

“Shibbolized iRODS” (and why it matters)

3rd TERENA Storage Meeting, Dublin,
February 12th -13th 2009

David Corney, for Jens Jensen, e-Science
centre, Rutherford Appleton Lab, UK

Overview

Introduction

A shared e-infrastructure – current status

One area of development: ASPiS



About STFC...

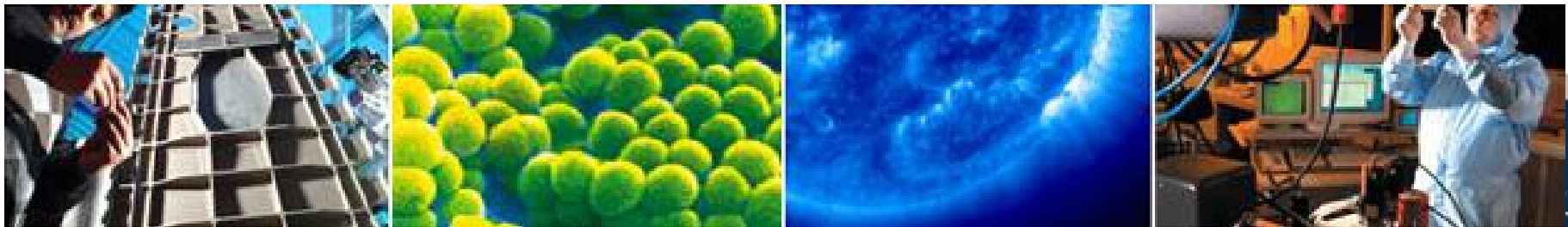
The Science and Technology Facilities Council (UK)

Created on April 1, 2007 (1 of 7 UK research Councils)

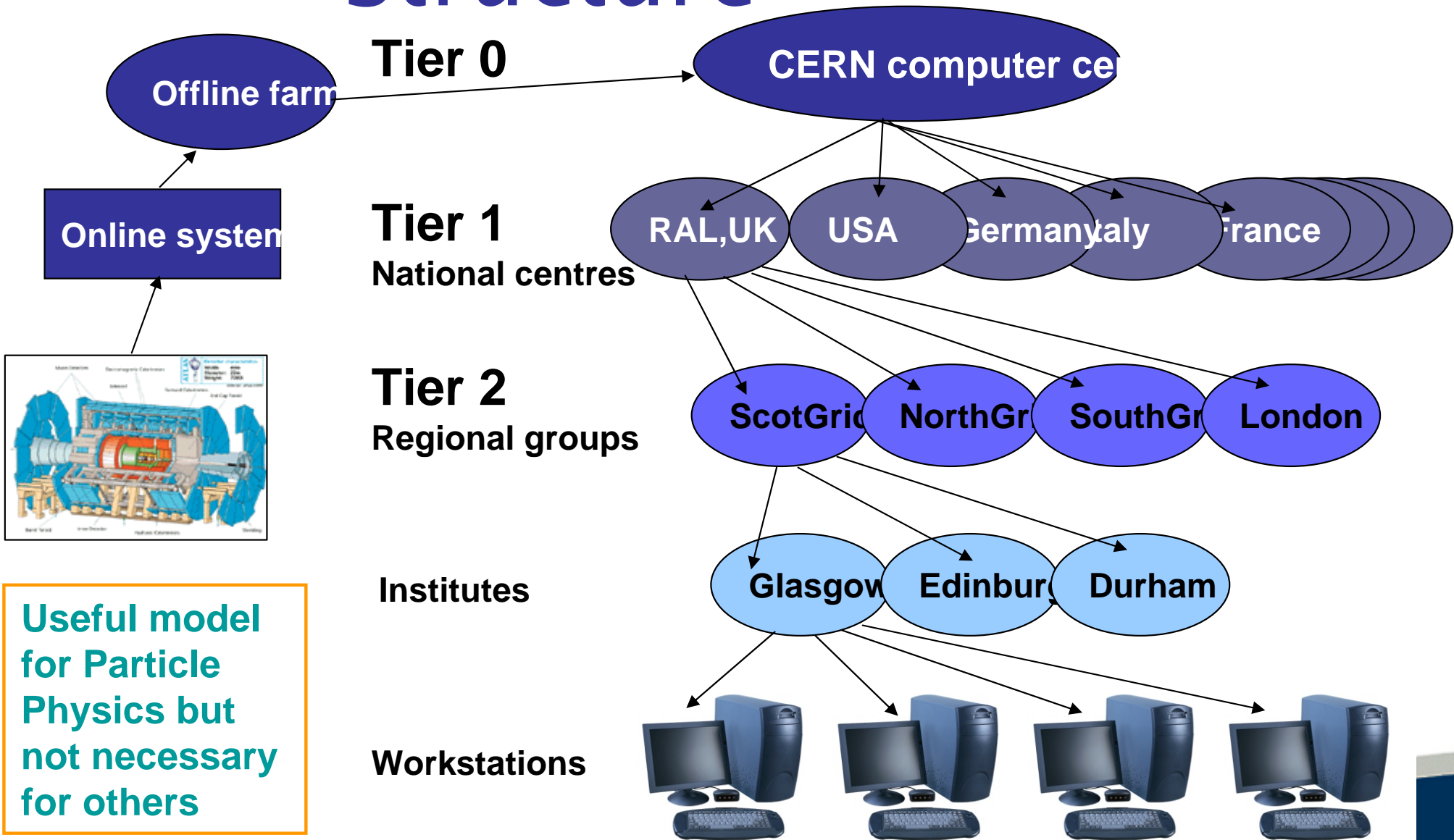
Responsible for:

- fundamental research in particle physics, nuclear physics, astronomy, space
- major UK facilities for the physical and life sciences
 - synchrotrons, light sources, lasers, neutrons
- national laboratories at RAL, Daresbury, UKATC
- international science projects
 - CERN, ESO, ESA, ILL, ESRF...

