

UKT0 Collaboration

WP2: Software Infrastructure

Frazer Barnsley
Software Engineering Group, SCD

WP2: Software Infrastructure

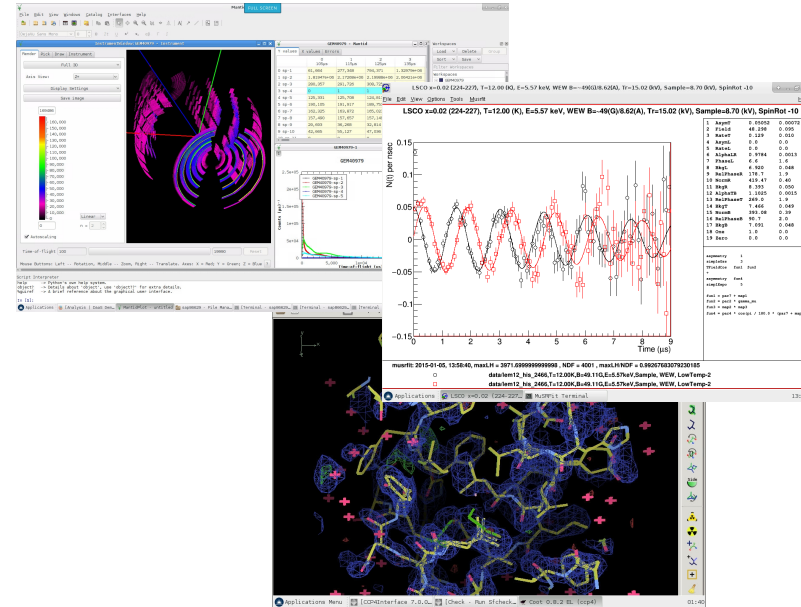
- Two software components for DAaaS
 - Virtual Machine Manager (VMM)
 - Data Movement Service (DMS)
- December → March (4 months)
- Developer effort from Tessella

DAaaS

- Data Analysis as a Service
- Improving the scientific output of the RAL facilities
 - ISIS Neutron and Muon Source
 - Central Laser Facility
 - Diamond Light Source
- Three main components
 - Software
 - Data
 - Compute
- Simplification and automation

Analysis Environments

- Remote desktop to VMs
- Customised to specific types of analysis
- Ready to go instantly
- Pre / in / post experiment analysis



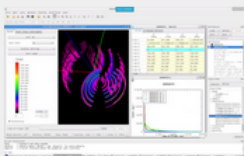


Create Analysis Environments



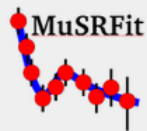
Excitations

This environment has been tailored for the analysis of in-elastic neutron scattering data from the ISIS Excitations Group. It comes complete with Mantid, Matlab, Horace and MSlice as well as access to the ISIS Experiment Archive.



Excitations Large

This environment has been tailored for the analysis of in-elastic neutron scattering data from the ISIS Excitations Group. It comes complete with Mantid, Matlab, Horace and MSlice as well as access to the ISIS Experiment Archive. This environment has been supplied with extra resources: 32GB RAM, 8 cores.



Muon

This environment has been setup to test the MuSRFit software.

Create

Close

Create

Analysis Environments



Muon! ✎

Host: vm560.nubes.stfc.ac.uk
Date created: 2017-08-22
Owner: Mr Frazer Barnsley

Share Launch ▾



Excitations ✎

Host: vm568.nubes.stfc.ac.uk
Date created: 2017-10-13
Owner: Mr Frazer Barnsley

Share Launch ▾



Home



Experiment
Archive



RB
Directories



Data

Software

Examples

Utilities

Run Program...

