Tackling Big Data Storage Challenges

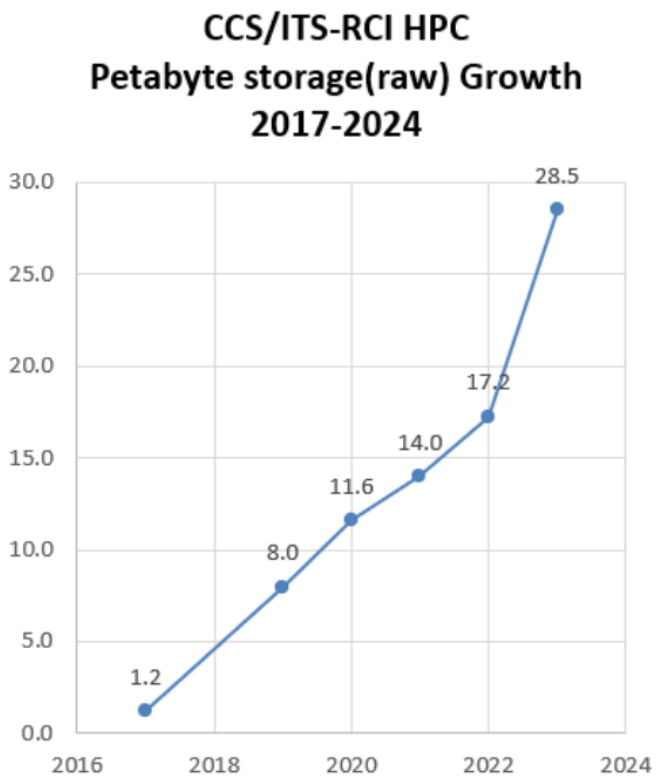
Tyler Burkett

Outline

- Considerations for Big Data
- UK/CCS Storage Options
- Other Storage Systems
- Data Transfer Options

Recent Trends

- Current Breakdown of 28 PB storage
  - 16.5 PB NAS
  - 6.5 PB GPFS
  - 5.3 PB Object Store
- Exponential Growth with increasing need for Big Data
  - Increases in AI workloads will only reinforce this trend.



Access Frequency (Temperature)

Important consideration is how often you access certain data sets or parts of them.

- Frequently accessed (hot) data on a slow storage systems wastes time in reading/writing data.
- Infrequently accessed (cold) data on fast storage wastes expensive space that could be better used for hot data.
- As we approach “big data” scales, incorporating multiple types of storage is required to effectively balance costs with performance.

Data Location & Bandwidth

Large storage means little if you can’t pull data fast enough.

- Increases in storage capacity necessitate increasing network capacity
- Where data is located is becoming more important as data size increases.
  - Generally, more physical distance = less bandwidth.
  - Moving data between NAS and LCC/MCC: 10’s Gbps
  - Moving the same data between UK and another KY institution: 1 Gbps, if your lucky.

Security & Maintenance

Big data sets are big targets for bad actors

- Verifying that data sets conform with regulations like HIPPA, FERPA, etc. to mitigate this becomes more challenging the larger the data set.
- Increased infrastructure complexity further exacerbates the issue.
- Has lead to push away from cloud storage and back to on-premises storage.

Preserving large data sets also becomes increasingly difficult.

- Partial data loss often can’t be tolerated, so increasing data size leads to increased likelihood of failure without additional actions.
- Data duplication adds redundancy, but making copies of large data sets can be difficult.

Files vs Object Storage

File system

- Most familiar format for users
- Hierarchical structure
  - Quicker to access data, but harder to expand systems
- Files can be easily modified after creation
- Better for smaller, structured data sets

Object System

- Uses less familiar APIs to access data (e.g. S3)
  - An entire object Must be read or written in one operation
- Flat structure
  - Slower access to data, but easier to expand
- Objects are static; changing an object requires re-create whole object
- Better for large or unstructured data sets

UK/CCS Storage Options

LCC/MCC Clusters

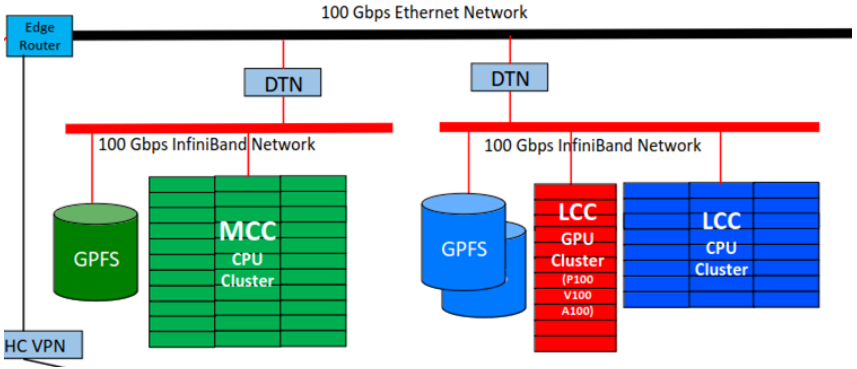
Storage space is largely temporary; meant to be used when running jobs.