Over 2000 staff and an annual budget of over £700M



Tier Structure



Tier-1 Hardware

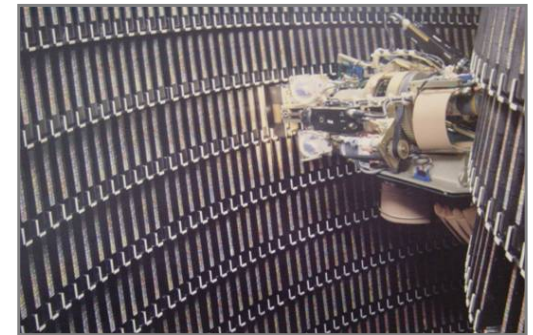
CPU Power (Reconstruction,
Simulation, User Analysis etc).
600 systems, 1250 cores, 1500
KSI2K



**Disk Storage
(Frequently Accessed)
138 Servers, 3200
drives, 750TB**

**'Tape' Storage – Long Term
retention – write once – read several
times a year – 1PB in SL8500 robot
+ 12 drives**

**Currently about 45 racks – with a further 25 due to arrive
for Xmass**



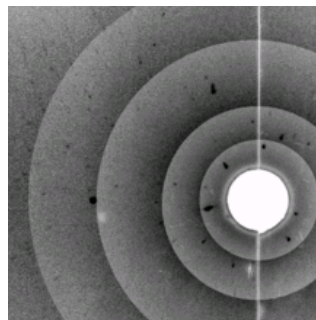
Rutherford Appleton Laboratory



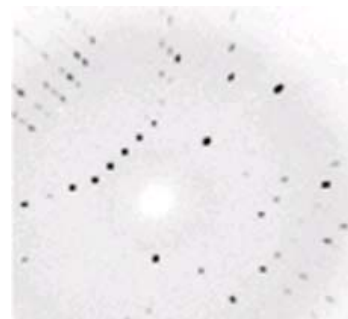
Science driver - enabling better science



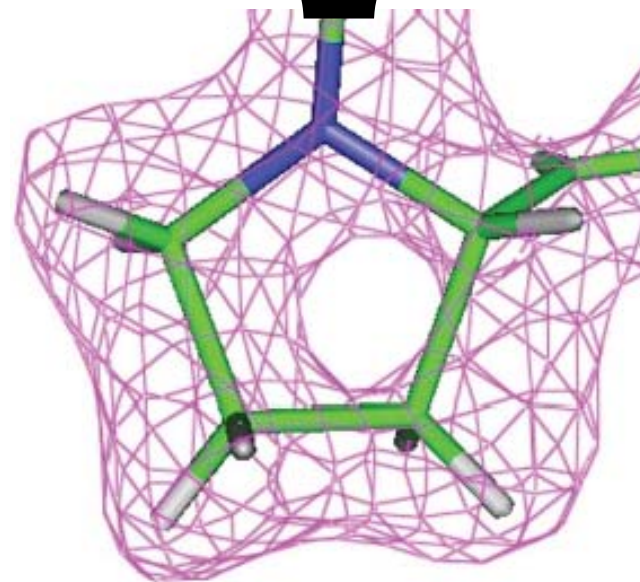
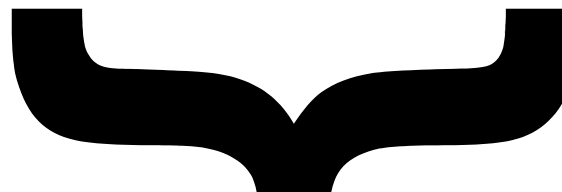
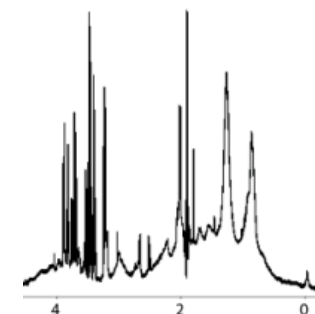
Neutron diffraction



