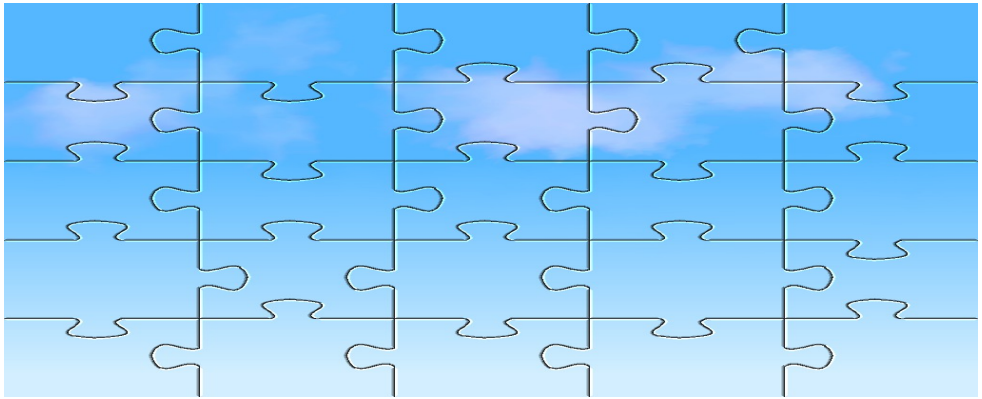


IBM **Spectrum Scale**

Cloudy Jigsaw Puzzles

Proof-of-Concepts with complex integrations

Harald Seipp, Leader, CoC for OpenStack Storage

With input from Rob Basham, Anbazhagan Mani and many more...

