

Storage Providers' Perspective

WLCG Workshop Data Session

Andrea Ceccanti, Patrick Fuhrmann, Andrew Hanushevsky, Oliver
Keeble, Andreas Peters

Where can the storage systems help?

- Reduce cost/volume
 - cost of storage management
 - integrating standard (non HEP) solutions e.g. ceph
 - protocol zoo, SRM-less operation
 - T2 storage as cache
 - multi-site storage
 - regional federations
 - cloud storage
 - system manageability
 - storage overheads
 - redundancy
 - replication, erasure, RAID levels etc
 - reduce system reliability requirements?
 - reduce cost/impact of data loss
 - component technology
 - shingled disks
 - consumer/enterprise disks
- Reduce volume used
 - reduced number of global replicas
 - remote access
 - latency hiding
 - applications, overcommitting
 - global federations
 - CPU-only resources (inc cloud)
 - data formats and lifecycle, intermediate products
 - resource reporting
 - monitoring usage
 - eliminating dark data
 - data "enrichment"
 - popularity
 - caching, avoiding unused data
 - promoting locality in workflows
 - trading disk for...
 - tape
 - data parking
 - CPU
 - maintain metadata enabling regeneration of data on demand

Minimizing Data Replicas

- Storage federations address this
 - Pool geographically separated storage
 - Better data placement to minimize duplication
 - One large pot is better than many small ones
 - But not all data is now co-located to CPU
 - Network resources become a critical factor
 - Storage sites must be willing to share resources
 - Non-local use of a site's resources
 - This may lead to tit-for-tat political problems

Andy's Input

Architectural Approaches

- Regional federations
 - Minimize RTT where storage is abundant
 - Makes cross-site communication easier
 - Regions more amenable to resource sharing
- Global federations
 - Maximize data coverage (RTT secondary)
 - Works best if not the primary means of data access
 - Sites less amenable to far-flung resource sharing

Andy's Input

Key Federation Road Blocks

- Site reliability to meet overall expectations
 - May need to tier federation by reliability
- Cross-site communication
 - Problem resolution across disparate domains
- Engendering need for the common good
 - Monitoring to prove everyone benefits
- Road blocks are not due to the technology
 - Federating different priorities is the challenge

Andy's Input

Storage Cost Reduction

- CERN IT testing consumer grade HDD with a potential cost saving of ~30%
 - evaluate also SMR drives
- Use erasure coding e.g. (12,4) in no-repair mode for non-random access data
 - no disk replacement after burning-in
 - >99% of data volume is write-once
 - orthogonal to CERN-Wigner setup
 - EC assembly on client side (avoid doubling network IO on read)
- openLab evaluation/integration of OpenKinetic object disks
- big data repository (today 100PB) with CEPH back-end is not (yet) cost effective (additional journaling disks, memory & network requirements for EC)
- CERN Tape Archive project with modular interface to couple to any storage system with notification, extended attributes and XRootD protocol bridge

EOS & Castor
Input

dCache: 'reducing the cost of storage'

- Full support of "High Availability", allowing to cover any 'one' subsystem outage without the necessity of immediate manual intervention,
 - resulting in
 - the reduction of operations team sizes
 - reducing 'full operation' hours.
 - HA available with dCache 3.1
 - See CHEP poster on dCache HA by Paul and Gerd.
 - "Towards redundant services in dCache", Poster Thursday afternoon
- Support of "Lights-Out Management" by migrating towards a consistent (REST) interface for system administration and data management,
 - Benefit:
 - simplifying the integration of dCache operations into existing monitoring and administration frameworks.

dCache: 'reducing the cost of storage'

- Building upon well established low level storage technologies.
 - Benefits:
 - "low maintenance",
 - "horizontally scaling",
 - "Open Source" is preferred (option for payed support.)
 - Example : CEPH
 - Step wise integration of CEPH into dCache
 - CEPH only prototype for other shared object or block storage systems.
 - See Tigran's CHEP talk
 - "dCache on steroids - delegated storage solutions", Monday 14:00
- Supporting standards in data transfer, data access and authentication mechanisms.
 - Benefit:
 - Simplify sharing of resources among different scientist communities.

dCache: 'reducing the cost of storage'

- Improving the 'federation feature' of dCache, already in place for a decade at NDGF and Uni Michigan.
 - Benefits:
 - Human Resource synergy effect between sites.
 - Single Storage Endpoint seen from experiments point of view, while storage highly distributed across sites or countries.
 - Small sites can benefit from their storage as cache system only.
 - (Not clear yet, if that's not only a political effort)
 - Demonstrator in Russia and Germany,
 - **Russia** : Kurchatov, Dubna and St. Petersburg, **Germany**: HGF Data Alliance
 - Planned work :
 - Improved authorization of components against each other.
 - Improved local caching setups (if necessary)
 - But mostly integration and verification of features, already in place.

