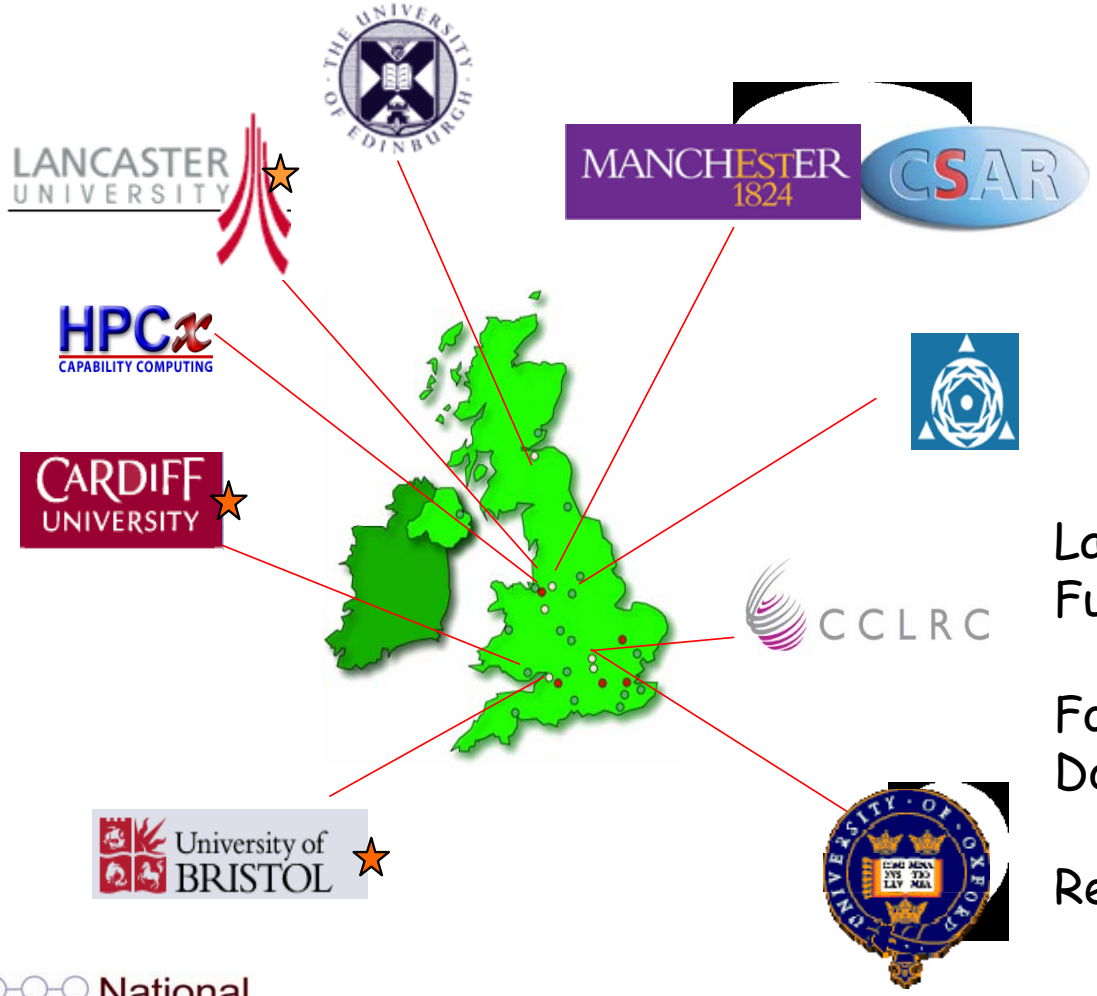
egee

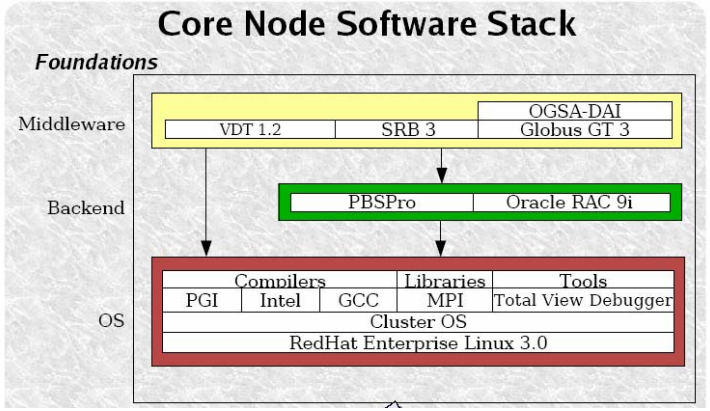
The National Grid Service

The National Grid Service

- The core UK grid, resulting from the UK's e-Science programme.
- Production use of computational and data grid resources.
- Supported by JISC, and is run by the Grid Operations Support Centre (GOSC).

The National Grid Service

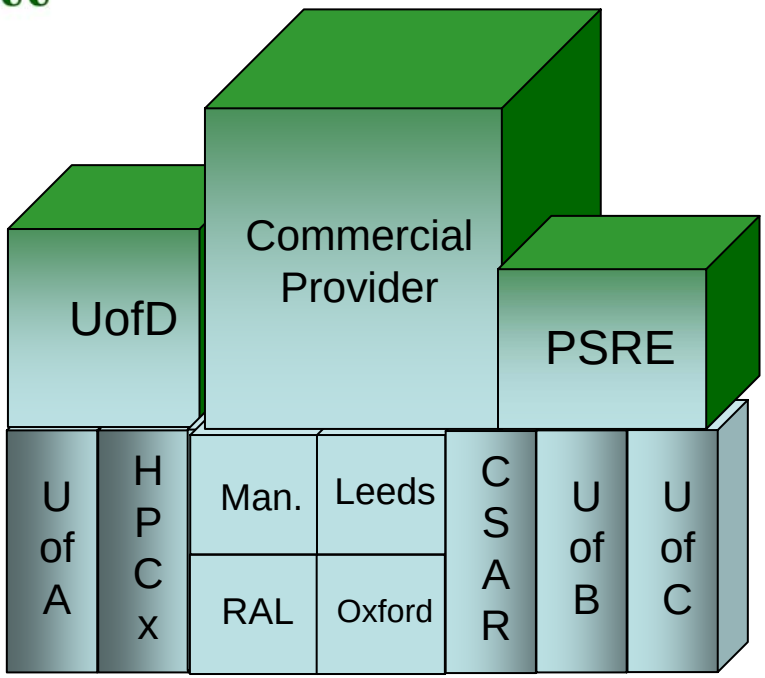




Launched April 2004
Full production - September 2004

Focus on deployment/operations
Do not do development

Responsive to users needs



NGS Core Nodes: Host core services, coordinate integration, deployment and support
+free to access resources for all VOs. Monitored interfaces + services

NGS Partner Sites: Integrated with NGS, some services/resources available for all VOs
Monitored interfaces + services

NGS Affiliated Sites: Integrated with NGS, support for some VO's
Monitored interfaces (+security etc.)