IBM

Puzzle 1: NFS-to-Object Gateway for Cleversafe Object Store

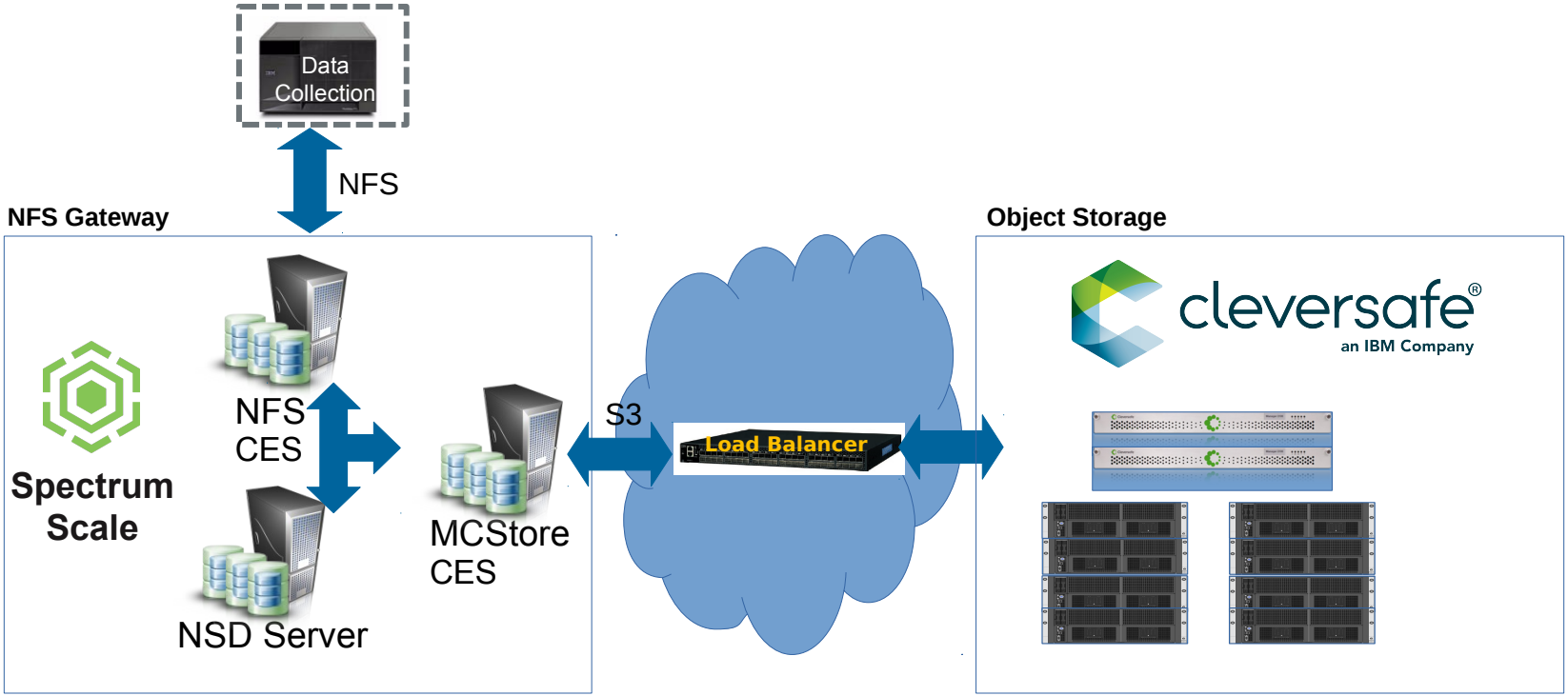
Requirements

- NFS native data access or NFS via Appliance
- HTTP RESTful S3 API and SWIFT API access
- More than 100 Million Objects
- More than 1000 Clients or client groups
- Multi-tenancy, authorization and authentication
- Encryption (out-of-scope for PoC)