X-ray diffraction

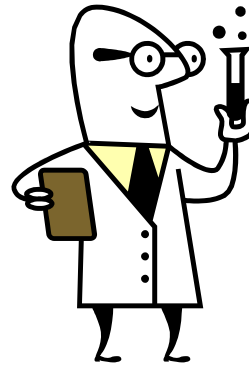


NMR



**High-quality
structure
refinement**

e-Infrastructure - Access to Multiple Facilities



SNS - ORNL



ANSTO - Australia



iCat



ISIS - TS1 + 2



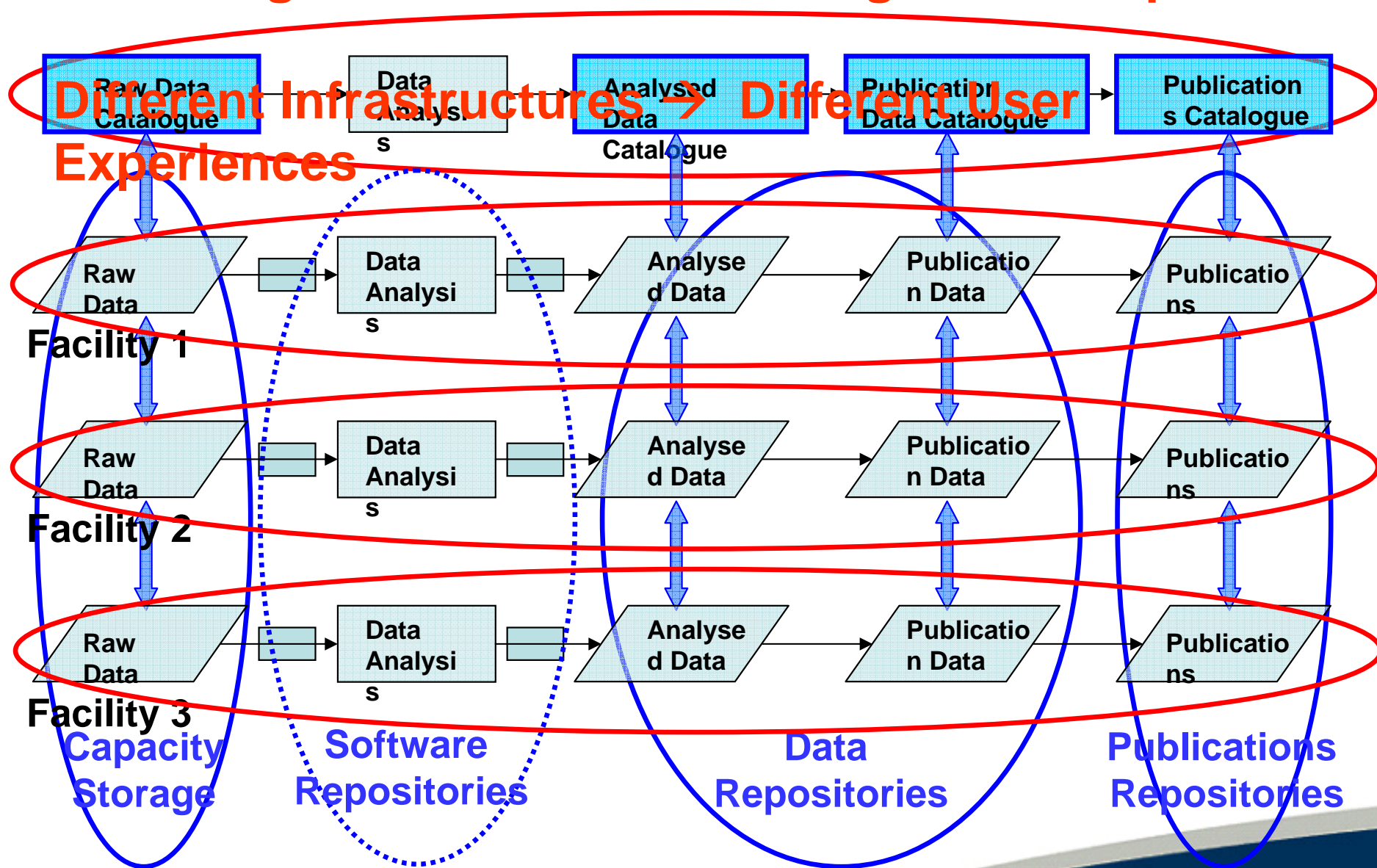