New partners

Over the last year, three new full partners have joined the NGS:

- Bristol, Cardiff and Lancaster
- Further details of resources can be found on the NGS web site: www.ngs.ac.uk.

- Resources committed to the NGS for a period of at least 12 months.

- The heterogeneity introduced by these new services has
  - provided experience in connecting an increasingly wide range of resources to the NGS
  - presented a challenge to users to make effective use of this range of architectures
  - basic common interface for authentication+authorisation is the first step towards supporting more sophisticated usage across such a variety of resources.

NGS Facilities

- **Leeds** and **Oxford (core compute nodes)**
  - 64 dual CPU intel 3.06GHz (1MB cache). Each node: 2GB memory, 2x120GB disk, Redhat ES3.0. Gigabit Myrinet connection. 2TB data server.
- **Manchester** and **Rutherford Appleton Laboratory (core data nodes)**
  - 20 dual CPU (as above). 18TB SAN.
- **Bristol**
  - initially 20 2.3GHz Athlon processors in 10 dual CPU nodes.
- **Cardiff**
  - 1000 hrs/week on a SGI Origin system comprising 4 dual CPU Origin 300 servers with a Myrinet™ interconnect.
- **Lancaster**
  - 8 Sun Blade 1000 execution nodes, each with dual UltraSPARC IIICu processors connected via a Dell 1750 head node.
- **HPCx** and **CSAR**
  - …

For more details: http://www.ngs.ac.uk/resources.html

NGS software

- Computation services based on Globus Toolkit 2
  - Use compute nodes for sequential or parallel jobs, primarily from batch queues
  - Can run multiple jobs concurrently
- Data services:
  - Storage Resource Broker:
    - Primarily for file storage and access
    - Virtual filesystem with replicated files
  - "OGSA-DAI": Data Access and Integration
    - Grid-enabling databases (relational, XML)
  - NGS Oracle service
  - GridFTP for efficient file transfer
- Authorisation and Authentication using GSI
- Portal to support collaboration and ease use

Data services on NGS

Simple data files

- Middleware supporting
  - **Replica files**
  - **Logical filenames**
  - **Catalogue**: maps logical name to physical storage device/file
  - **Virtual filesystems**, POSIX-like I/O
- **Storage Resource Broker**
  - **(3.3.1 since December 2005)**

Structured data

  - RDBMS, XML databases
  - Often pre-existing
  - Do NOT want to replicate
- Require *extendable* middleware tools to support
  - Move computation near to data
  - Underpin integration and federation
- **OGSA –DAI**
  - DAI: Data access and integration

Grids and OGSA-DAI

- Early VOs used their own data
  - Experiments
  - Simulations
- These data were held on grid resources
  - You join the VO and use their grid
- Many communities have large databases
- Duplicate these??? NO!!
- Build links from Grids
- Effect: opens new domains for e-Research

NGS: Gaining Access

NGS nodes

- data nodes at RAL and Manchester
- compute nodes at Oxford and Leeds
- partner nodes at Bristol, Cardiff and Lancaster

- all access is through digital X.509 certificates
  - from UK e-Science CA
  - or recognized peer

National HPC services

- HPCx

- CSAR

- Must apply separately to research councils

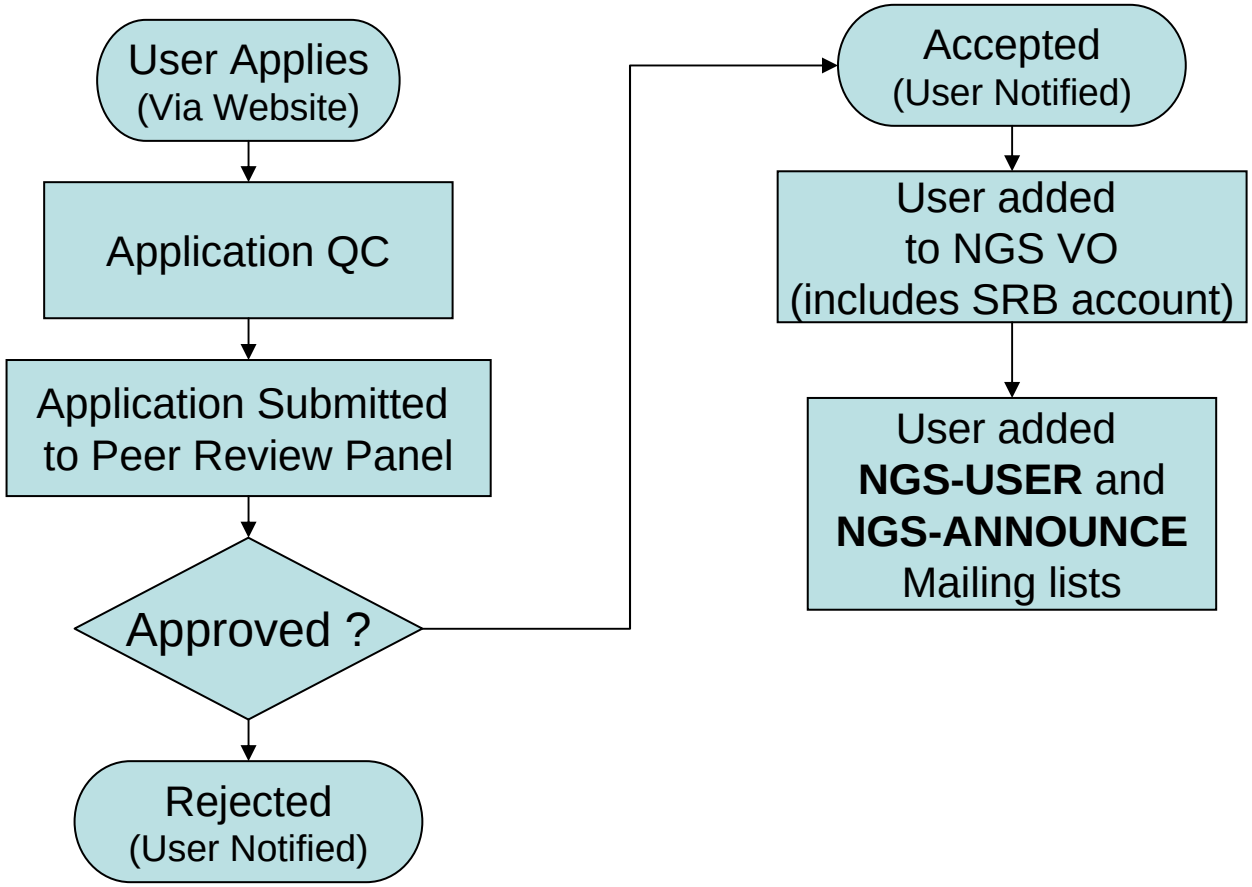
- Digital certificate and conventional (username/ password) access supported

