iRODS

Enabling eResearch with Automated Data Management from Ingestion Through Distribution

David Fellinger
Data Management Technologist
iRODS Consortium
October 17, 2018

File System and HSM evolution in HPC

- Early Beowulf clusters utilized the Network File System (NFS) to enable distributed data services.
  - Technology dating to the mid 1980s to provide service to individual users.
  - Network traffic and resource contention creates performance problems with no real parallelism.
  - Machine performance was limited because the I/O cycle was extended due poor file system and storage system performance.
- GPFS and Lustre were developed as revolutionary **parallel** file systems.
  - Multiple "gateway nodes" could be utilized as part of a cluster to enable simultaneous transfers of file system data.
  - Cluster nodes could run efficient file system clients.
  - In the case of Lustre, an asymmetrical MetaData Server (MDS) enabled centralized, accessible, metadata and lock management for an entire parallel file system.
- Hierarchical Storage Management (HSM) software evolved to enable data migration from the HPC file systems to archival and distribution systems.
- Data distribution systems like dCache enabled "push" publication to users.

HPC applications evolved driving greater requirements for I/O

- Initial HPC applications were primarily based in simulation or visualization.
  - Early large scale clusters were utilized to advance the understanding of nuclear fusion.
  - Difficult multi-dimensional matrices could be solved in parallel.
  - Visualization models were designed to allow researchers to understand plasma characteristics which could not be physically created.
  - Entire systems such as jet or rocket engines could be modeled in whole or part to shorten the design process.
- The proliferation of sensor data drove an entirely new set of HPC applications.
  - Genomic research.
  - High Energy Physics research.
  - Climate modeling.
- File systems and networks evolved to enable higher I/O bandwidth.
  - Infiniband and higher speed ethernet networks handled greater data rates.
  - Both storage and file systems evolved to solve “data burst” problems.
- **HSM systems remained proprietary servicing only archive requirements.**

Data management has become a critical requirement in analytical HPC.

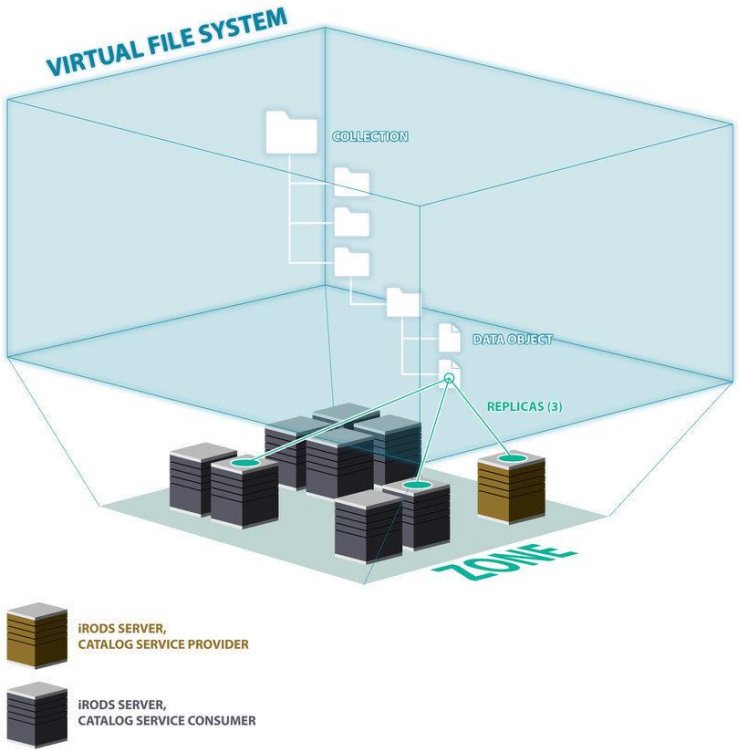
- The compute cluster is no longer the primary author of the collection but rather an element in the workflow.
- Huge amounts of data must be moved to and from the parallel file system adjacent to the compute cluster.
- Data must be organized based on content or specific metadata.
- Processes must be tracked.
- Process results may require additional compute time and that must be reported, allocated, and tracked.
- Output products must be cataloged and moved off of the scratch file system.

