



# UIS Storage Services

Wojciech Turek

[wjt27@cam.ac.uk](mailto:wjt27@cam.ac.uk)

Research Computing Platforms



## Agenda

Research data storage services - service overview, cost and roadmap

Storage service for institutions - update on the progress, introduction of a pilot service

Storage services platforms - deep dive into how we built our storage platforms and how they work

Utilising OpenStack infrastructure for Research Storage services



## Storage Services - the road so far

Early 2015 - Data Storage Survey conducted by the UIS, across University institutions, which concluded the following:

Approximately 15PB of data was being stored across various storage platforms. The majority of data were research data.

The most requested service was a DropBox-like storage service.

The second most requested service was a centrally provided filestore.



## **Storage Services - the road so far**

In 2016 UIS introduced two cloud based Storage Services:

DropBox for Business (significantly reduced price for members of the University)

OneDrive for Business (as part of the Microsoft ESS agreement).

In early 2017 UIS delivered three filestore services for research data:

Research Data Store (RDS)

Research Cold Store (RCS)

Research File Share (RFS)

Self service portal for storage services - facilitating quotes, purchases and management of data storage.



## Storage services for research data - key benefits

- Secure storage located in Cambridge
- Highly resilient with multiple copies in different locations
- Accessible from all university locations and home
- Designed specifically for researchers with medium to large data storage requirements
- Competitively priced

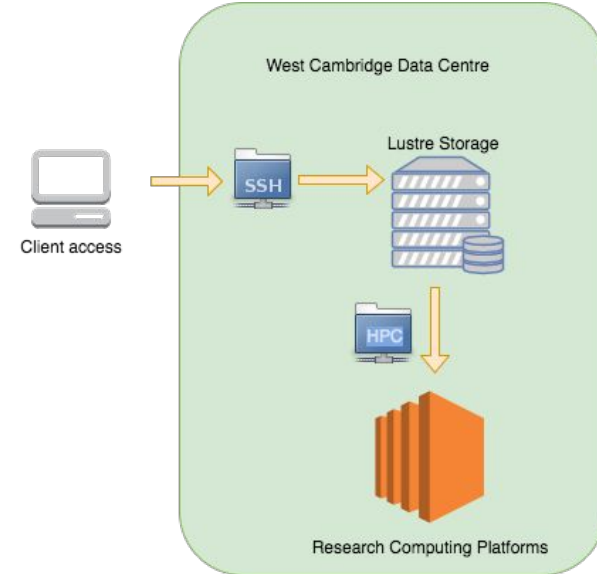


## Storage services for research data - key benefits

- Provided by self service portal
- Accessible through third party apps (e.g. Filezilla, Cyberduck)
- High performance access on HPCS
- Support available through expert storage-services support team

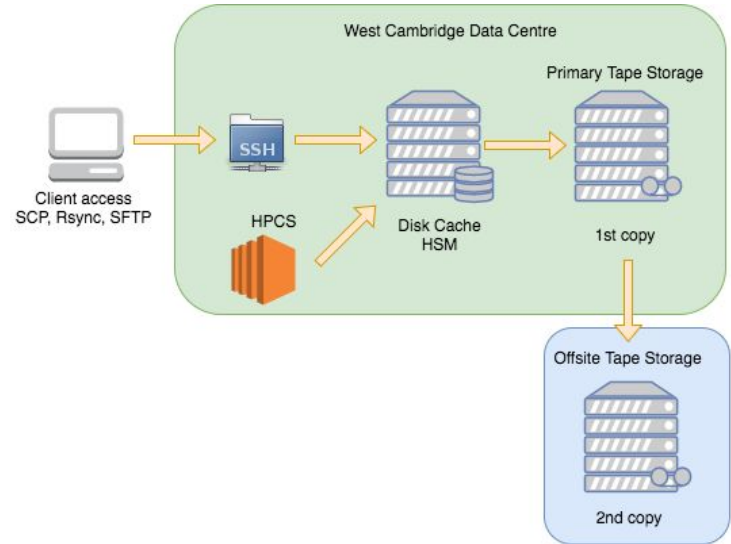
## Research Data Store (RDS)

- Capacity/performance cost optimised
- Mounted on Research Computing Platforms (HPCS, CSD3)
- Resilient platform with data protection (RAID6 and Failover)
- Accessible via RDS SSH gateway (SFTP, RSYNC)
- Optimised for medium to large files
- High data throughput
- Good for large scale storage requirements