Applications

Matlab

MantidPlot

MantidPlot Nightly

Horace

Horace Planner

MSlice

JournalViewer

Virtual Machine Manager

Motivations

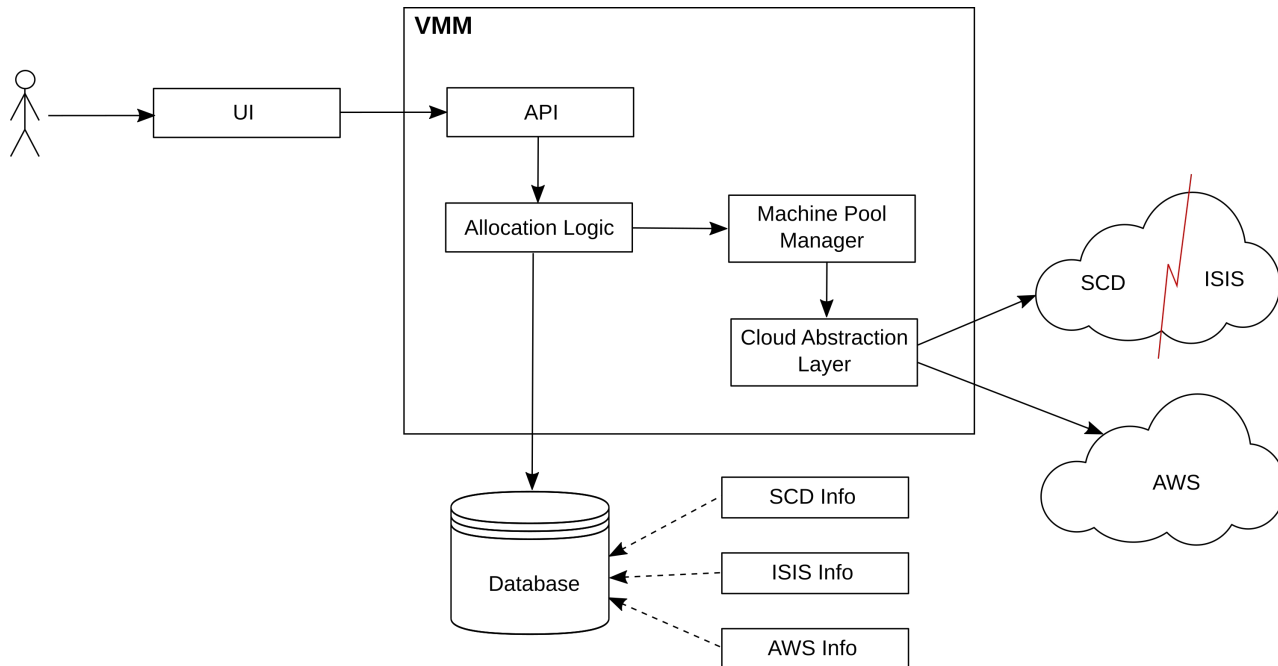
- Currently analysis environments only accessible through the TopCAT frontend
 - Open it up to other frontends
- Make use of resources at universities, other institutions, commercial clouds
- Bring 'compute' to the user / data

Requirements

- Provide API that allows clients to acquire different type analysis environments
- Cloud agnostic
 - Openstack, Azure ...
- Allow selection based on:
 - Location
 - Resources (CPU, RAM)
 - GPUs
- Fast - analysis environments available in seconds

Design

- 4 main components
 - API
 - Pool manager
 - Cloud abstraction layer
 - Allocation logic