The Integrated Rule-Oriented Data System middleware can provide all of the required data management to enable HPC analytics.

iRODS is;

- Open source
- Distributed
- Metadata driven
- Data Centric

Data Virtualization

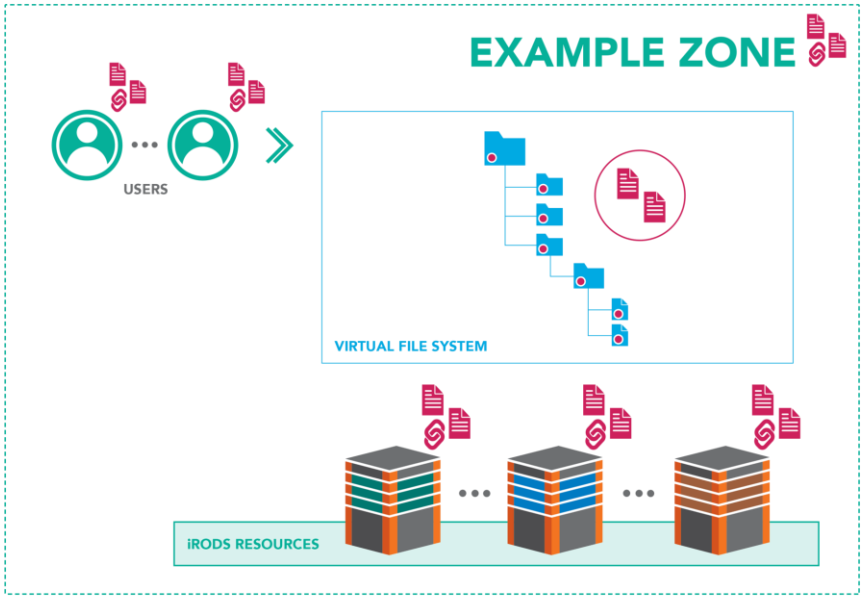


Combine various distributed storage technologies into a Unified Namespace

- Existing file systems
- Cloud storage
- On premises object storage
- Archival storage systems

- iRODS provides a logical view into the complex physical representation of your data, distributed geographically, and at scale mapping physical to logical paths

Data Discovery

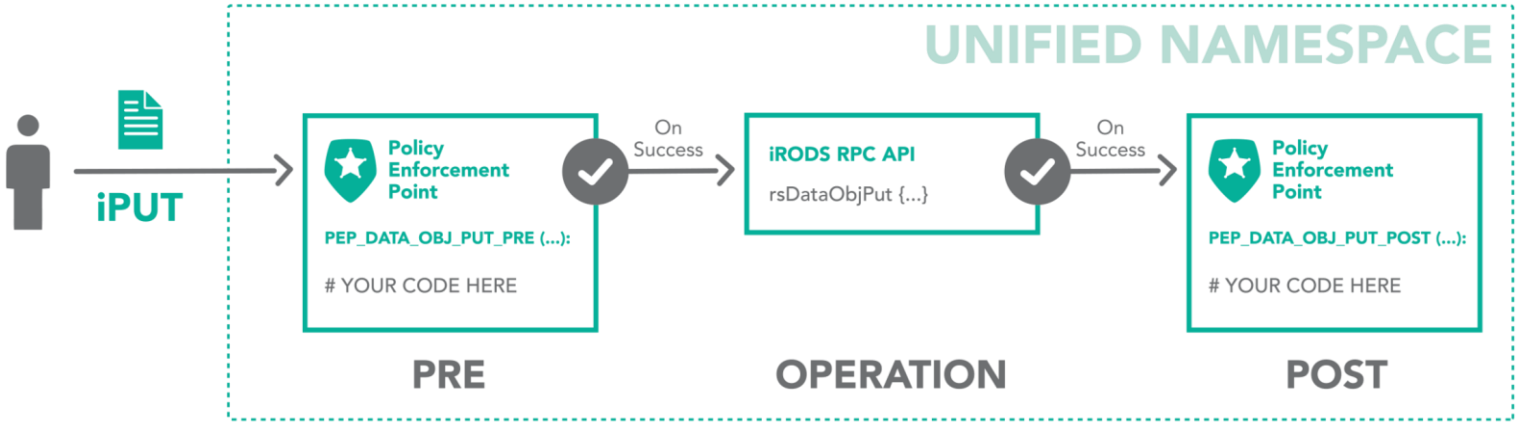


Attach metadata to any first class entity within the iRODS Zone

- Data Objects
- Collections
- Users
- Storage Resources
- The Namespace

iRODS provides automated and user-provided metadata which makes your data and infrastructure more discoverable, operational and valuable.