## Research Cold Store (RCS)

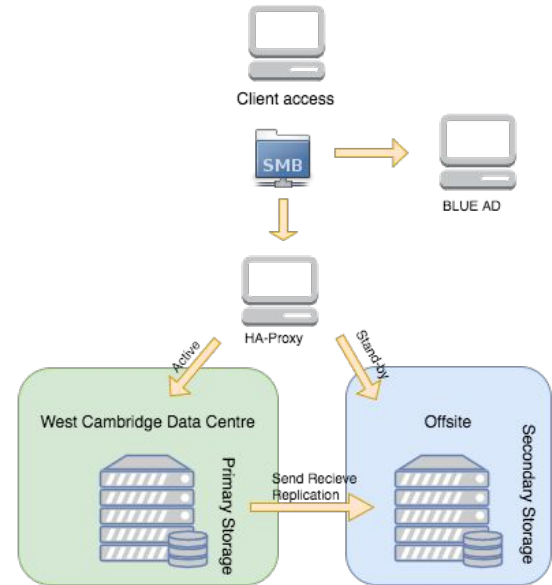
- Capacity cost optimised
- Accessible on HPC and via RCS ssh gateway (SFTP, RSYNC)
- Fast disk cache for uploading data
- Data must be bundled into large sets e.g. using tar to ensure good retrieval time (remember it's tape!)
- Data are moved to tape by HSM and stubs left on disk cache
- Data are pulled back from tape when stubs are accessed
- Specifically for “cold” data - not changing, not frequently accessed





## Research File Share (RFS)

- Resiliency/Availability optimised
- Highly resilient ZFS storage platform
- Replicated across two locations for backups
- User accessible snapshots going back one month
- Efficient storage with in-line compression
- CIFS mounted project shares
- Self service management of project groups and users
- Integrated with BLUE Active Directory
- Ideal replacement for local NAS boxes





## Self Service portal for storage

- Get quotes and save them in your self-service account or send them to your email
- Purchase storage licence
- Allocate storage to project groups
- Manage who has access to your storage projects
- Extend storage licence or buy more capacity



## How to buy and get access

Before you make a purchase make sure you have the following information ready:

- **Data Owner:** Principal Investigator (PI) who owns the data and takes responsibility for appropriate data being stored on the provided storage and its disposal after licence expiry
- **Data Managers:** List of persons who will have technical control over who has access to the data and can manage the storage licence and storage projects (Optional)
- **Data Users:** List of users who will need access to data (Optional)
- Length of the storage licence required (1,3 or 5years)
- Required Capacity




## How to buy and get access

- Use the Self Service portal to obtain quotes for the required licence, choosing the term and capacity
- Obtain a PO from your institution
- Use Self Service to upload the PO and make the purchase
- Configure the storage licence by assigning the Data Owner (mandatory) and Data Managers (optional)
- Create the first storage project by allocating space from the licence and choosing a name for the project, add Data Managers and users to the project
- Once storage is provisioned, all assigned users will receive an email with instructions for accessing the storage allocated to the storage project.



## Prices

- Changes to pricing for research storage services
- Fixed pricing independent of the length of the storage licence
- Volume pricing
  - 200TB+ for RDS and RCS
  - 100TB+ for RFS and IFS
- Customers who already purchased one of our services will be able to take advantage of the new prices - we will contact them to discuss options.



## Price - cont.

### Departments

Service	Price (£ per TB)			
	Per month	Per year	Per 3 years	Per 5 years
<b>Research Data Store (RDS)</b>	£4.50	£54	£161.99	£269.99
<b>Research Cold Store (RCS)</b>	£2.56	£30.67	£92.01	£153.35
<b>Research File Share (RFS)</b>	£11.64	£139.72	£419.17	£698.62



## Price - cont.

### Colleges & other institutions

Service	Price (£ per TB)			
	Per month	Per year	Per 3 years	Per 5 years
<b>Research Data Store (RDS)</b>	£5.40	£64.80	194.39	323.99
<b>Research Cold Store (RCS)</b>	£3.07	£36.80	110.41	184.02
<b>Research File Share (RFS)</b>	£13.97	167.67	503.01	838.35

Note: Colleges & other institutions prices include VAT.



## Price comparison

Supermicro storage server with 36 8TB Drives providing 200TB usable capacity	£25,000
Staff cost for 3 years - grade 8, 4h per week storage operation and maintenance	£15,000.00
<b>Total Cost over 3 years</b>	<b>£40,000.00</b>
<b>RDS cost over 3 years</b>	<b>£ 23,615.07</b>





## Storage services - roadmap

Institutional File Store - centrally provided filestore services for University of Cambridge institutions

- Service pilot to start in August
- Pilot study will evaluate functionality and requirements
- IFS project board will decide if requirements vs functionality metrics are a good match and decide if pilot service passes into production as is or requires further changes.

Infrastructure as a service - pilot service to start Q4 2017

- OpenStack compute and storage platform

Secure research computing platform

- Platform for storing and processing highly sensitive data (ISO27001 certified)



## Institutional File Store (IFS)

- Designed based on feedback we received from RMs and the IT community
- Using the same platform as RFS
- Currently in development, pilot service to start in mid August
- Cambridge institutions provision storage licence through Self Service portal
- Licence terms of 1, 3 and 5 years
- Fixed pricing per TB not dependent on licence term (minimum purchase 2TB)
- Integrated with BLUE Active Directory
- Get quotes and purchase IFS licence using Self Service



## Institutional File Store (IFS) - cont.

- Self Service integration for management of the IFS
- COs can allocate capacity available in their institutions licence to create group shares
- COs can allocate users to groups
- COs can choose retention policy (pre-set snapshot schedules) for shares
- COs can set default user quota for shares

Institutions interested in the pilot service please contact [storage-services@uis.cam.ac.uk](mailto:storage-services@uis.cam.ac.uk).