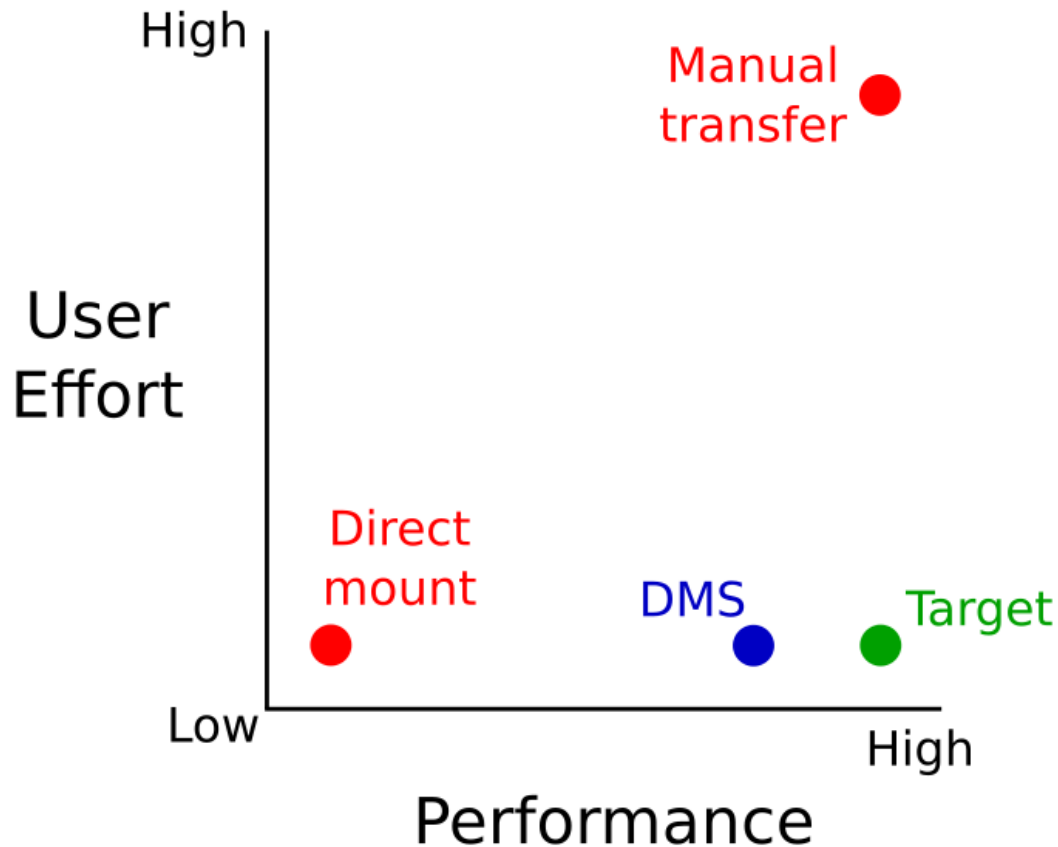




Data Movement System

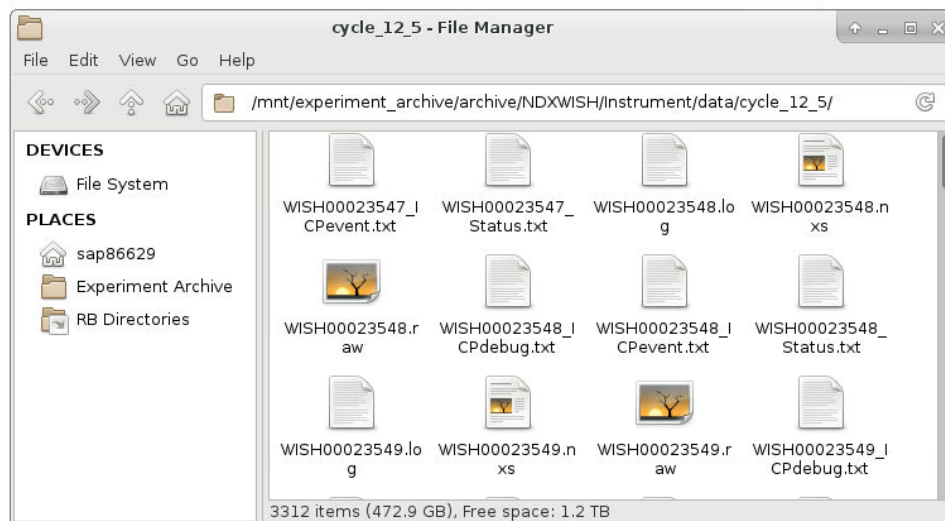
Motivations

- Simplify data access for users



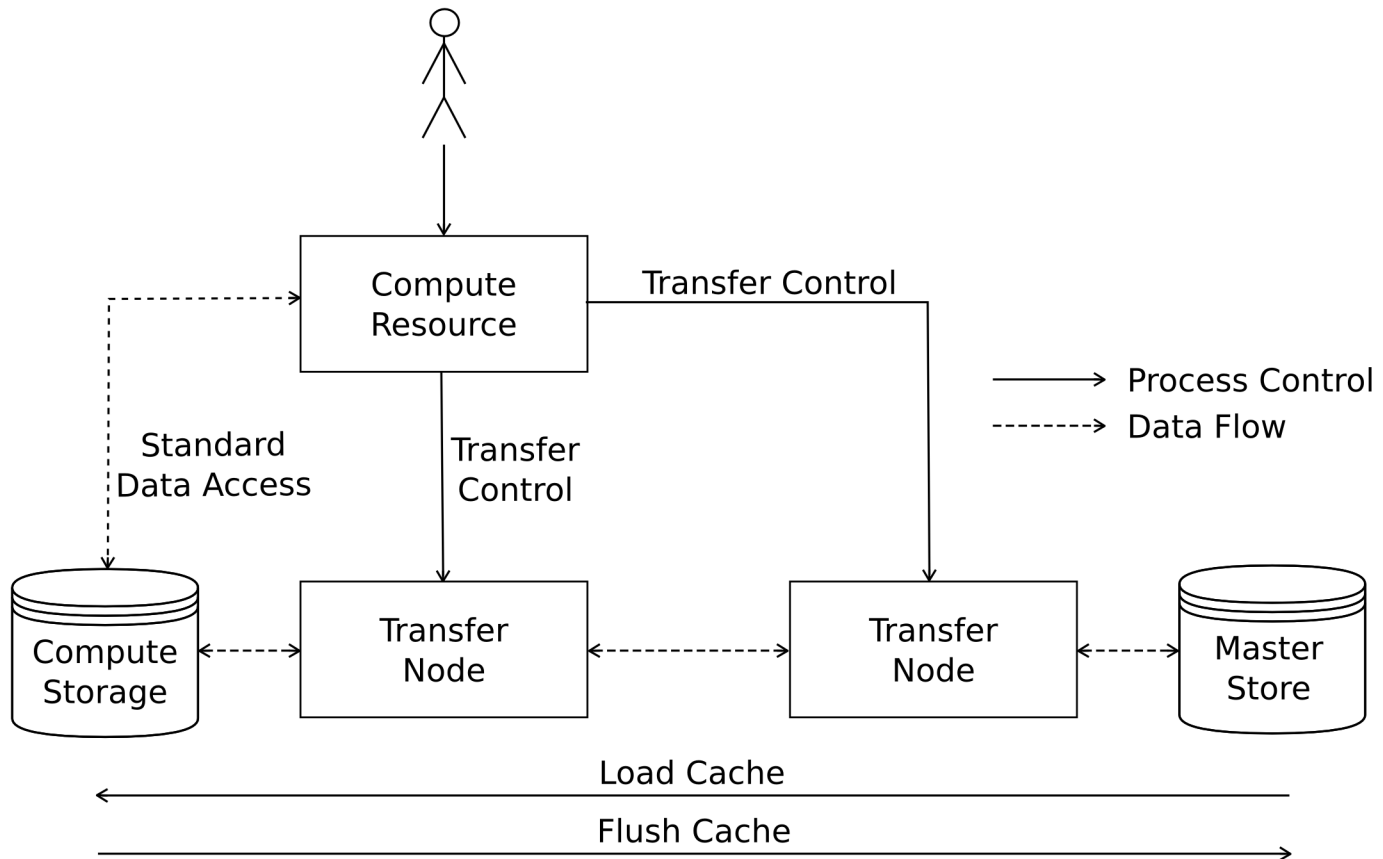
Requirements

- Filesystem view of data for the user
- Passive - independence of storage resources
- Scalable for large data volumes
- Transfers over long distances