Dynamic Policy Enforcement



The iRODS rules may;

- Control access
- Build a log for audit and reporting
- Send notifications to users, schedulers, etc.
- Extract metadata
- Provide additional context

Integrating iRODS Enabling Workflow Automation in HPC

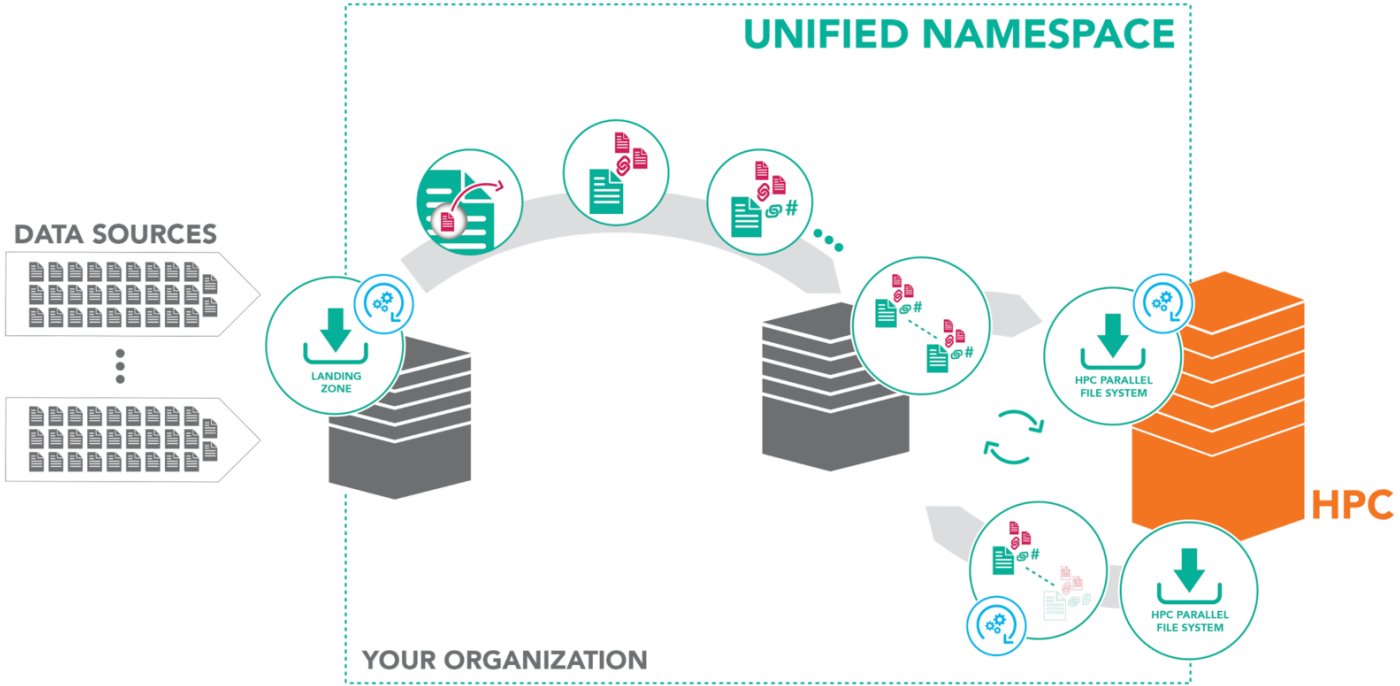
In order of increasing complexity...

iRODS as a compute orchestrator
- Launch a job via irule, or as part of a PEP
- Implement a Landing Zone for product capture

iRODS as part of a compute job script
- Stage the source data via replication for the application
- Capture the products and ingest them into iRODS

