National Federated Compute Services NetworkPlus

A Network+ of Federated Compute Services

NFCS+: Federated data movement and access

Alastair Basden, George Beckett, Shaun de-Witt, Manu Antony, Jonathan Hollocombe, Kieran Leach, Mark Lovell, Chris Mountford, James Perry, John Taylor, Mark Wilkinson

Durham University

COSMA

DiRAC

epcc

UK Atomic Energy Authority

UKRI

Reporting on two NFCS+ projects around large scale data

Central premise:
User-driven data movement and sharing at scale is not a solved problem!



Federated data movement

- Enabling researchers to easily and predictably transfer their data
- Removing access barriers: user-driven
  - No intervention from “tech-support” required
  - Autonomy and control

Federated data access

- Zero-copy data sharing
  - Without copy
  - Without staging
  - Making datasets available to collaborators
- User-driven
  - Federated across the DRI

NFCS+

- Seeking to make recommendations to UKRI
  - Focused on federation across the DRI
  - 12-18 month projects
  - Prototypes, PoCs and technical demonstrators
    - Production facilities not expected
  - Landscape surveys, questionnaires
  - Workshops, tutorials
- 20+ projects in total

Outcomes

- A final report from each project
  - With recommendations
- A final overall report summarising these to UKRI
- To enable direction of future funding

Our approach

- Study of different transfer tools: rsync, Globus, S3, Rclone, RUCIO
  - Performance benchmarking between multiple sites
  - Setting up of endpoints
- Study of sharing tools
  - S3, Versitygw, s3fs, hdfstream
- Questionnaire sent to large projects
- Landscape survey of national and international projects and capabilities
- Identifying and removing obstacles
- Documentation and guidance – including online live project updates

UK user survey summary (abridged)

- Dataset sizes are >100TB
- File sizes typically 10GB-1TB
- Both data to user and user to data models
- Transfers can be continuous or discrete files
- Data can be cut out of larger files
- RUCIO has been adopted into some workflows
- Rsync is achievable for 100TB scales (at night) – though not recommended
- Some access restrictions may be required
- Mixed user community expertise
- There are some powerful community-specific tools – e.g. HDF5 streaming over http, with a Python client.

The recommendations

- 9 for data movement
- 7 for data sharing
- MUST/SHOULD/COULD

DRI sites MUST adopt a common practice for user-driven and user-led data movement

- Current situation is too site-dependent
- Globus seems like a good option here for general purpose
  - Gaining a critical mass (helped by this project)
  - Free – or with paid support and features
  - Transfer (the free component) as a minimum
- A centralised object storage (e.g. StorJ) is highly desirable

A monitoring infrastructure MUST be created which should link all DRI sites, to measure and track likely data-transfer performance and topical availability.

- So that automation is a given
- To enable up to date transfer rate information
- To facilitate tracking of problems
- Similar to the JISC PerfSonar

A regular (for example, every six months) meeting of UK-involved data-transfer research activities, plus infrastructure providers and other stakeholders SHOULD be organised.

- A data-transfer community
  - So that it is not forgotten
- Can be coupled with other events (CIUK, etc)

An application process for access to DRI SHOULD be created that allows applicants to request resources at multiple sites in a single step, as is done for the DiRAC Resource Allocation Committee or IRIS Resource Scrutiny and Application Panel. These kinds of experiment are what has motivated the Federated Data Movement project

- Avoiding the problem where only parts of a complex workflow get given a time allocation
  - The ability to move data is inherent

Provision SHOULD be made to allow UK-based research teams to contribute data to or exploit data from international experiments or activities.

- e.g. making use of international repositories
- Increase UK visibility
- Reduce barrier to exploitation of free datasets

A multi-site Data Transfer testbed SHOULD be used to test new transfer technologies and approaches, to develop good practice for the community, and to provide a space in which research groups can experiment with data movement at a realistic scale.

- New stuff without breaking production
- To encourage development

User training in data management technologies and use MUST be provided, SHOULD should include consideration of sustainability issues and how responsible data management and transfer can reduce environmental impact.

- Good documentation
- Regular workshops
  - Including community feedback
- Basic and advanced levels
- Good practice in data curation and management
- Providing information to users

DRI sponsors SHOULD aim to locate compute resources close to significant data acquisition and data generation sites, so as to eliminate unnecessary data transfer tasks, and improve time to solution for research campaigns.

- A single compute facility is not a good solution
  - Bring the compute to the data

A DRI-wide catalogue of services MUST be created and SHOULD include details of each service's (or host institution's) support and technologies for data ingress/ egress.

- Seems obvious: have a central UK source of information for researchers
  - How do I get my data from A to XXX

Requirements for data sharing

- Key concept:
  - I have a large dataset
  - Someone needs to access it
  - No one has space to copy or stage it
    - It has to remain where it is
    - Shared for partial access
- Zero-copy
- Partial file access
- Remote mounting
- One-off operations are a good use case

DRI facilities MUST provide read access data sharing mechanisms (e.g. S3 or web access)

- This would simplify data sharing for many users
- Technically not difficult
  - At least as a minimum to get started

Options for open (no credentials) and protected (credentials) sharing MUST be provided

- Open access is to be encouraged
  - But some data may have embargo periods
  - Or be sensitive in some way

Data access mechanisms SHOULD facilitate POSIX file system mounting of shared datasets (e.g. s3fs)

- To ease existing workflows
- To simplify life for researchers
- To reduce data access barriers

Data sharing mechanisms SHOULD be user driven and self-service without requiring intervention from technical support staff

- To allow it to happen in a timely manner
- Reducing the burden on sys admins

Partial data transfer access SHOULD be facilitated (e.g. it should be possible to transfer only a subset of a file transparently to the user)

- e.g. accessing part of a dataset within a file
  - Accessing file metadata
  - Selecting a particular object (e.g. a galaxy)
  - Slicing particular indices based on metadata
  - May only need a few kB from a GB file
  - S3 facilitates this

Sharing mechanisms SHOULD allow for both permanent (long-term) and temporary access with automatic removal of shared access upon expiry

- It is easy to forget what has been shared
- And even if reminded, easy to miss that, or not get around to removal
- Policy engines such as RUCIO can facilitate this

Sharing mechanisms, both for the instigator and those accessing the data SHOULD offer API interfaces

- To allow programmatic data access
- Automated sharing of generated datasets
- Reducing the researcher burden
- RUCIO and S3 offer good options here

Data Movement Landscape

- Many large scale infrastructures with data movement considerations – these were investigated for best practice
  - EOSC
  - US IRI
  - CSC
  - NFDI
  - RFI
  - NXCT
  - Royce
  - MRC
  - Cryo-EM
  - EBI
  - JASMIN

Governance

- Should be a key consideration
  - Who owns the data
  - What permissions do they have
  - With whom is it being shared
  - What are the risks

Conclusions

- Many recommendations around data access, transfer and sharing
- There are some low hanging fruits
  - We can get a long way without significant complexity
  - Many things can be done quickly
- Keeping it simple is important
- Effort is required at sites to implement tools