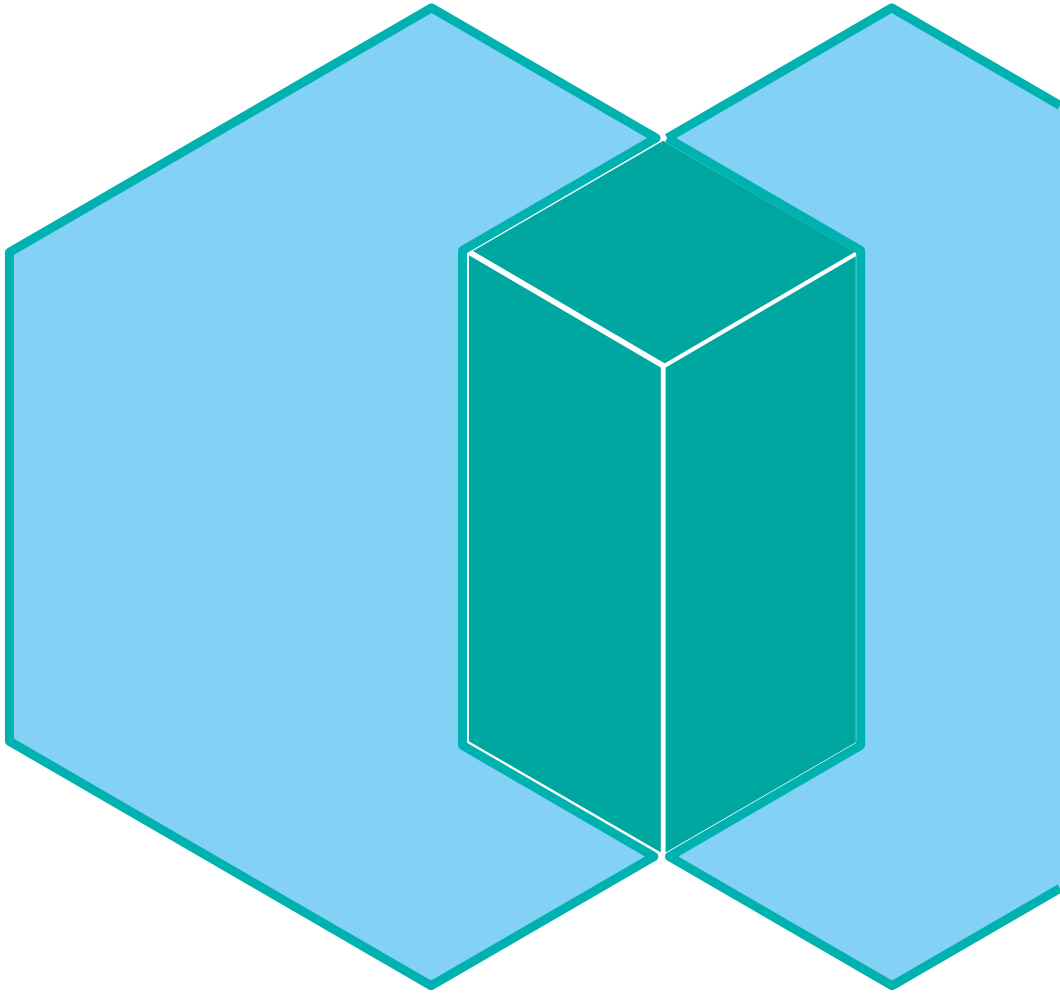


Multi-cloud with Transparent Cloud Tiering

Harald Seipp, Center of Excellence for Cloud Storage

Rob Basham, Hybrid Cloud Storage Architect

Agenda

- Spectrum Scale Transparent Cloud Tiering introduction
  - Capabilities
  - Cloud integration scenarios
- Real-world use cases
  - Autonomous driving measurement data
  - Bodycam recordings
  - Tape backup replacement
  - StaaS provider leveraging public cloud
  - Video surveillance „anti-use-case“
- Future: Multi-site support considerations