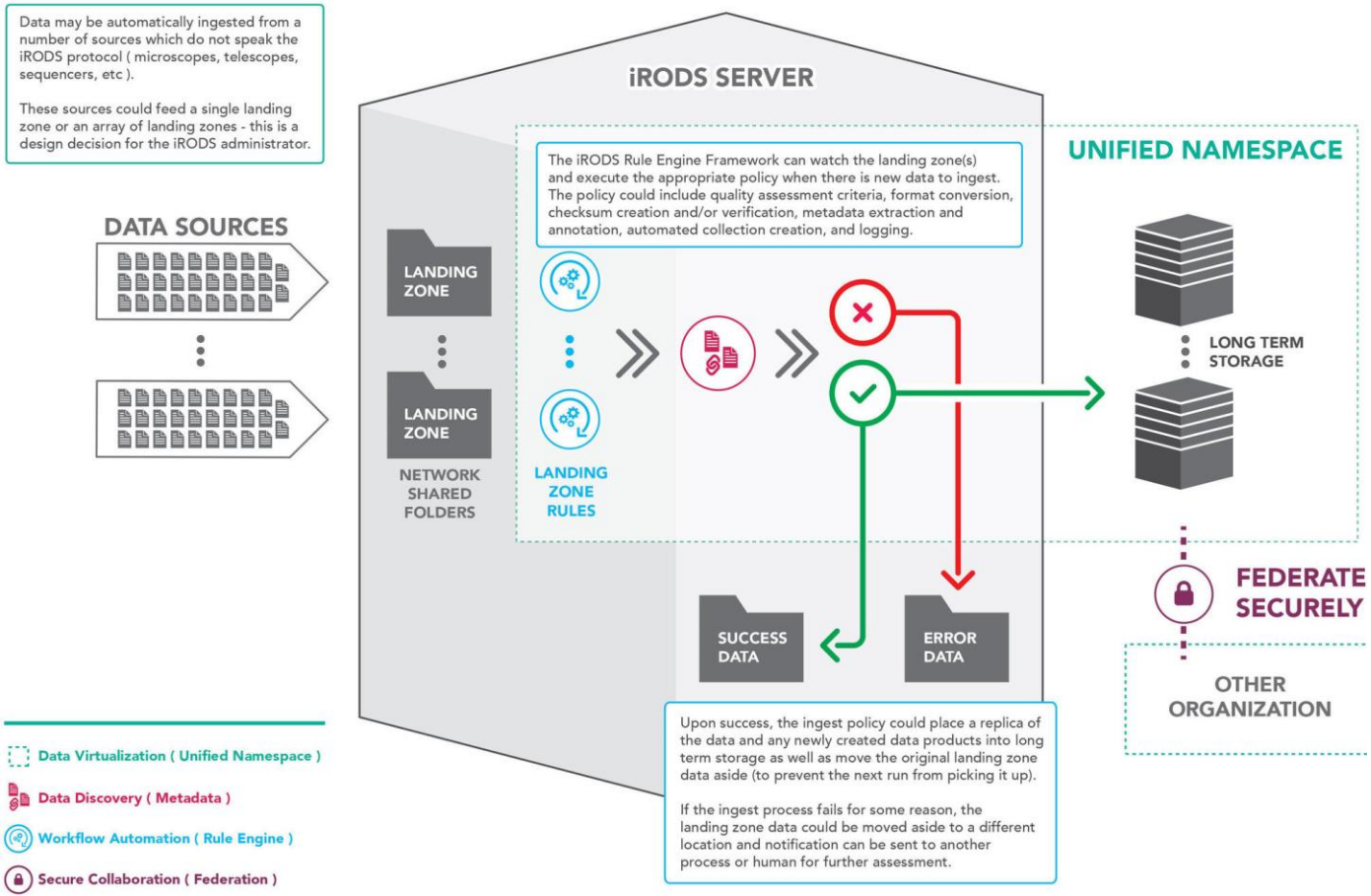
iRODS as part of the compute application
- Compute application directly leverages the iRODS API to open, read, and write data