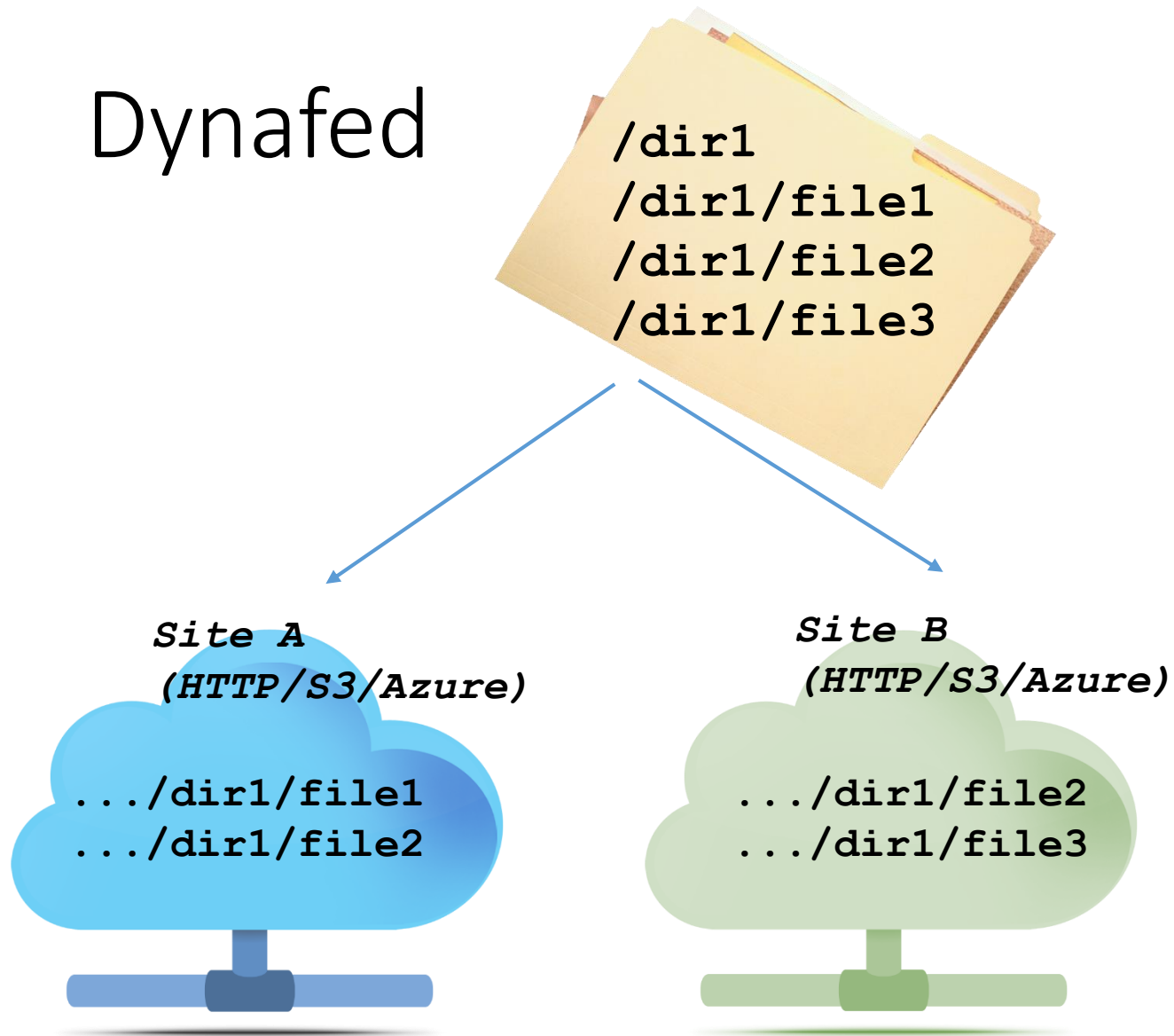
StoRM

- - SRM-less operations (mainly provide space usage information via webdav)
- - StoRM CDMI plugin: provide a CDMI interface for bring onlines
- - StoRM WebDAV integration with INDIGO AAI (i.e. OpenID-connect)

DPM 1.9

- SRM free operation
 - Quotas
 - “space token” concept generalised and mapped into namespace
 - Space reporting – used/free via HTTP/DAV
 - Reporting on “space tokens” and subdirectories
 - GridFTP redirection enables scalability
- Simplified system
 - All internal communication over HTTPS
 - Control and data tunnelling
- Caching laboratory
- Improved dmlite-shell
 - Improved manageability

Dynafed



- Global HTTP federations
 - Reduced number of replicas
- Regional HTTP federations
 - E.g. sharing a dataset transparently between sites
- Integration of cloud storage

Where can the storage systems help?

- Reduce cost/volume
 - cost of storage management
 - integrating standard (non HEP) solutions e.g. ceph
 - protocol zoo, SRM-less operation
 - T2 storage as cache
 - multi-site storage
 - regional federations
 - cloud storage
 - system manageability
 - storage overheads
 - redundancy
 - replication, erasure, RAID levels etc
 - reduce system reliability requirements?
 - reduce cost/impact of data loss
 - component technology
 - shingled disks
 - consumer/enterprise disks
- Reduce volume used
 - reduced number of global replicas
 - remote access
 - latency hiding
 - applications, overcommitting
 - global federations
 - CPU-only resources (inc cloud)
 - data formats and lifecycle, intermediate products
 - resource reporting
 - monitoring usage
 - eliminating dark data
 - data "enrichment"
 - popularity
 - caching, avoiding unused data
 - promoting locality in workflows
 - trading disk for...
 - tape
 - data parking
 - CPU
 - maintain metadata enabling regeneration of data on demand