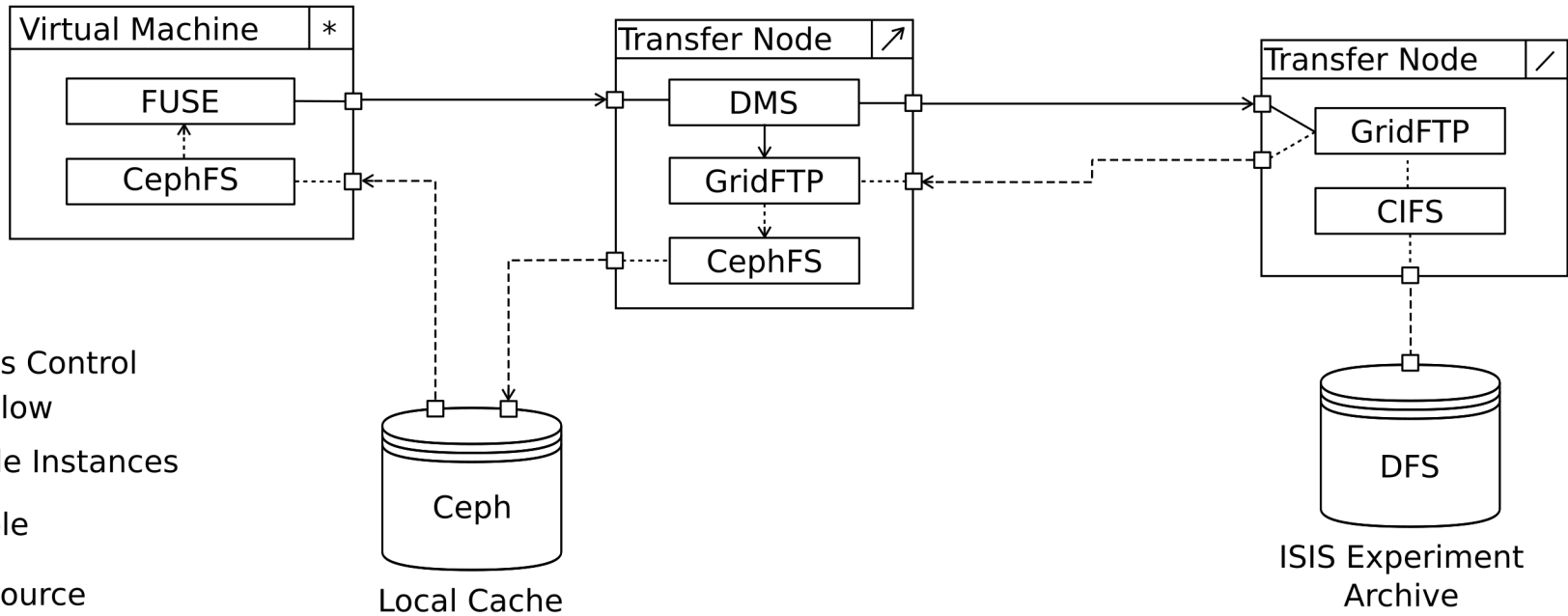
Design

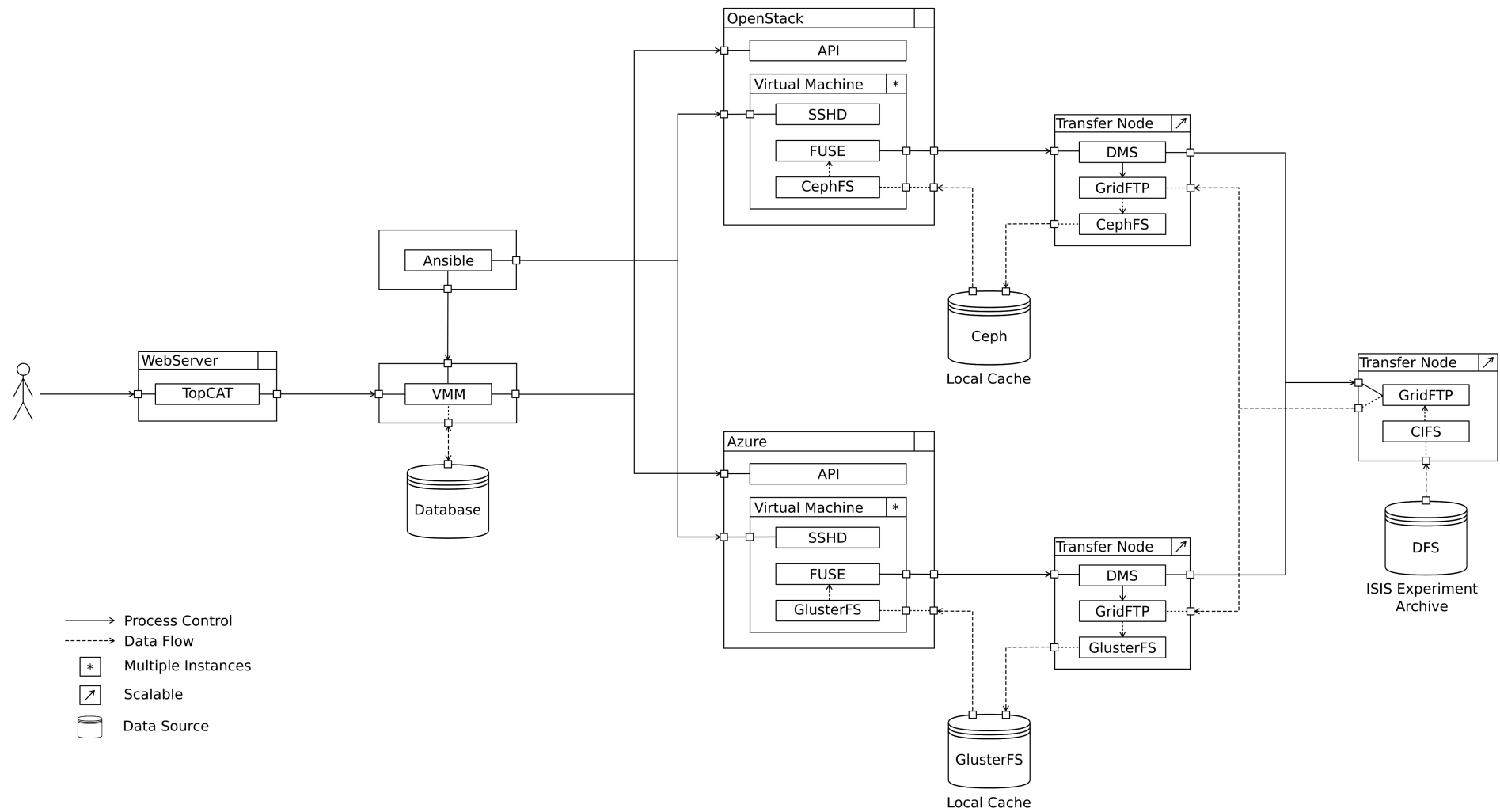
- On demand data caching



Design

- 2 main parts
 - FUSE client
 - DMS service
- Plus GridFTP machines





Achievements

- VMM
 - API, cloud abstraction layer, pool manager complete
 - On-going development on allocation logic
 - Basic functional tests run against Openstack
- DMS
 - Can handle ISIS experiment archive (read-only, shared data)
 - Private user data almost complete (home dirs)
 - To-do, shared private data (ACL rules)
 - Performance tuning
- On-going development to integrate with DAaaS

frazer.barnsley@stfc.ac.uk



Science & Technology
Facilities Council