Joining the NGS - Users

- Users
  - Most apply for NGS accounts as individuals
  - Projects (Virtual Organisations) – talk to GOSC
    - NGS provides a basis for collaboration with UK colleagues

free to UK academic users!

User Registration (Process)



Why Join?

- Users increasingly want resources as services and not as complicated bits of kit
  - common interfaces across a range of facilities
- Funders of regional and national facilities want common interfaces to lower barriers to access
- By joining you leverage the national expertise in running these services
  - technical advice and support
  - security procedures and incident response
  - tools to help monitor and patch
  - Get it at lower cost by joining the NGS
- Members get a say in the technical decisions about the NGS

How Do I Get A Certificate?

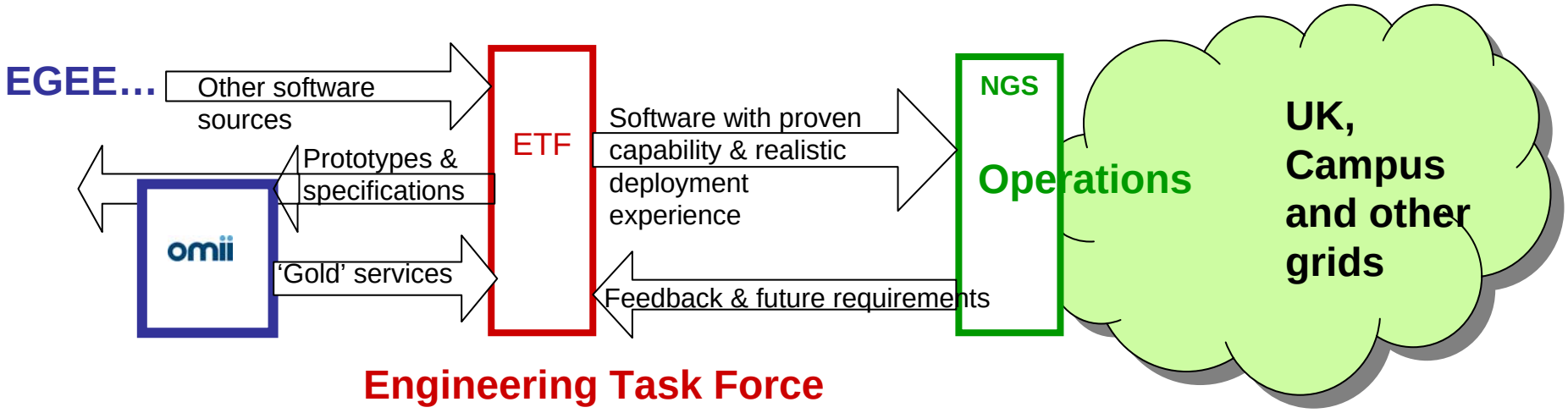
- You need a valid UK certificate before applying for an NGS account:
  - https://ca.grid-support.ac.uk, the UK Certificate Authority.
  - You will probably need to provide non-electronic proof of identity to your local representative of the CA.
    - For example: show your passport.
    - See http://www.grid-support.ac.uk/archive/ca/ralist.htm.
  - Always keep this certificate secure.
    - E.g. DO NOT LEAVE IT ON ANY NGS CORE NODE !
    - Do store it on a USB drive, that you keep safe!!

Projects and VOs

- Just need access to compute and data resources for users in your project?
  - Currently, mainly applications from individuals
  - project-based applications being dealt with case-by-case, as procedures are established – for these, talk to GOSC!
- Want to host your data on NGS?
  - consider SRB, Oracle, or OGSA-DAI
  - NGS maintains infrastructure
  - you populate and manage data
- Want to use NGS resources to provision services, portals for a community of users?
- Want researchers to access your data?