DLS



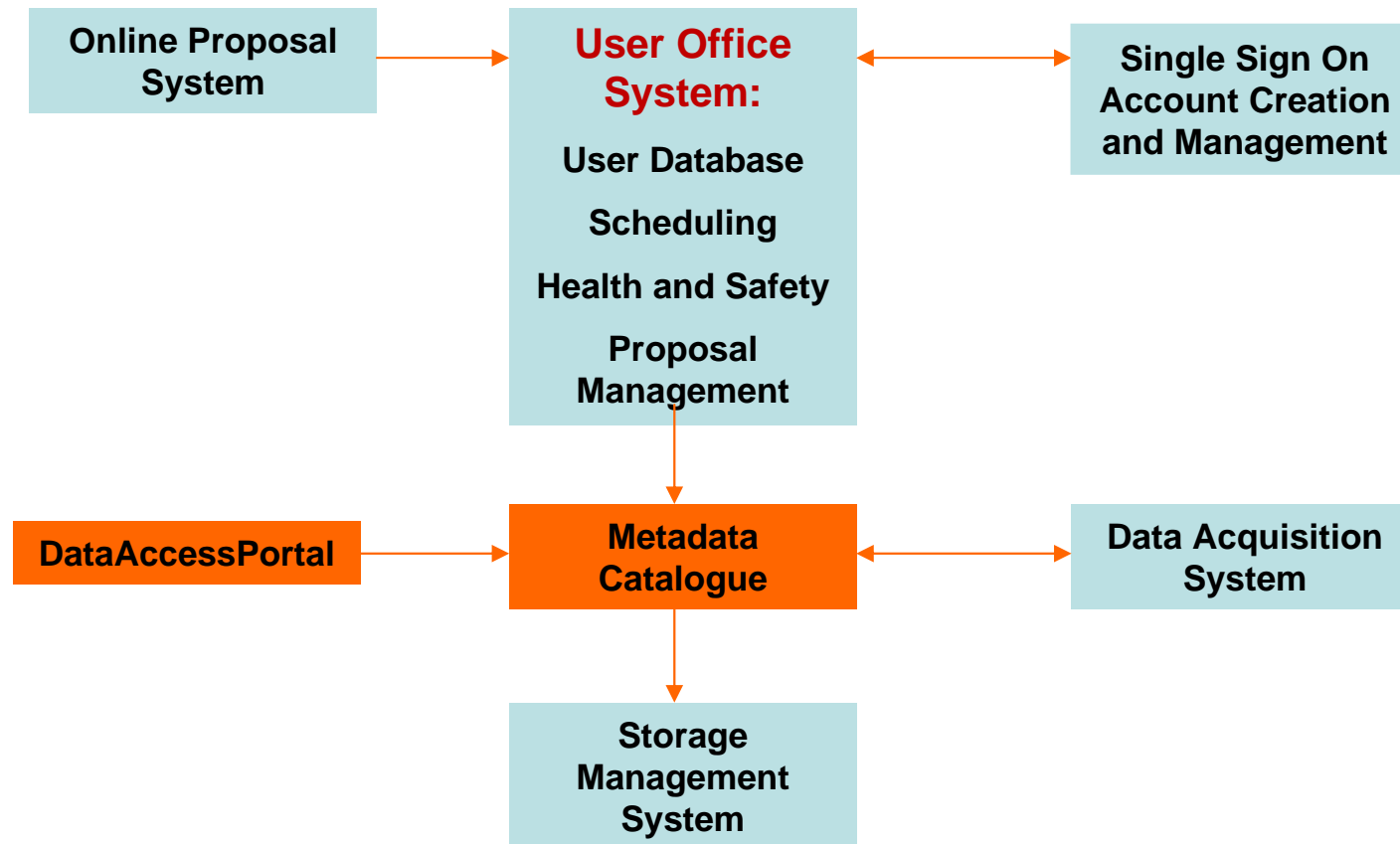
CLF

Technology Driver - integration and interoperation

Single Infrastructure → Single User Experience



Underlying Data Infrastructure



ICAT Software Suite, providing the crucial integration of key functions.