  - Data-in-transition encryption (at least HTTPS)
  - Data-at-rest encryption (encrypted disks)

Jigsaw Pieces

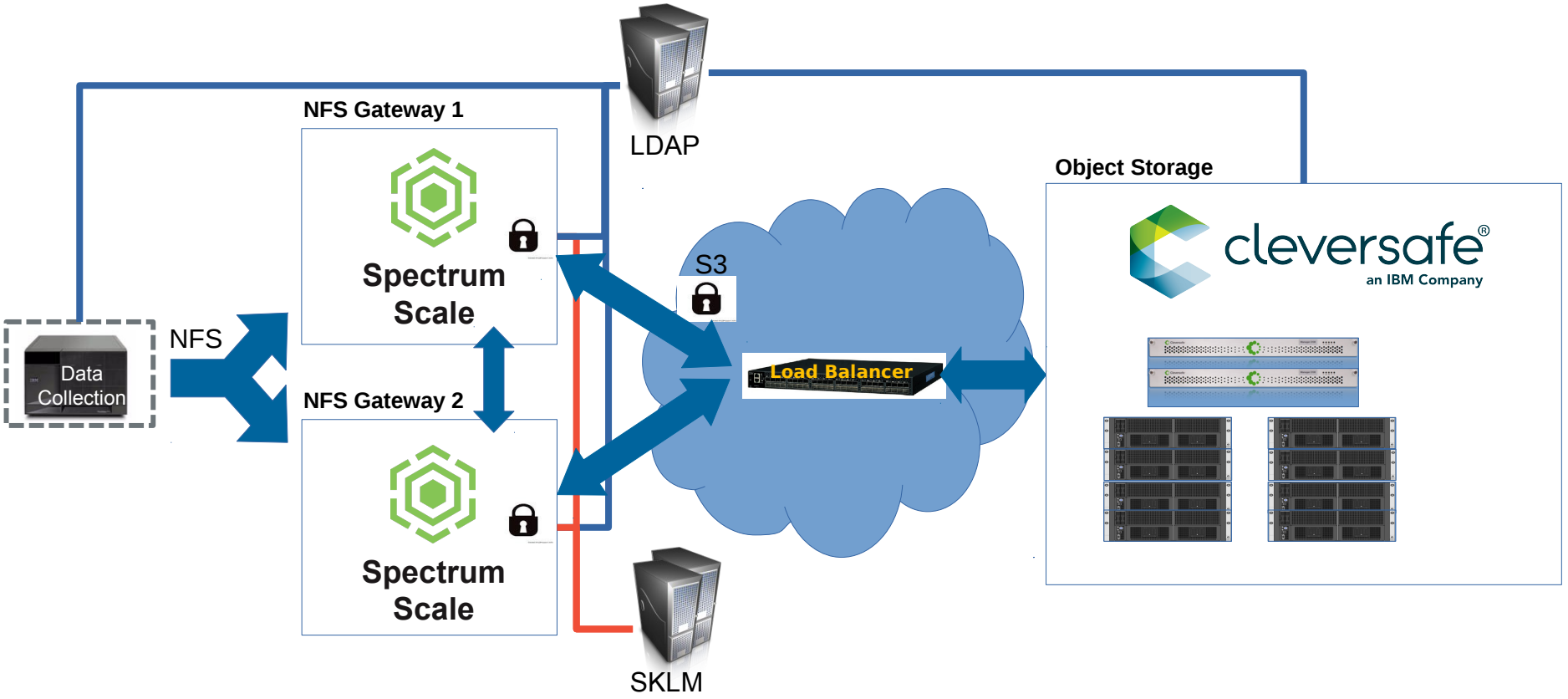
- Cleversafe as a cost-effective highly scalable Object store
- Spectrum Scale as a scalable file system with NFS cluster export services
- Spectrum Scale transparent cloud tiering (MCStore) to connect the above