Managing middleware evolution

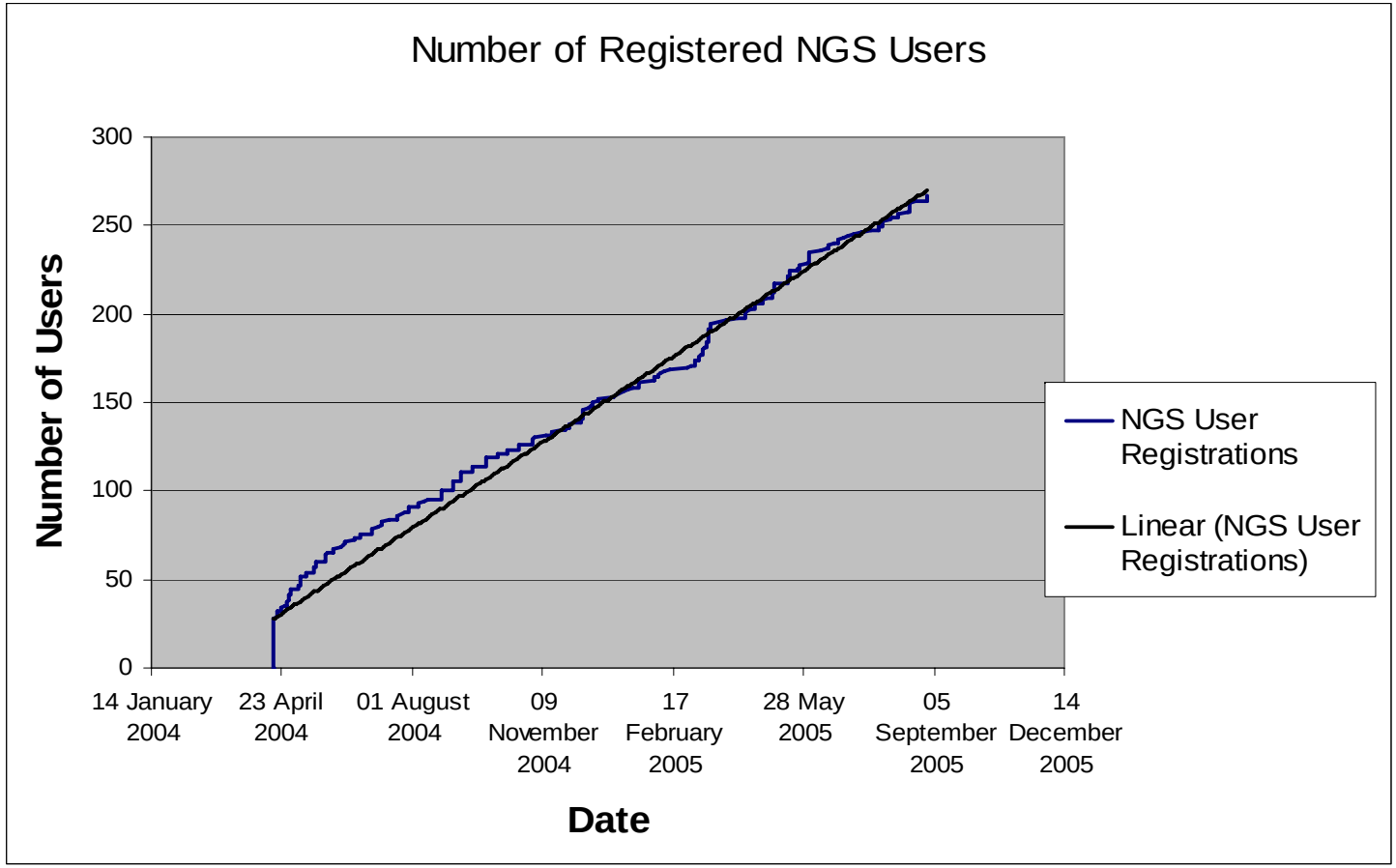
- Important to coordinate and integrate with deployment and operations work in EGEE and similar projects.
- Engineering Task Force makes recommendations for deployment on NGS.



Engineering Task Force

Deployment/testing/advice

NGS Users


Number of Registered NGS Users
Number of Users
300
250
200
150
100
50
0
14 January 2004
23 April 2004
01 August 2004
09 November 2004
17 February 2005
28 May 2005
05 September 2005
14 December 2005
Date
NGS User Registrations
Linear (NGS User Registrations)

NGS Organisation

- Operations Team
  - led by Andrew Richards (RAL)
  - representatives from all NGS core nodes
  - meets bi-weekly by Access Grid
  - day-to-day operational and deployment issues
  - reports to Technical Board
- Technical Board
  - led by Stephen Pickles
  - representatives from all sites and GOSC
  - meets bi-weekly by Access Grid
  - deals with policy issues and high-level technical strategy
  - sets medium term goals and priorities
  - reports to Management Board
- GOSC Board meets quarterly
  - representatives from funding bodies, partner sites and major stakeholders
  - sets long term priorities

NGS: Key facts

- Production: deploying middleware after selection and testing – major developments via Engineering Task Force.
- Evolving:
  - Middleware
  - Number of sites
  - Organisation:
    - VO management
    - Policy negotiation: sites, VOs
- International commitment – EGEE, Teragrid
- Gathering users' requirements – National Grid *Service*

Grid Operations Support Centre

GOSC

The Grid Operations Support Centre is a distributed “virtual centre” providing deployment and operations support for the UK e-Science programme.

GOSC Roles

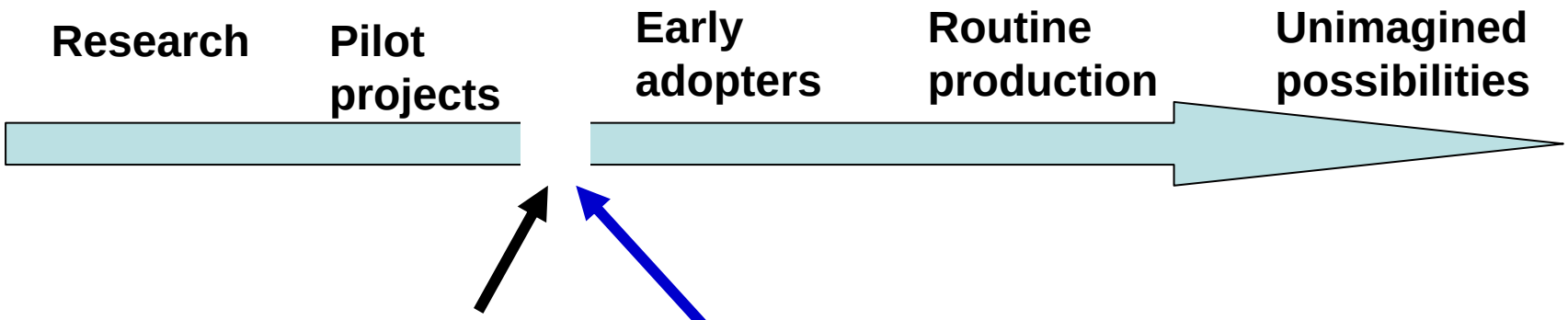
- **UK Grid Services**
  - **National Services**
    - Authentication, authorization, certificate management, VO management, security, network monitoring, help desk + support centre.
    - support@grid-support.ac.uk
  - **NGS Services**
    - Job submission, data transfer, data access and integration, resource brokering, monitoring, grid management services, operations centre,…
  - **NGS core-node Services**
    - CPU, (meta-) data storage, key software
  - **Services coordinated with others** (eg OMII, NeSC, LCG, EGEE):
    - Integration testing, compatibility & Validation Tests, User Management, training
- **Administration:**
  - Security
  - Policies and acceptable use conditions
  - Resource providers: service level agreements,…
  - Coordinate deployment and Operations