Taking Data to Compute



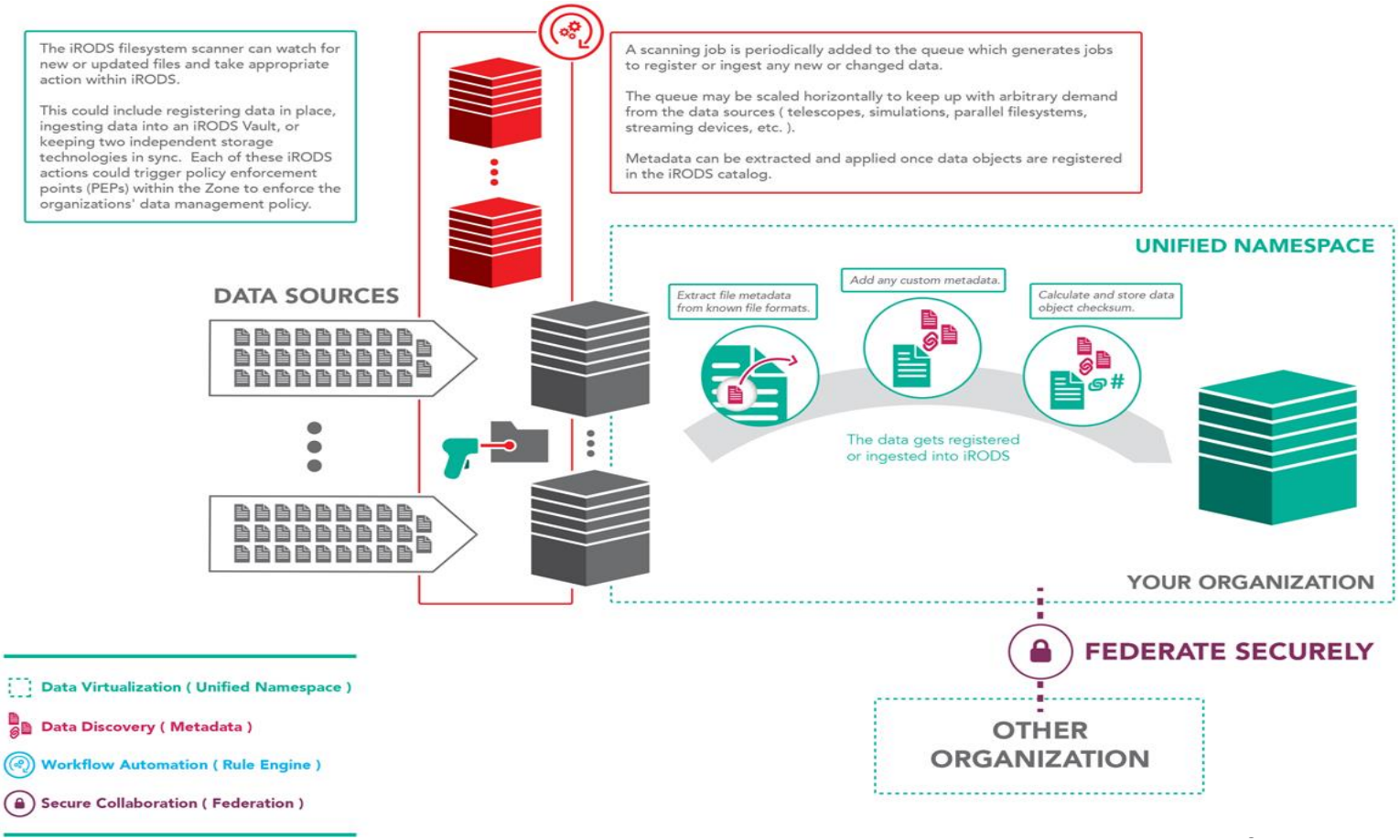
iRODS stages the data to a high performance scratch file system but is not in the compute path.