Solution Architecture*



* Simplified, as used for PoC

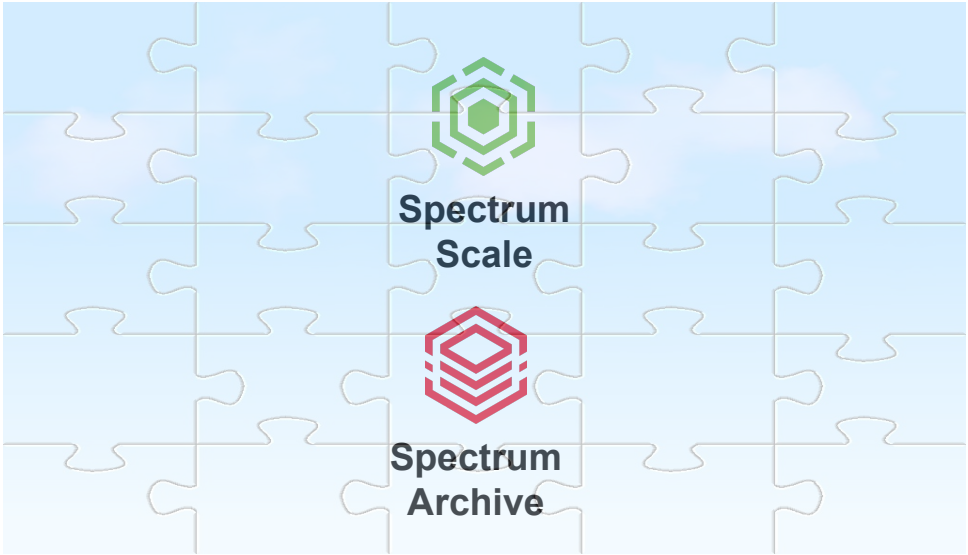
Solution Architecture – Future



Observations

- Multi-tenancy though multiple MCStore nodes that map 1:1 to file systems
  - MCStore nodes are connected to same cloud storage, different tenant or
  - different cloud storage
- Data deleted through NFS requires reconcile job to also delete cloud data
- In-flight encryption through HTTPS to Cleversafe object store requires self-signed certificate installed on MCStore node (not yet supported with MCStore OpenBeta)
- NFS CES should run on separate node than MCStore to optimize resource usage
  - Note that MCStore requires a CES node, so NFS protocol will get installed there – just do not use it
- MCStore does work with an encrypted file system
  - On migration, data will get (GPFS-)decrypted, then encrypted for transmission by MCStore