Summary

- NGS is a production service
  - Formal commitments to quality of service
  - GOSC runs the NGS and UK's Certificate Authority
  - Focus is on computation and data services

- NGS is evolving
  - OMII, EGEE, Globus Alliance all have m/w under assessment by the ETF
  - New sites and resources being added
  - Organisation establishing procedures (e.g. for VOs)
  - International context

Building e-Research



Researchers are not funded to provide production quality software for others to use

OMII-UK was created to bridge this gap!

Open Middleware Infrastructure Institute

To be a leading provider of reliable interoperable and open-source Grid middleware components services and tools to support advanced Grid enabled solutions in academia and industry.

- Formed University of Southampton (2004)
  - Focus on an easy to install e-Infrastructure solution
  - Utilise existing software & standards
- Expanded with new partners in 2006
  - OGSA-DAI team at Edinburgh
  - myGrid team at Manchester

Activity

- By providing a software repository of Grid components and tools from e-science projects
- By re-engineering software, hardening it and providing support for components sourced from the community
- By a managed programme to contract the development of “missing” software components necessary in grid middleware
- By providing an integrated grid middleware release of the sourced software components

The Managed Programme:

- Integrated with the Distribution
  - OGSA-DAI (Data Access service)
  - GridSAM (Job Submission & Monitoring service)
  - Grimoires (Registry service based on UDDI)
  - GeodiseLab (Matlab & Jython environments)
  - FINS (Notification services using WS-Eventing)
- Delivering into the repository
  - BPEL (Workflow service)
  - MANGO (Managing workflows with BPEL)
  - FIRMS (Reliable messaging)

Overview

- The UK e-science programme
- The National Grid Service
- GOSC – Grid Operations Support Centre
- OMII – Open Middleware Infrastructure Institute
- **JISC – Joint Information Service Committee**
  - **Services: NGS, Networking, Data Centres**
  - **Programmes**

JISC

Super JANET

JISC provide budget to UKERNA

JISC guide UKERNA through the JCN

UKERNA provide and manage JANET

Present incarnation is SJ4

SJ5 en route

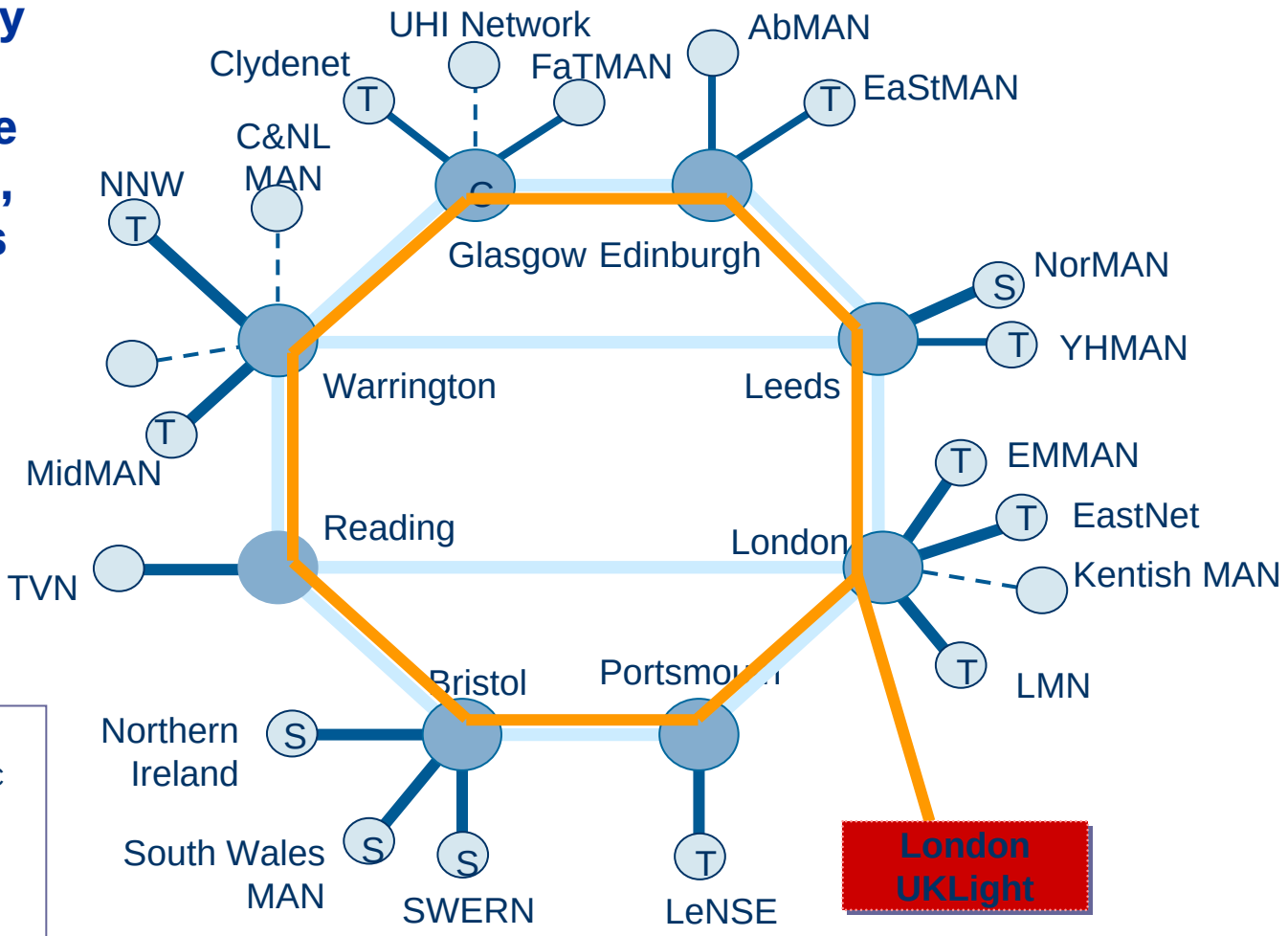
UKLight

- Funded by HEFCE and managed by UKERNA
- UK's first national switched circuit optical network
- Complements the SuperJanet4 production network
- National dark fibre facility for use by the photonics research community
- 10Gbit/s backbone to selected points in the UK
- Channels can be multiplexed, e.g. 4 x 2.5Gbit/s
- Connects to global optical networks via 10Gbit/s links to Chicago (StarLight) and Amsterdam (NetherLight)
- ESLEA is the first widely scoped project to exploit UKLight for a range of scientific applications