Scratch space (GPFS)

- 90-day deletion policy
- 25 TB individual
- 50 TB project scrap

Permanent

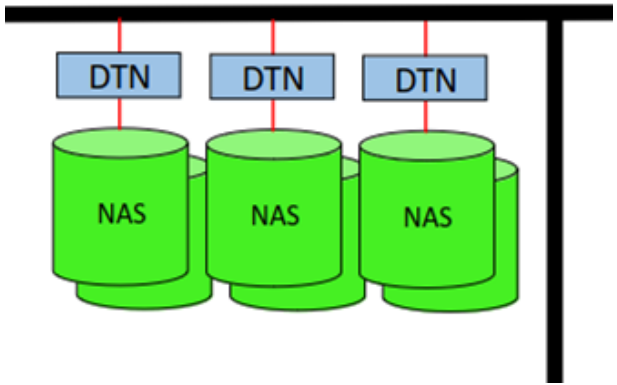
- Home: 10 GB
- Project: 1 TB
- Only intended for smaller items like scripts, test data, etc.



NAS (Gemini)

Meant for longer-term storage of data.

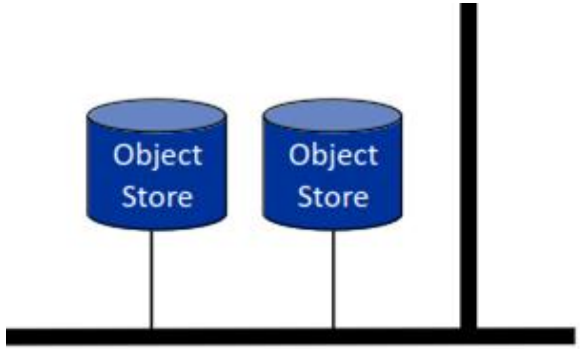
- Most are part of the paid condo model
  - Minimum purchase of 100 TB
- Separate storage system from compute systems
- Expected Workflow
  - Transfer data to work with to LCC/MCC
  - Run jobs
  - Transfer back any results back



Ceph Storage

Object storage system

- Good for large chunks of data that will be read/write (streamed) all at once.
- Group Quota of 30 TB
- Currently reaching EOL; not highly used



LabArchives

https://www.research.uky.edu/ERN

Electronic Research Notebook (ERN)
Cloud-storage

- Researchers are encouraged to store their research data in lab archived, but not required.
- Not designed to store all kinds of research data.
  - Not often a good fit for RCD data
- Max individual file size 16 GB

Tape Storage

Different section of ITS manages a tape archive system.

- ~$70/TB
- Intended as a complete system backup, not just select data sets.
- At this point, not the best place for tape storage.
    - The only exception might be if you already manage your own server with a large set of data.

Other Storage Options

Other Cloud Storage

- Google Drive
  - Free storage dropped to 15 GB free for UK users
- Microsoft OneDrive/Sharepoint
  - Soon to be greatly reduced (timeline unknown yet), but currently 5 TB of storage (up to 25 TB; contact ITS help desk to up)
- Other Cloud Storage providers (Box, DropBox, etc.) don’t have much better options for free storage compared to Google.
- AWS and Azure also offer cloud object storage
  - Pricing is often complicated to determine
  - Egress charges affect where/how transferring data impacts costs

Open Storage Network (OSN)

Object storage provided as a part of the ACCESS project.