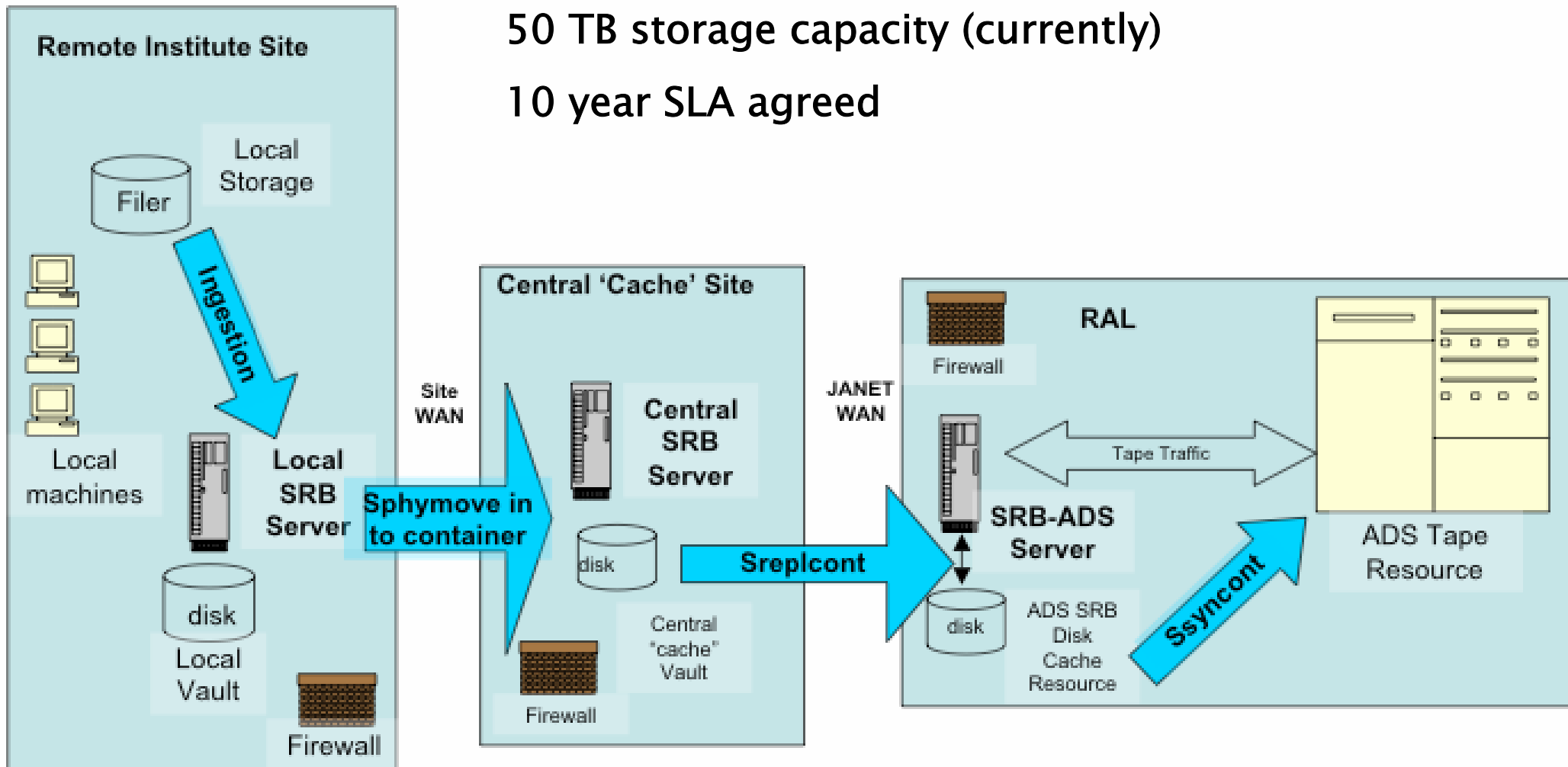
BBSRC Archive system

All (12) Institutes of the BBSRC

6000 scientists across the UK

50 TB storage capacity (currently)

10 year SLA agreed



Data Archive/Management Services

High Energy Physics Experiments (CMS, Atlas, LHcb, Alice, H1 ,...)

ISIS (Neutron Muon Source)

Diamond Light Source

British Atmospheric Data Centre

EISCAT (Radar research)

National Earth Observation Data Centre

BBSRC archive

Solar Physics World Data Centre

CICT (Standard IT backups)

Central Laser Facility

National Crystallography Service, University of Southampton

Hartley Library, Southampton University

WASP, VIRGO Consortium

SOLAR-B (Hinode)

Data Policy

- Data Policy (ISIS)
 - 3 year embargo on data (+1 if requested)
 - Commercial data is never made public
 - Instrument Scientists can access all data from their beamline
 - Calibration data is public
 - Any data that involves IPR (e.g. analysed) is private for perpetuity unless explicitly shared by user
- Automatic Enforcement of policy
 - A research area



EDNP



European Data Infrastructure for Neutron and Photon Sources



Combining European Neutron and Synchrotron Facilities

Already a common user community



Across many disciplines

- Materials, chemistry, proteomics, pharmaceuticals, nuclear physics, archaeology ...