UKLight

National Connectivity

Potentially any institute can connect to the R&D network, and to access UKLIGHT through it

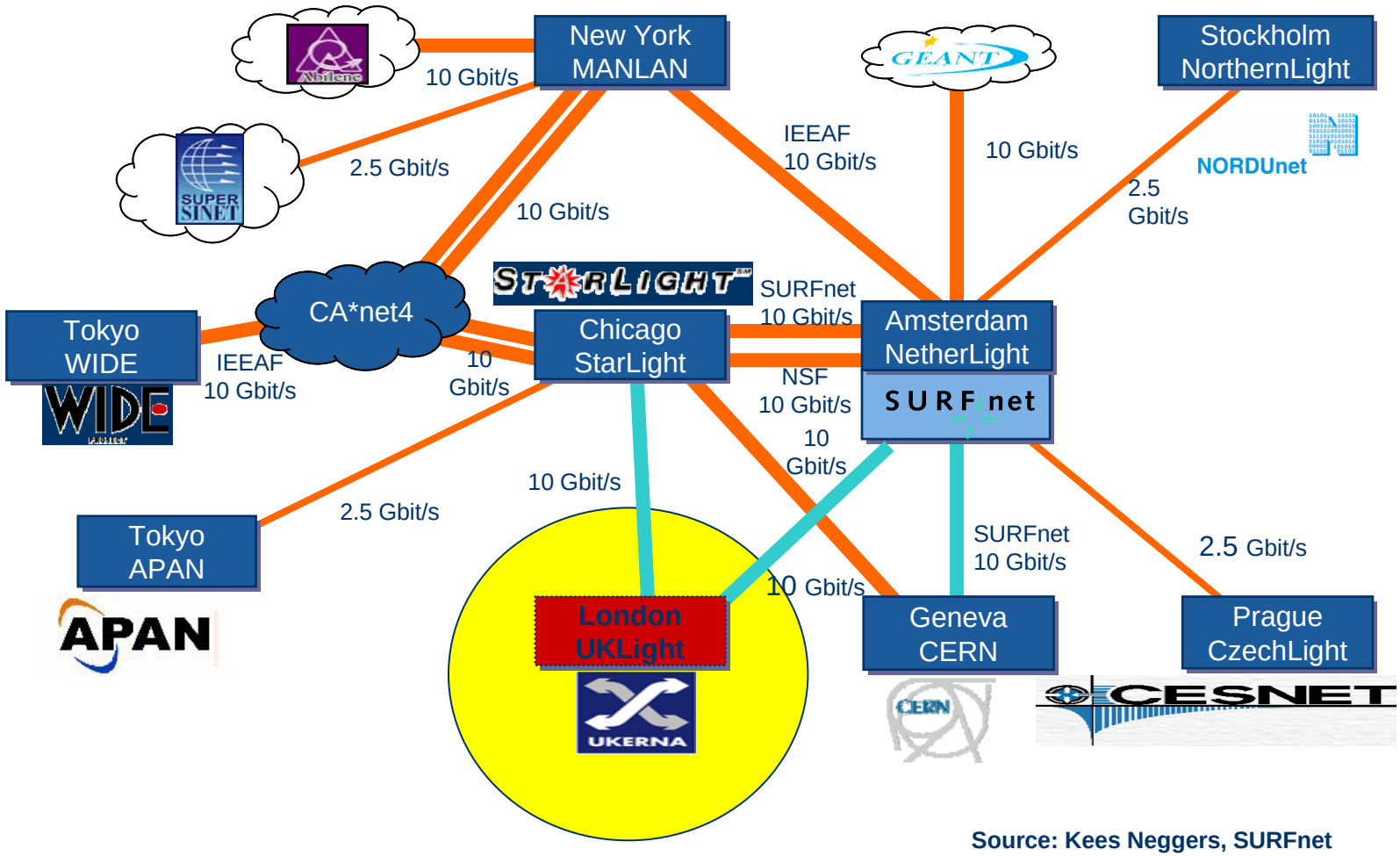


Key for components.

C-PoP	C
Extension to 2.5G BAR	S
Extension to 10G BAR	T

UKLight

International Connectivity



Source: Kees Neggers, SURFnet

What are (JISC) Data Centres?

- BIDS http://www.bids.ac.uk/
  - Bath Information and Data Services
  - Bibliographic
- EDINA http://www.edina.ac.uk/
  - Mulitmedia
  - Geospatial
- MIMAS http://www.mimas.ac.uk
  - Census
- What do they do?
  - Content delivery – with AA (currently

What are (JISC) Data Centres?

- BIDS http://www.bids.ac.uk/
  - Bath Information and Data Services
  - Bibliographic
- EDINA http://www.edina.ac.uk/
  - Mulitmedia
  - Geospatial
- MIMAS http://www.mimas.ac.
  - Census
- What do they do?
  - Content delivery – with AA (currently ATHENS)
  - Licensing authority

Challenge to grids!

How to manage IPR in workflow with licensed data

VOMS ? PERMIS??