- Requires submitting an allocations request.
- Initial Minimum of 10 TB allocation
- Can increase up to 50 TB
- Website: https://www.openstoragenetwork.org/
- Creating Allocation Requests: https://openstoragenetwork.readthedocs.io/en/latest/allocations.html

Tape Archiving (OURRstore)

Cheap tape archiving options provided by the University of Oklahoma

- Pay only for the tapes you send
  - Currently using LTO-8 tapes, which should cost about ~$10/TB
- Can transfer data via Globus
- Will auto-duplicate data when writing to tapes
  - Can have tapes shipped back when done writing
- Writing to tapes can be awkward if your data isn't already tape-friendly
  - Hard file size limit: 1 GB - 1 TB
  - Optimal file size: 20 - 200 GB
  - Tools like tar and split exist to group/divide files, but adds a step of reconstructing data when retrieving data off tape
- Best option for archiving; helps fulfills the 3-2-1 rule
  - At least 3 copies of the data
  - On at least 2 different storage mediums
  - With at least 1 in a physically distant location
- For more info: https://www.ou.edu/oscer/resources/ourrstore--ou---regional-research-store

Data Transfer Options

UNIX-Oriented File Transfer Options

- SCP
  - Simplest
  - One-off transfers across systems
  - Slowest method; no parallel download
- SSHFS
  - Network filesystem access; make remote directories accessible on a local filesystem
  - Similarly slow like SCP to copy files to local system
- Rclone
  - Multi-threaded transfers; faster
  - Works with a wide variety of kinds of storage systems than just NFS (e.x. Cloud storage)

Globus

Globus offers GUI access to data on a number of our systems.

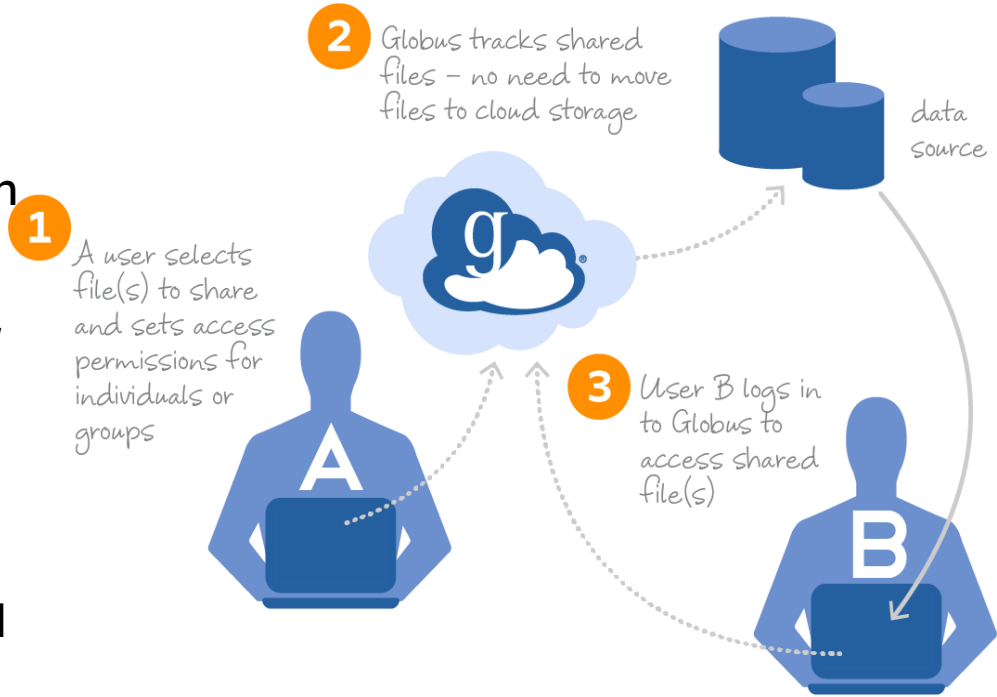
- Parallel transfers.
  - similar or (usually) better speed compared to using Rclone
- Available on most/all CCS resources and a number of storage systems outside UKy.
- Recommend method of transferring data.

https://www.globus.org/data-transfer

Sharing Data with Globus

For whatever collection you have access to, you can (if enabled) create guest collections to share your data on the system.

- You can create your own rules for who can access the data.
- Can share with those who don't have an account on the system you have access to.
- All CCS systems have this enabled with read-only permissions.



https://www.globus.org/data-sharing

Conclusion

- Playing catch-up with data storage will be the trend for most if not every institution, including UK.
  - Demanding more cross-institutional and vendor support for storage.
- Building out storage capacity and deciding where to store data is more than just deciding “how much space is there?”.
- Effective Big Data management will entail
  - Integrating many kinds of storage systems.
  - Matching data sets with the most appropriate storage systems.
  - Maintaining and synchronizing multiple copies of data.
  - Identifying if/when data sets need to migrate storage systems.