The ASPiS project

Jens Jensen, STFC

via David Corney, STFC

Terena Storage TF, Dublin

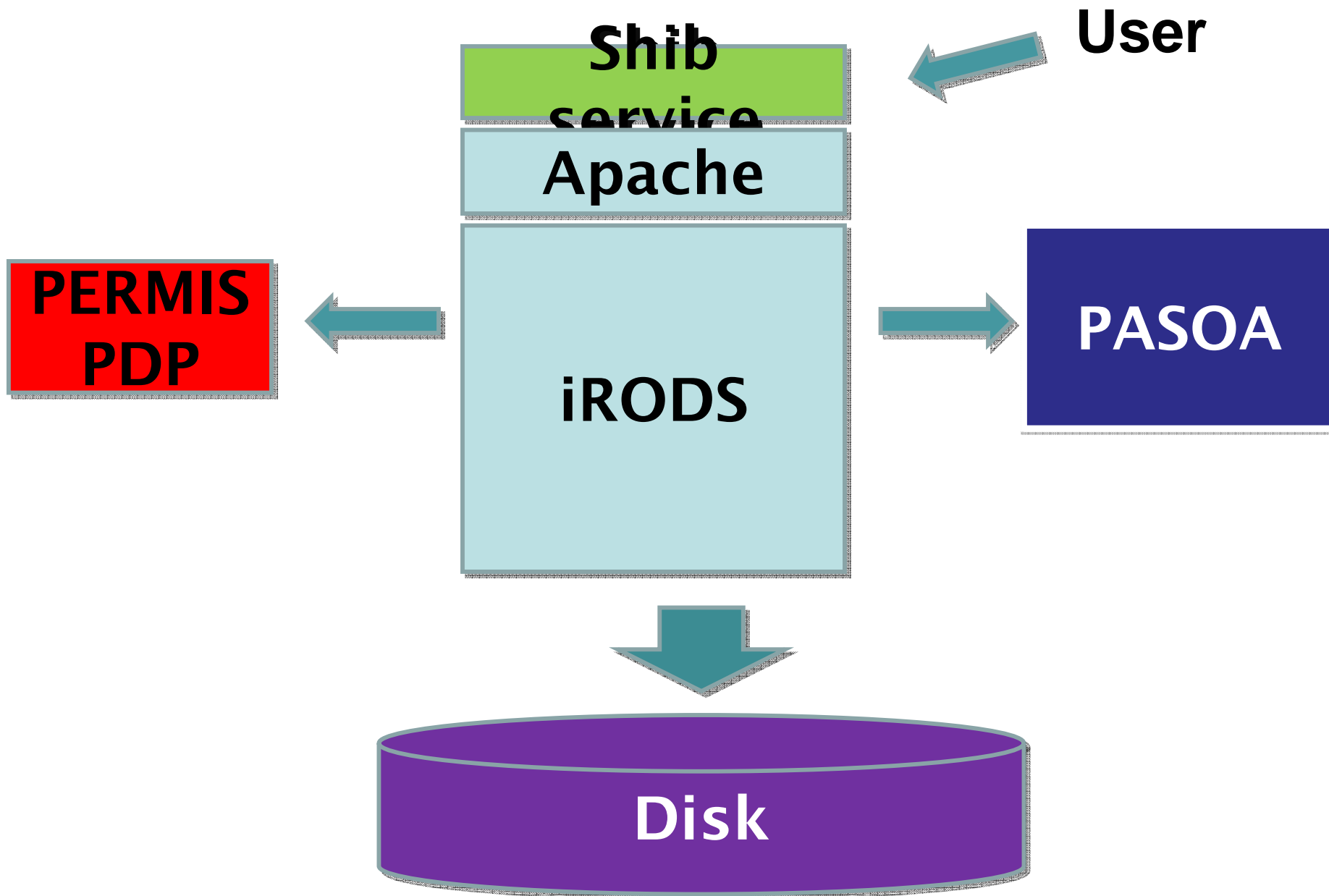
February 2009

ASPiS: people

- M Hedges, E Liao, T Blanke, CeRCH KCL
- A Weise, Reading
- A Hasan, Liverpool
- J Jensen, R Downing, STFC

ASPiS

- iRODS as datastore
- SSO login via Shibboleth
- PERMIS access control policy
- Provenance metadata in PASOA
- Funded by JISC

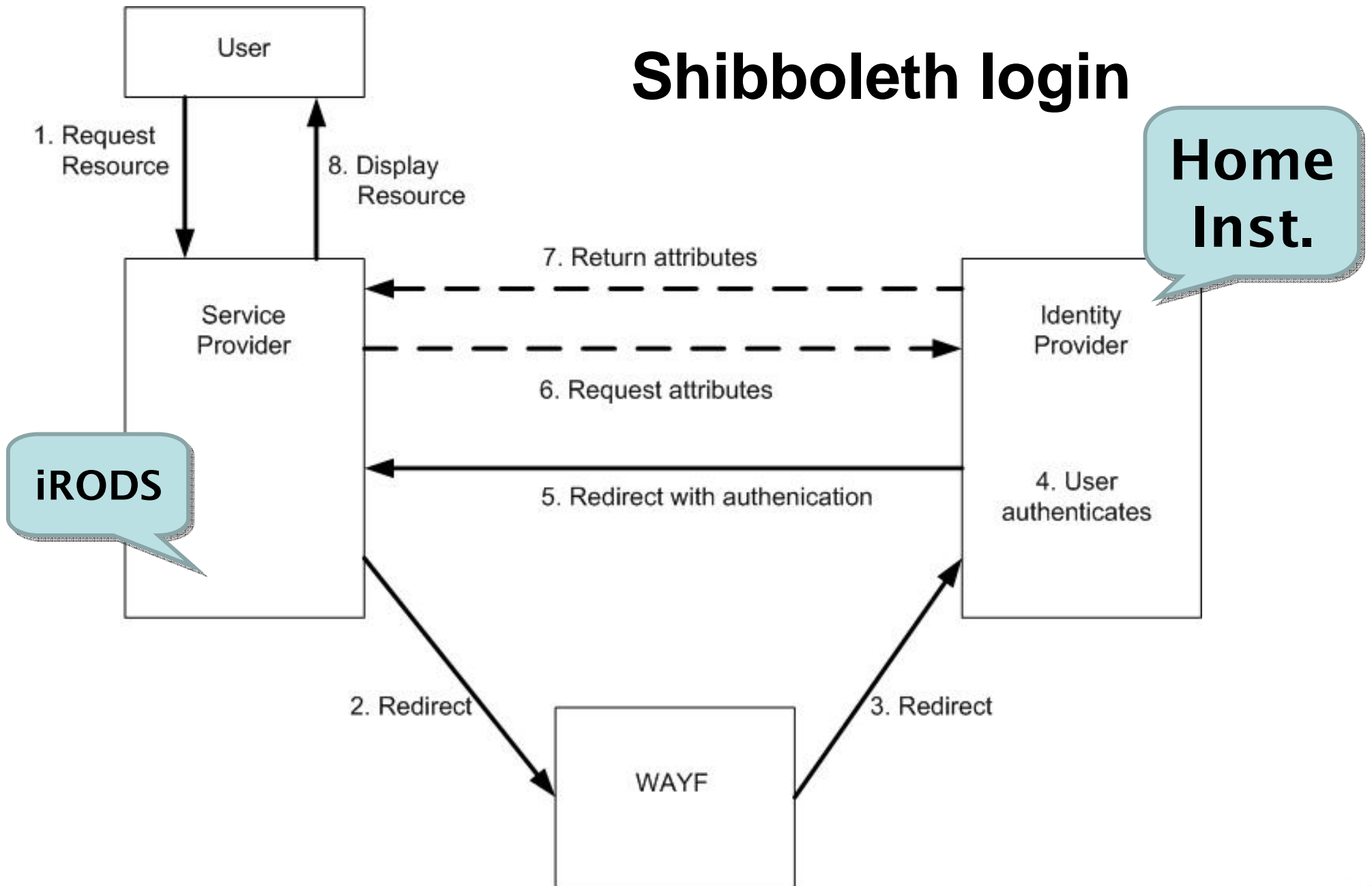


Shib login

So what does it do?