Puzzle 2: Hybrid Spectrum Scale Cloud with off-premise Tape Archiving



Requirements

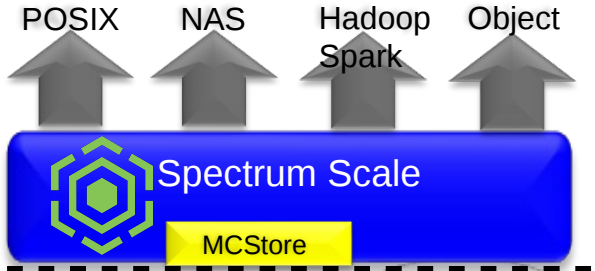
- On-premises Storage
  - Multi-protocol on-premises storage solution
  - Seamless and efficient cold data migration to Cloud
- Cloud Storage
  - Easy to use enterprise-class cloud based storage service
  - Optimally store huge amount of cold data or inactive data
  - Capacity of storage to be efficiently utilized
  - Very low-cost storage, affordable and predictable retrieval costs
  - Data retrieval in minutes

Jigsaw Pieces

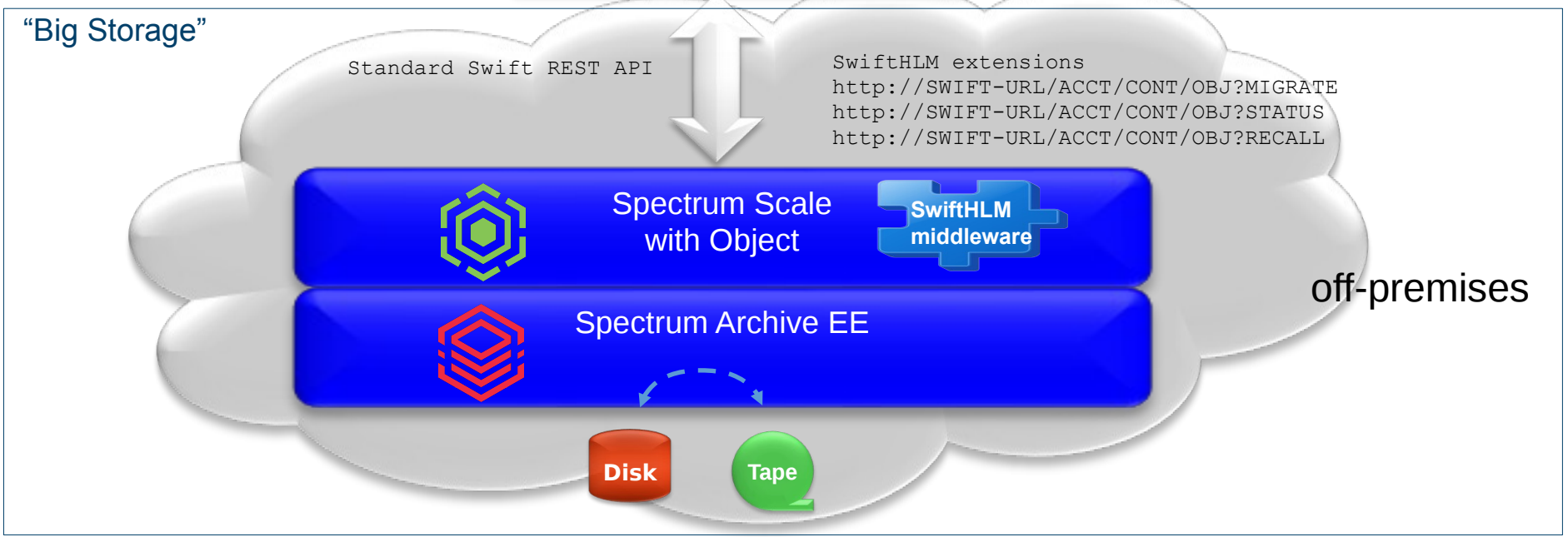
- On-premises
  - Spectrum Scale as scalable multi-protocol primary storage
  - Spectrum Scale transparent cloud tiering (MCStore) as Cloud connector
- Off-premises (“Big Storage”)
  - Spectrum Scale with Object as Cloud storage
  - Spectrum Archive Enterprise Edition as highly scalable solution to migrate data to Tape with minimum Storage TCO
  - IceTier prototype with SwiftHLM* OpenStack Swift middleware and backend for controlled data movement between Disk and Tape