The Landing Zone; Organizing Data based on Policy



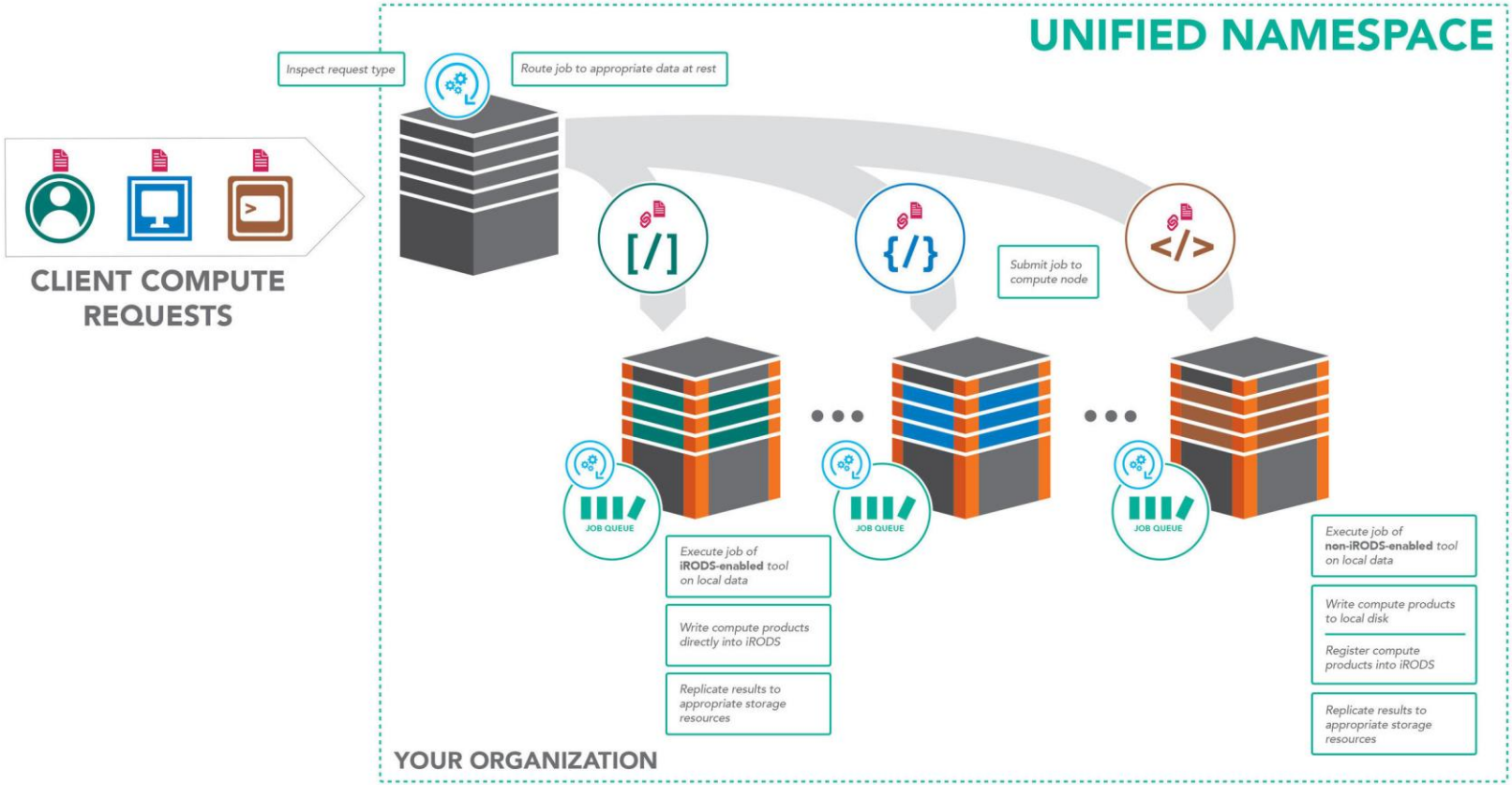
iRODS CONSORTIUM

Automating data ingestion

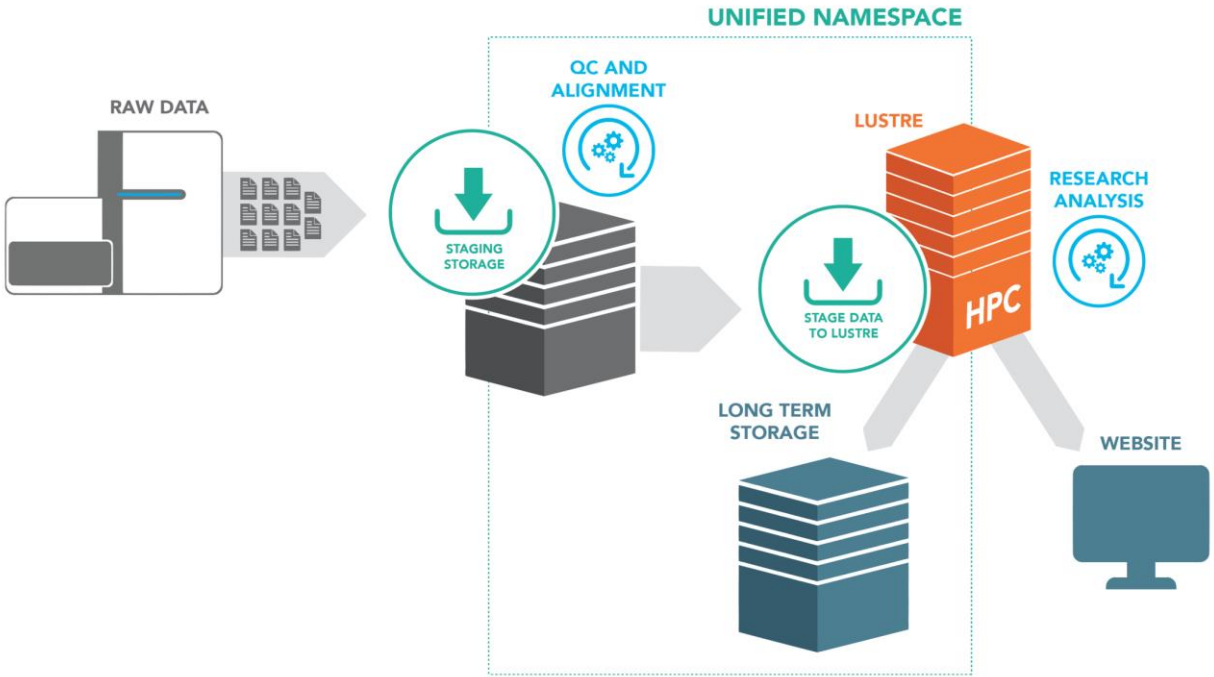


iRODS CONSORTIUM

Compute to data use case enabling distribution

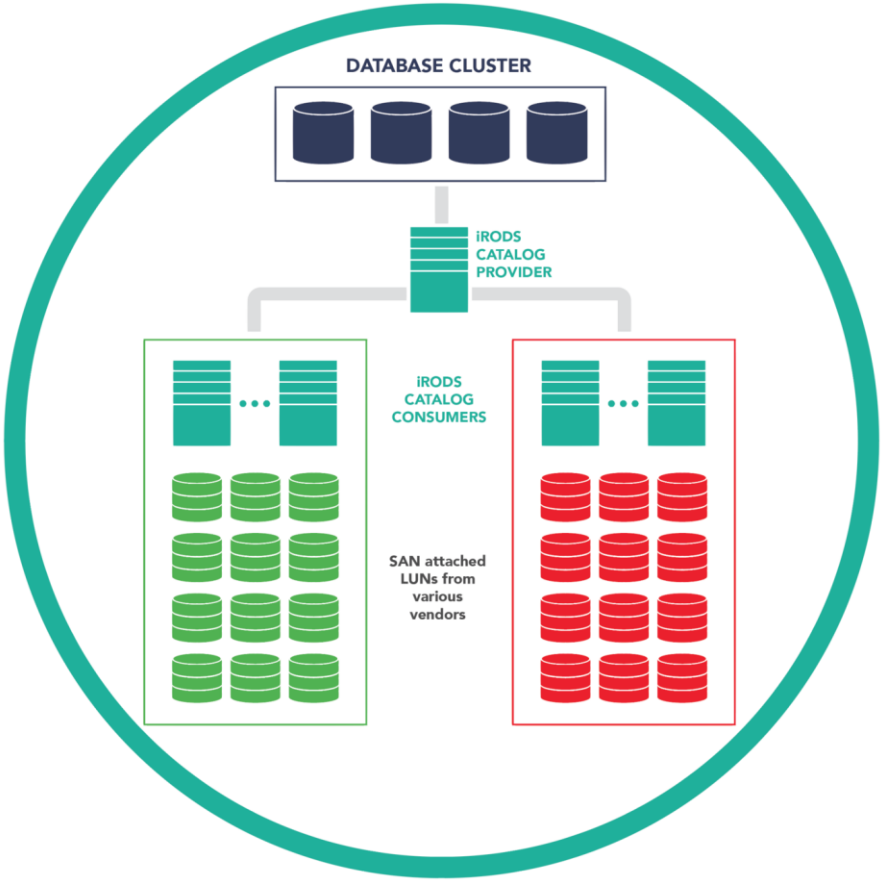


Use case; The Wellcome Sanger Institute



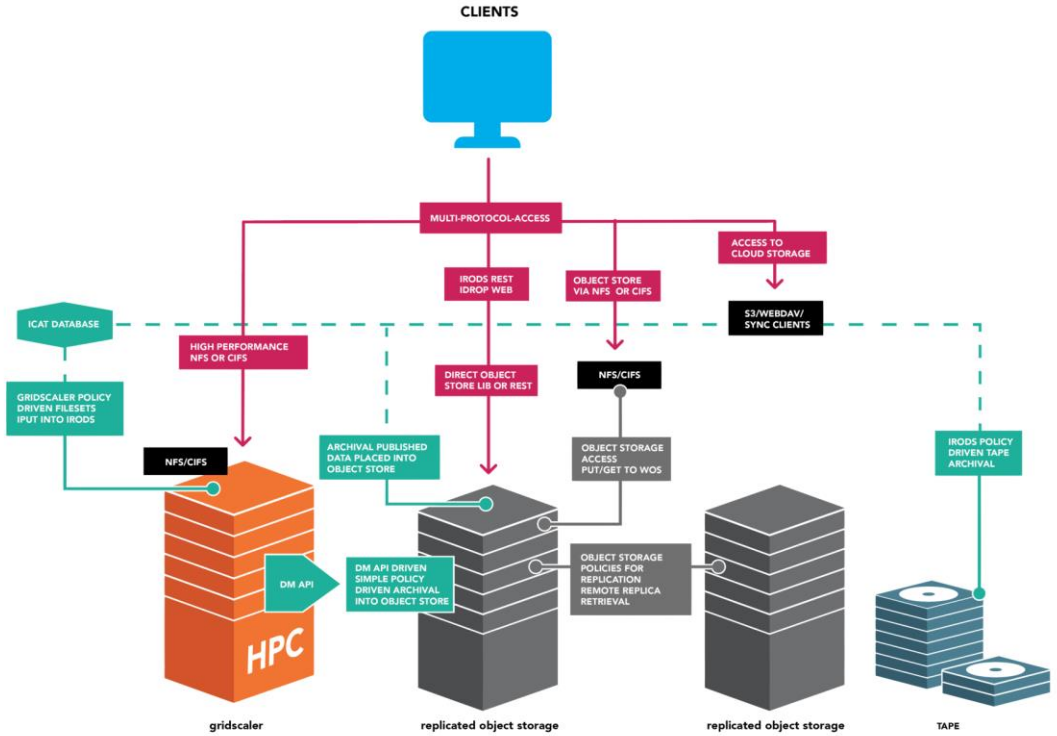
Sanger-Replication

- Data preferentially placed on resource servers in the green data center (fallback to red)
- Data replicated to the other room.
- Checksums applied
- Green and red centers both used for read access.



Use case; The University College London

- UK sponsored data requirements; last date of citation plus 10 years.
- iRODS tiers data across storage technologies and provides a virtual unified namespace.
- Federated access between other centers is enabled.
- Complete audit and reporting functionality within iRODS allows demonstration of compliance.



Conclusion

- Traditional HSM solutions generally manage data transfers to an archive based on file attributes not metadata properties or file content.
- HPC data management has become far more complex in dealing with sensors and “big data”.
- Automated, policy-based, data reduction has become a requirement in life sciences and related fields of study.
- iRODS can enable complete workflow control, data lifecycle management, and present discoverable data sets with assured traceability and reproducibility.

The iRODS Consortium (iRODS.org)

The iRODS Consortium

- Leads software development and support of iRODS
- Hosts iRODS Events
- Tiered membership model


Questions?

Thank you!

David Fellinger

davef@renci.org

iRODS.org