- Single password
- Password managed by home institution
 - S.E.P.
- Home institution provides attrs
 - ASPIs can use these for access control
 - And for provenance

Shibboleth login



iRODS

- Rule Engine to manage data workflow
- Microservices calling out to ext'l services
- No changes to iRODS itself
 - Improves maintenance



iRODS

Rule Engine

Log attrs

Access Ctrl

Update metadata

Branch on file type

Image metadata

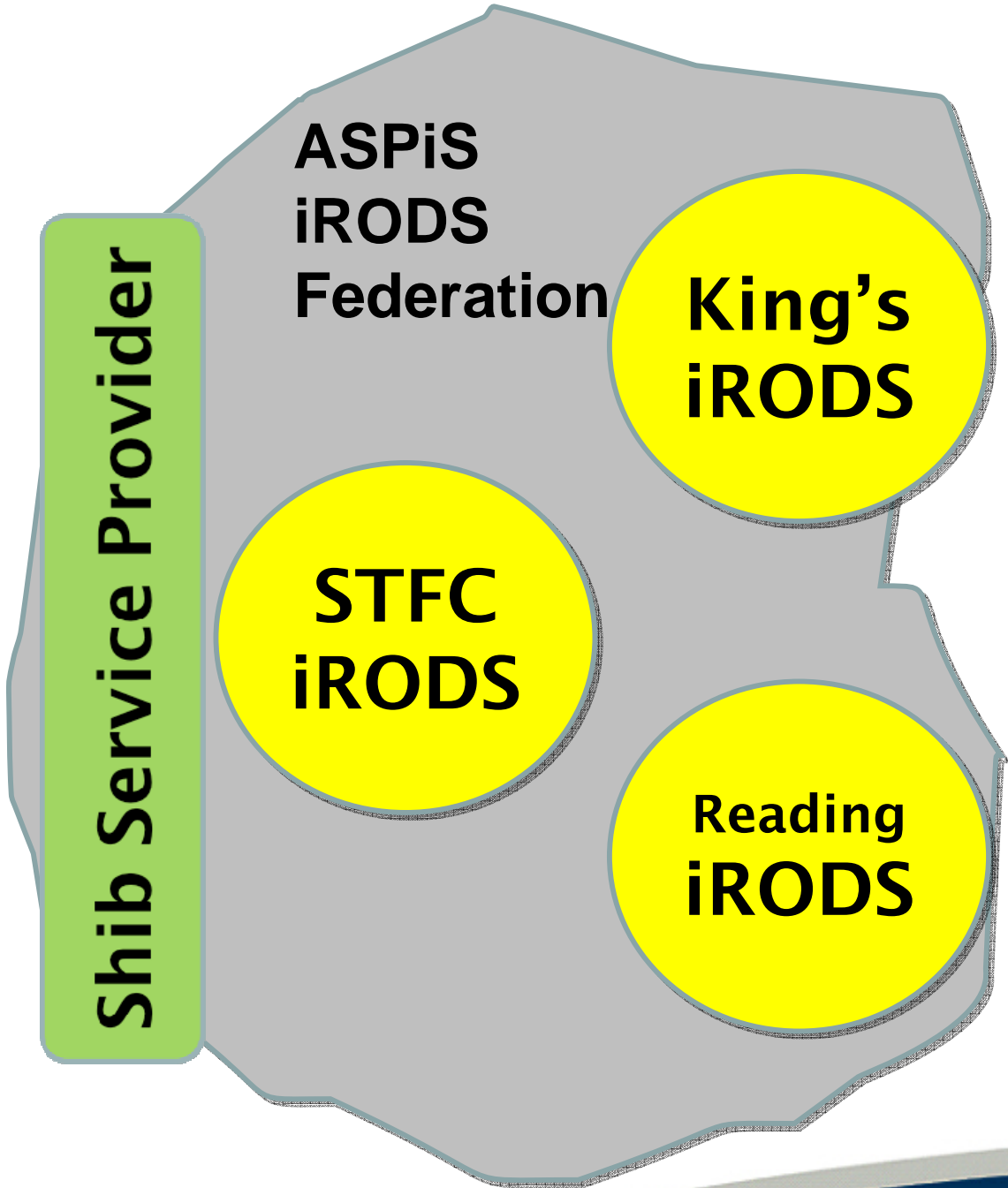
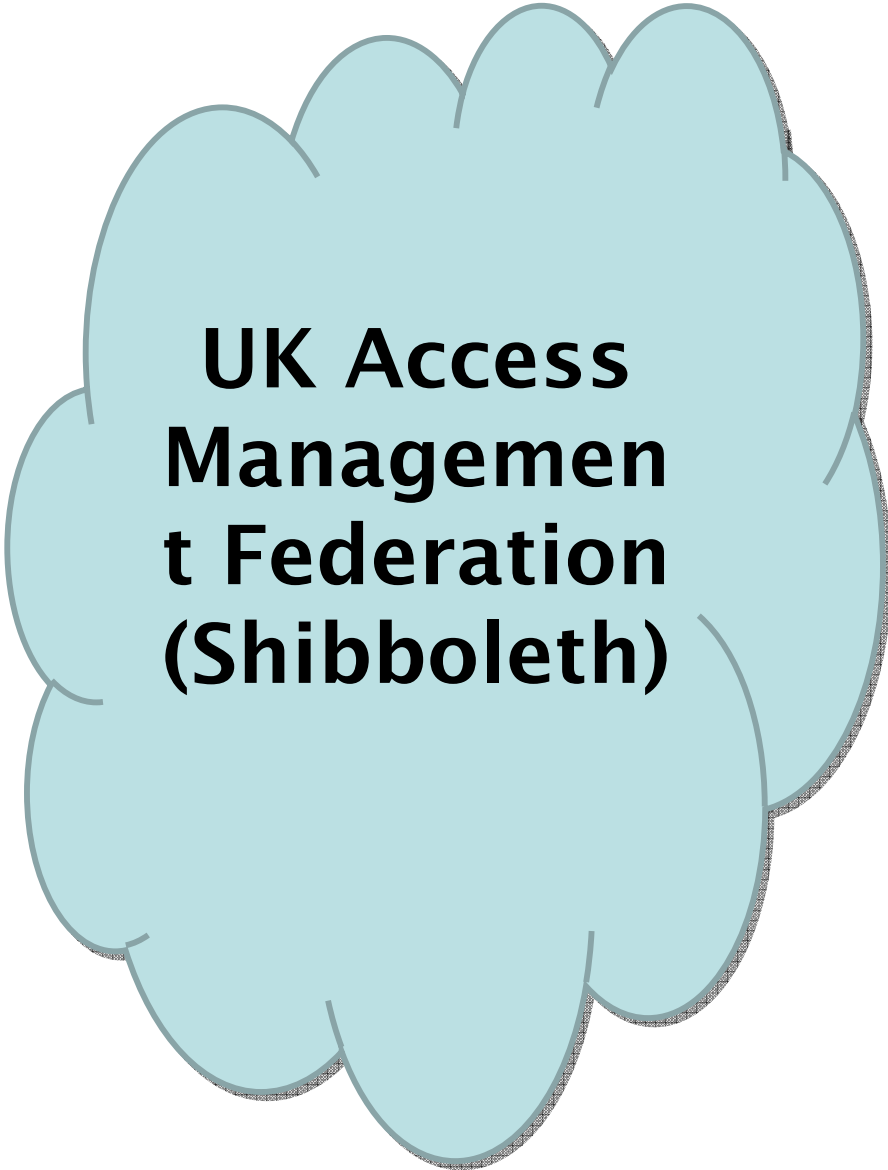
Document metadata