*HLM=High-latency media

Solution Architecture



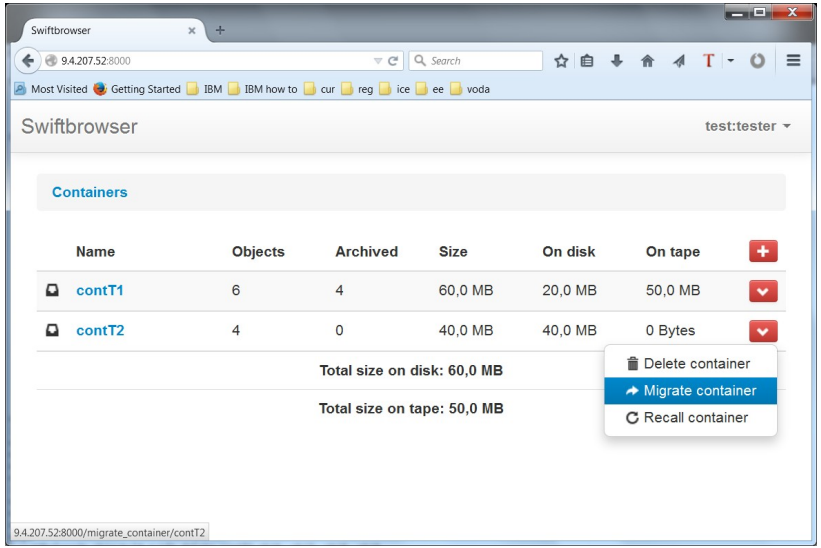
on-premises



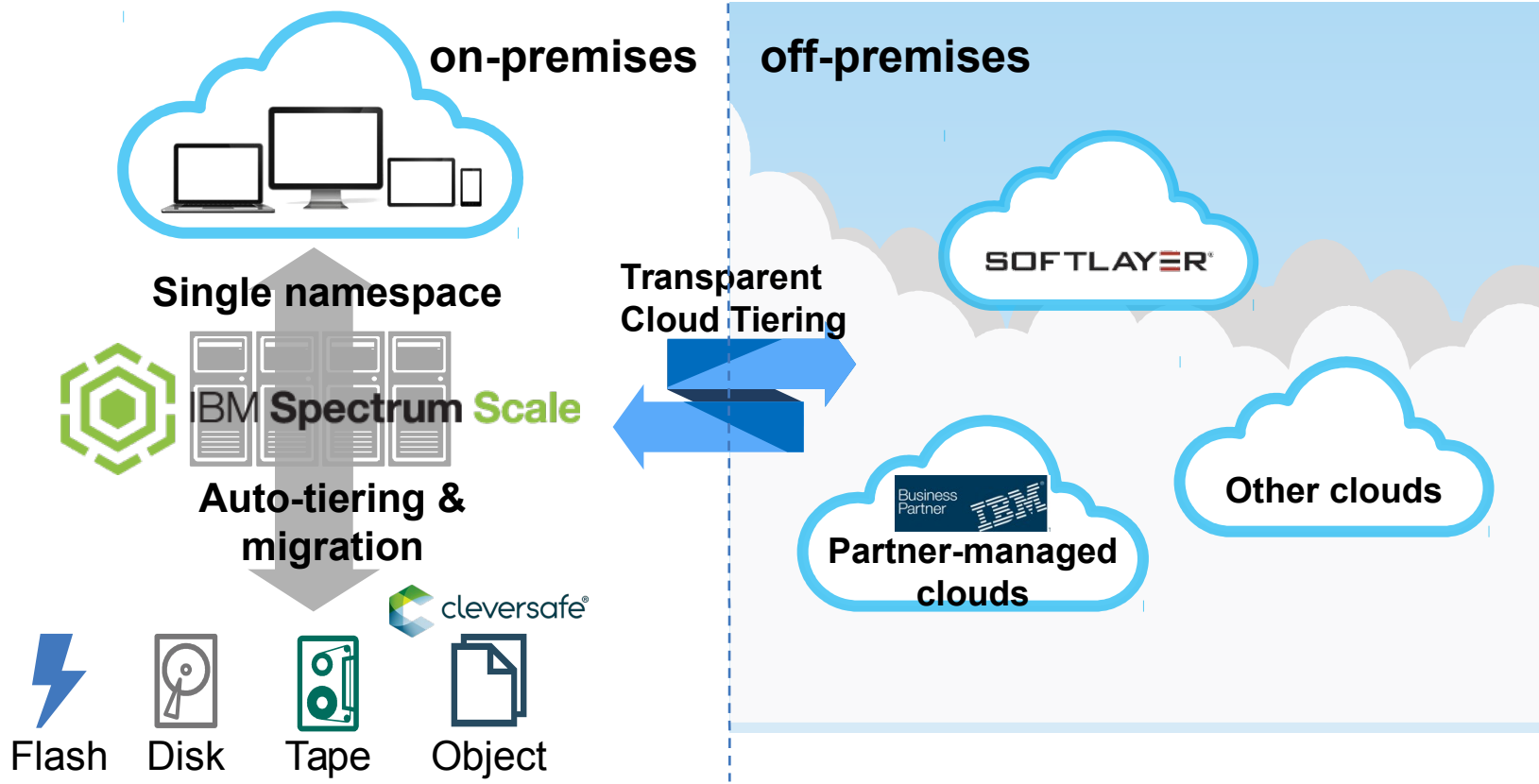
off-premises

Observations

- Do not use for small files (<1MB), avoid deletions
- Migration of on-premises S3/Swift Objects is not yet supported by MCStore
- Migration / Recall control
  - Manually through prototype Web Interface on the Spectrum Scale Object node (see picture on the right)
  - Could be automated later
    - Would require SwiftHLM extensions built into MCStore



IBM Spectrum Scale – Intelligent data placement for hybrid cloud



References

- Download Spectrum Scale transparent cloud tiering Open Beta from http://www.ibm.com/developerworks/servicemanagement/tc/gpfs/evaluate.html
- Get Spectrum Archive EE trial VM (soon) from: https://www.ibm.com/developerworks/servicemanagement/storage/sar/index.html
- Redpaper on Spectrum Scale Object and Spectrum Archive: https://www.redbooks.ibm.com/redbooks.nsf/RedpieceAbstracts/redp5237.html?Open
- “Big Storage”: https://www.ibm.com/developerworks/servicemanagement/sds/bs/
- IceTier and SwiftHLM : http://www.research.ibm.com/labs/zurich/sto/tier_icetier.html https://wiki.openstack.org/wiki/Swift/HighLatencyMedia https://github.com/ibm-research/SwiftHLM

Trademark and disclaimers

Trademarks and disclaimers (cont.)

NOTES:

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area. Prices are suggested US list prices and are subject to change without notice. Starting price may not include a hard drive, operating system or other features. Contact your IBM representative or Business Partner for the most current pricing in your geography. Any proposed use of claims in this presentation outside of the United States must be reviewed by local IBM country counsel prior to such use.

Information is provided “AS IS” without warranty of any kind. The information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any performance data contained in this document was determined in a controlled environment. Actual results may vary significantly and are dependent on many factors including system hardware configuration and software design and configuration. Some measurements quoted in this document may have been made on development-level systems. There is no guarantee these measurements will be the same on generally-available systems. Users of this document should verify the applicable data for their specific environment.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM makes no representation or warranty regarding third-party products or services.