JISC Programmes

- **Research is only the start !**

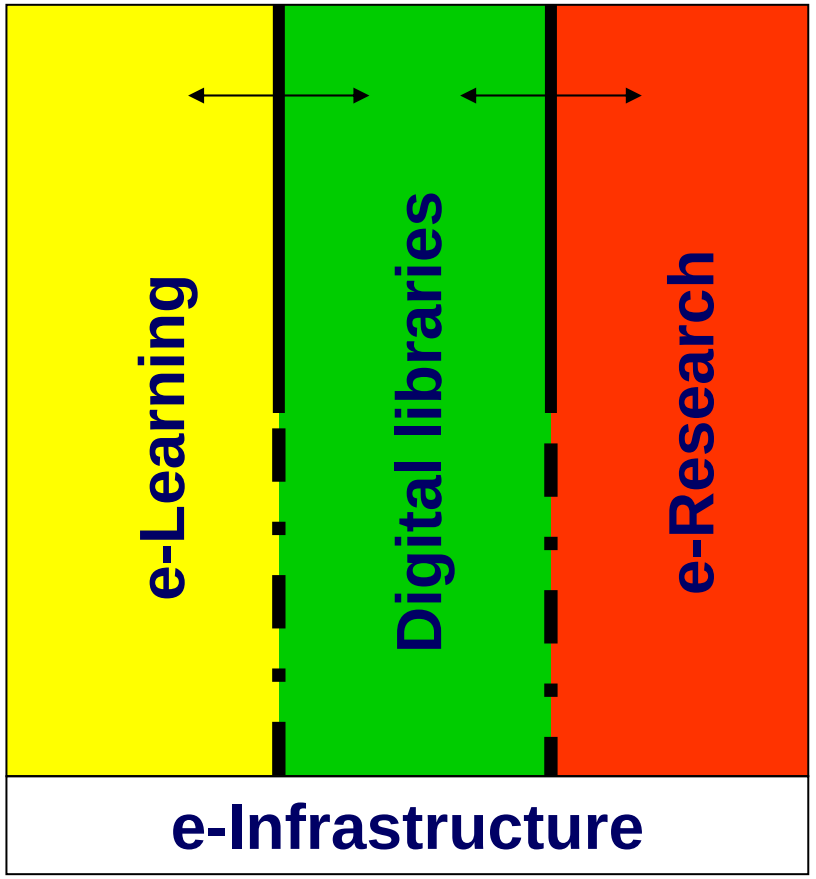


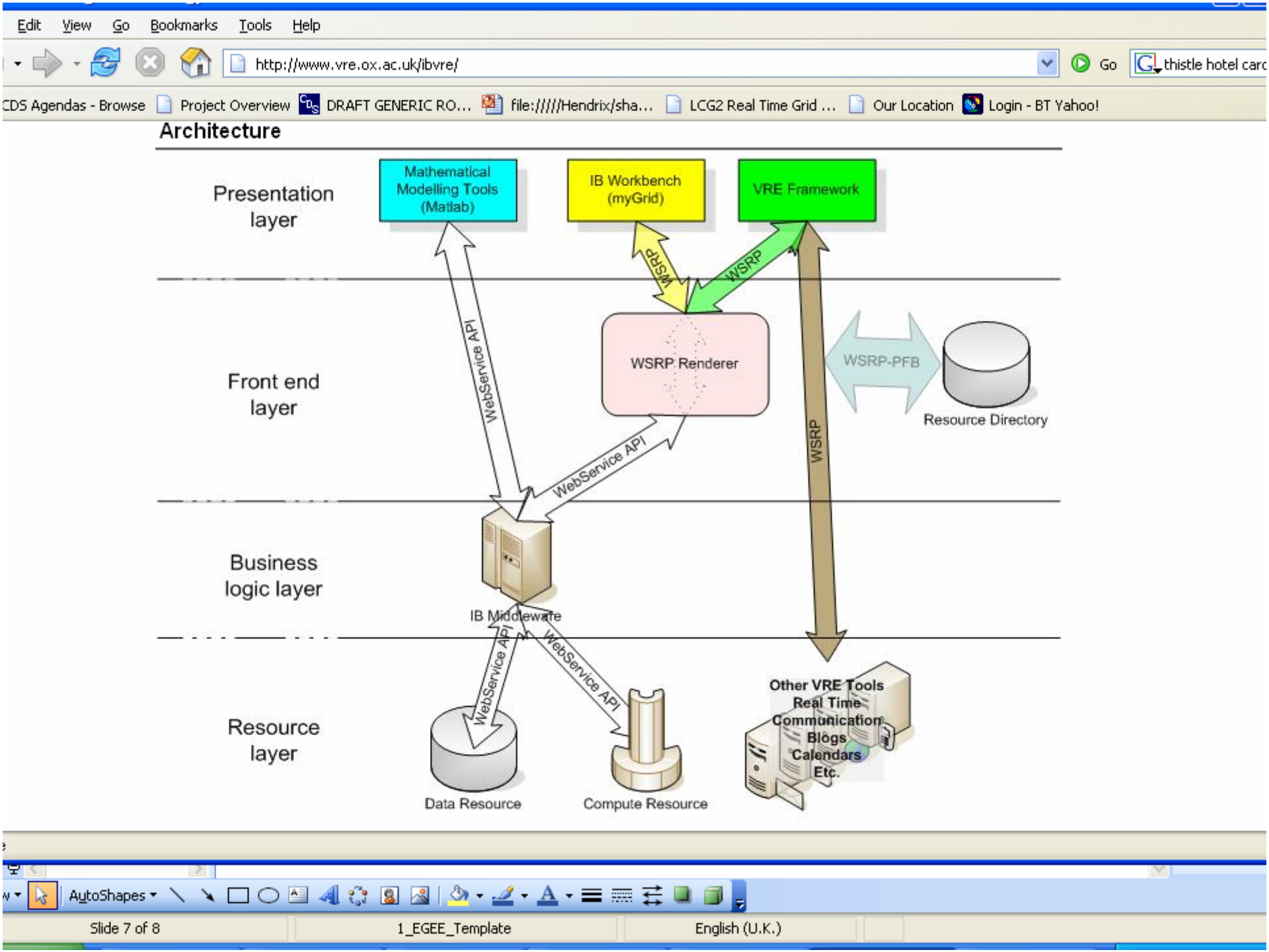
Figure derived from JISC

JISC Programmes

British Library/JISC Online Audio Usability Evaluation Workshop
Core Middleware Infrastructure
Core Middleware: Technology Development Programme
Digital Libraries in the Classroom Programme
Digital Prreservation and Records Management
Digital Repositories Programme
Digitisation Programme
Distributed e-Learning Strand
e-Learning Programme
e-Learning Tools Projects - Phase 2
Exchange for Learning (X4L) Phase 2
Exchange for Learning (X4L) Programme
Focus on Access to Institutional Resources (FAIR) Programme
JISC Framework Programme
JISC-SURF Partnering on Copyright
Network Development Programme
Portals Programme
Presentation Programme
Semantic Grid and Autonomic Computing Programme
Shared Services Programme
Supporting Digital Preservation and Asset Management in Institutions
Virtual Research Environments Programme