Example Rule workflow

PERMIS PDP

PASOA

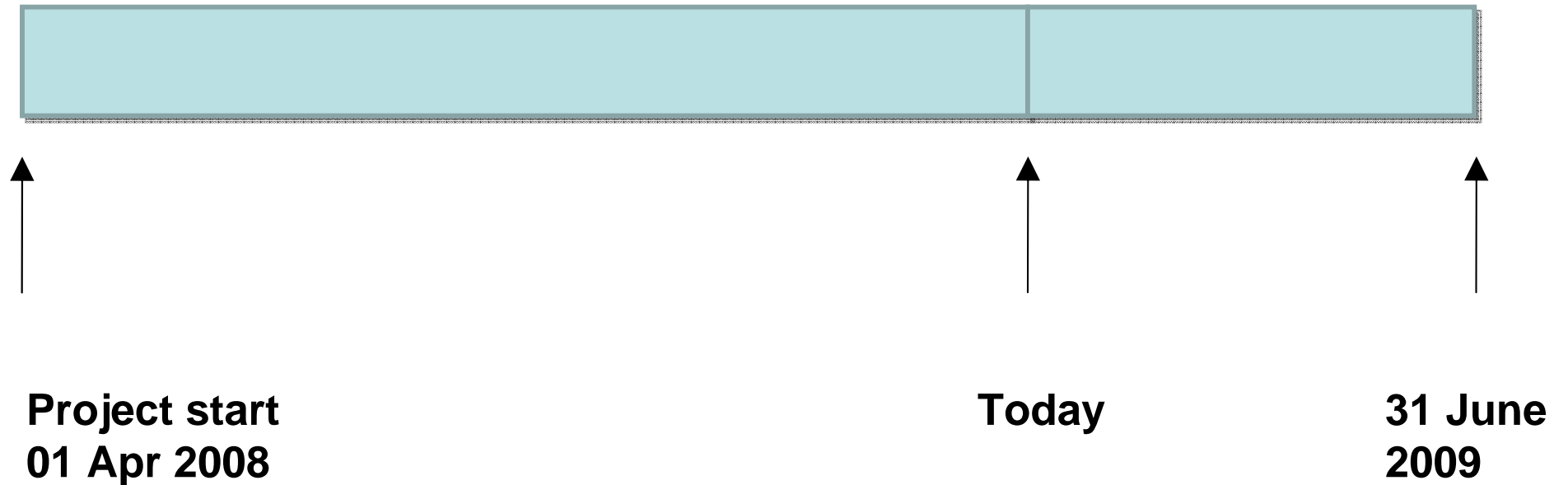
Two Federations



Target Users

1. Arts and Humanities
2. STFC facilities
 - Was Diamond Light Source (no IdP)
 - Now ISIS Neutron Source
3. SRB users on the National Grid Service

Timescale



Questions

Thanks for your attention

- and to David for giving the presentation

For questions, please contact

j dot jensen dot ral at gmail dot com