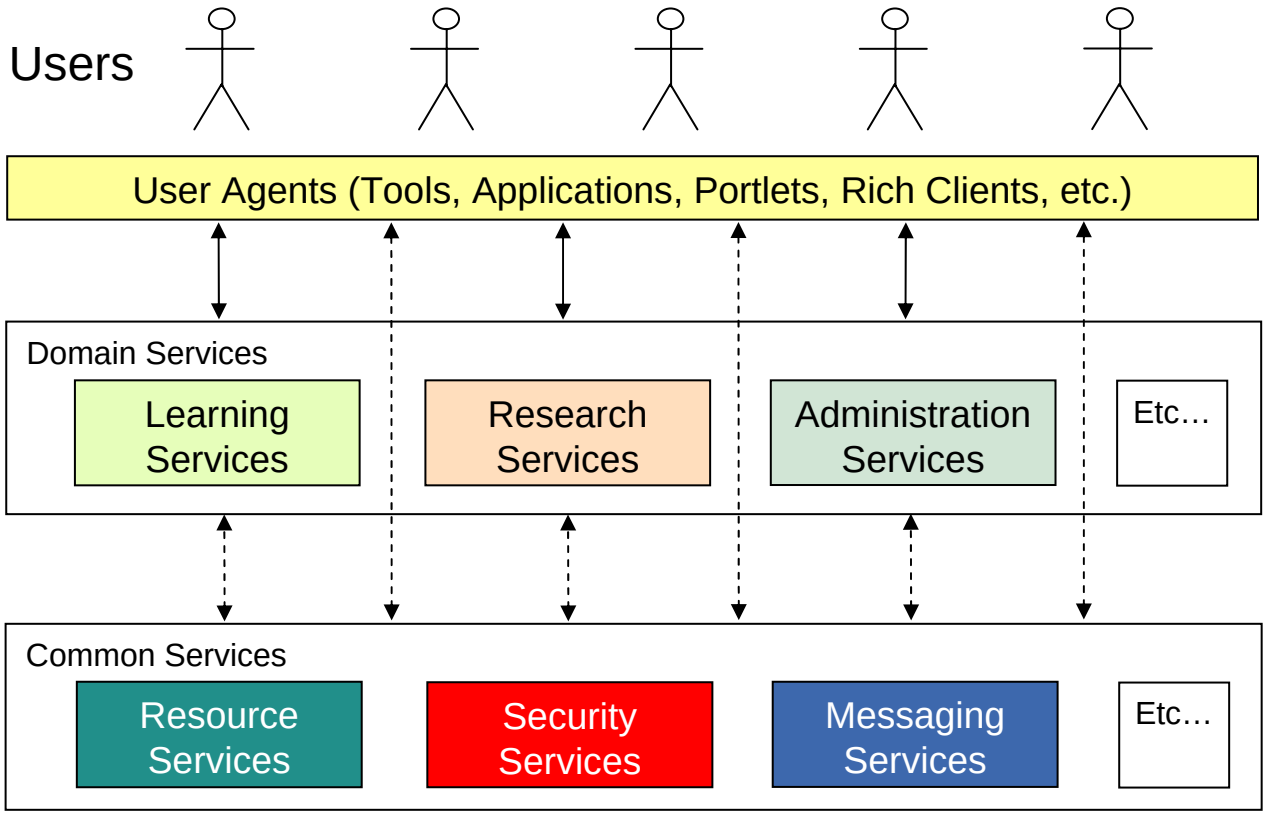
VRE development for Integrative Biology

- http://www.vre.ox.ac.uk/ibvre/
- “Whereas the existing IB Grid services have focussed on supporting the core IB experimental workflow - moving, processing and visualising data on the Grid - the IBVRE will support the research process in its widest sense i.e. activities such as identifying research areas and funding sources, building and managing projects/consortia, real-time communication, disseminating results, and provision of training to new researchers entering the field (learning and teaching support tools)”.


Edit View Go Bookmarks Tools Help
http://www.vre.ox.ac.uk/ibvre/
Go
thistle hotel carc
CDS Agendas - Browse
Project Overview
DRAFT GENERIC RO...
file:////Hendrix/sha...
LCG2 Real Time Grid ...
Our Location
Login - BT Yahoo!
Architecture
Presentation layer
Front end layer
Business logic layer
Resource layer
Mathematical Modelling Tools (Matlab)
IB Workbench (myGrid)
VRE Framework
WSRP
WSRP
WebService API
WSRP Renderer
WSRP-PFB
Resource Directory
WSRP
WebService API
IB Middleware
WebService API
WebService API
Other VRE Tools
Real Time Communication
Blogs
Calendars
Etc.
Data Resource
Compute Resource
AutoShapes
Slide 7 of 8
1_EGEE_Template
English (U.K.)

e-Frameworks Programme

- **Everyone (e-Learning, Research, Core middleware,…) are all developing services**



⇒ **The Frameworks programme is attempt to engender a coherent approach across all JISC programmes where possible**

UK e-Infrastructure providers

- The UK e-science programme
- The National Grid Service
- GOSC – Grid Operations Support Centre
- OMII – Open Middleware Infrastructure Institute
- JISC - Joint Information Service Committee

Web Sites

- NGS
  - http://www.ngs.ac.uk
  - To see what's happening: http://ganglia.ngs.rl.ac.uk/
- GOSC
  - http://www.grid-support.ac.uk
- OMII
  - http://www.omii.ac.uk/
- CSAR
  - http://www.csar.cfs.ac.uk
- HPCx
  - http://www.hpcx.ac.uk
- Grid Operations Support Centre http://www.grid-support.ac.uk
- National e-Science Centre http://www.nesc.ac.uk
  - UK Training events http://www.nesc.ac.uk/training

- JANET http://www.ja.net/