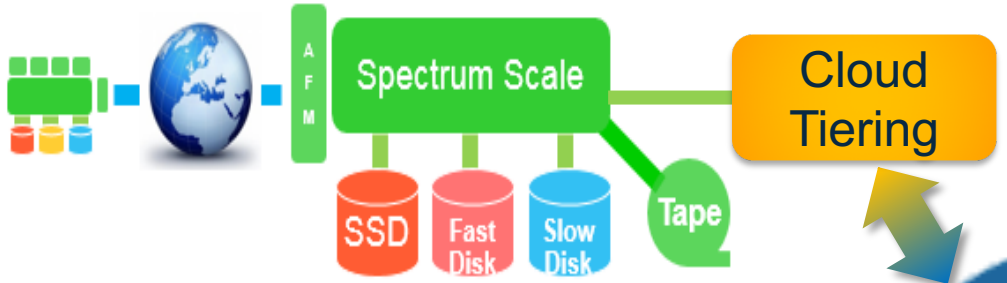
Transparent Cloud Tiering Is Used Across IBM Storage

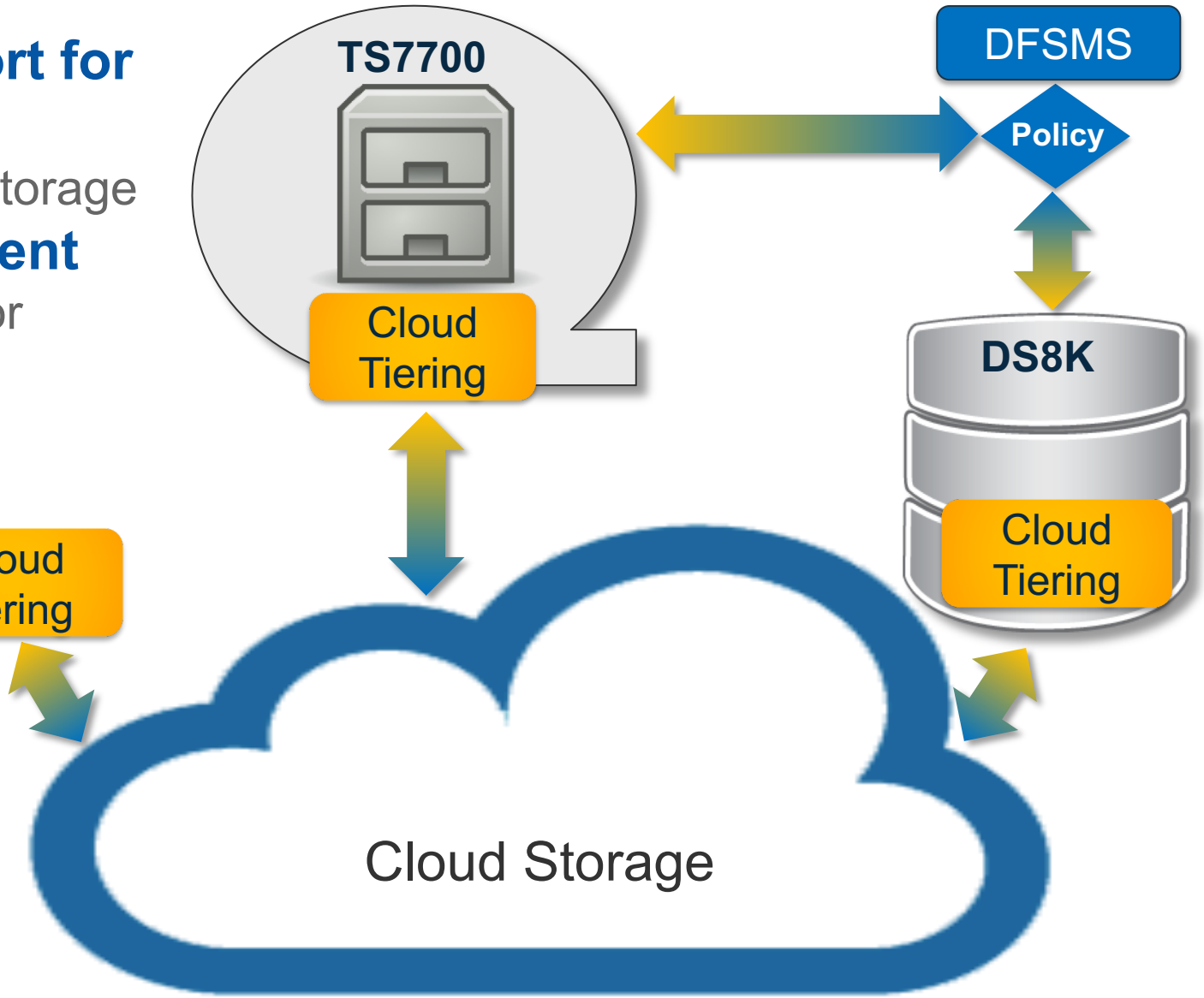
Native IBM Storage Support for Object Storage

- On or off-premises object storage

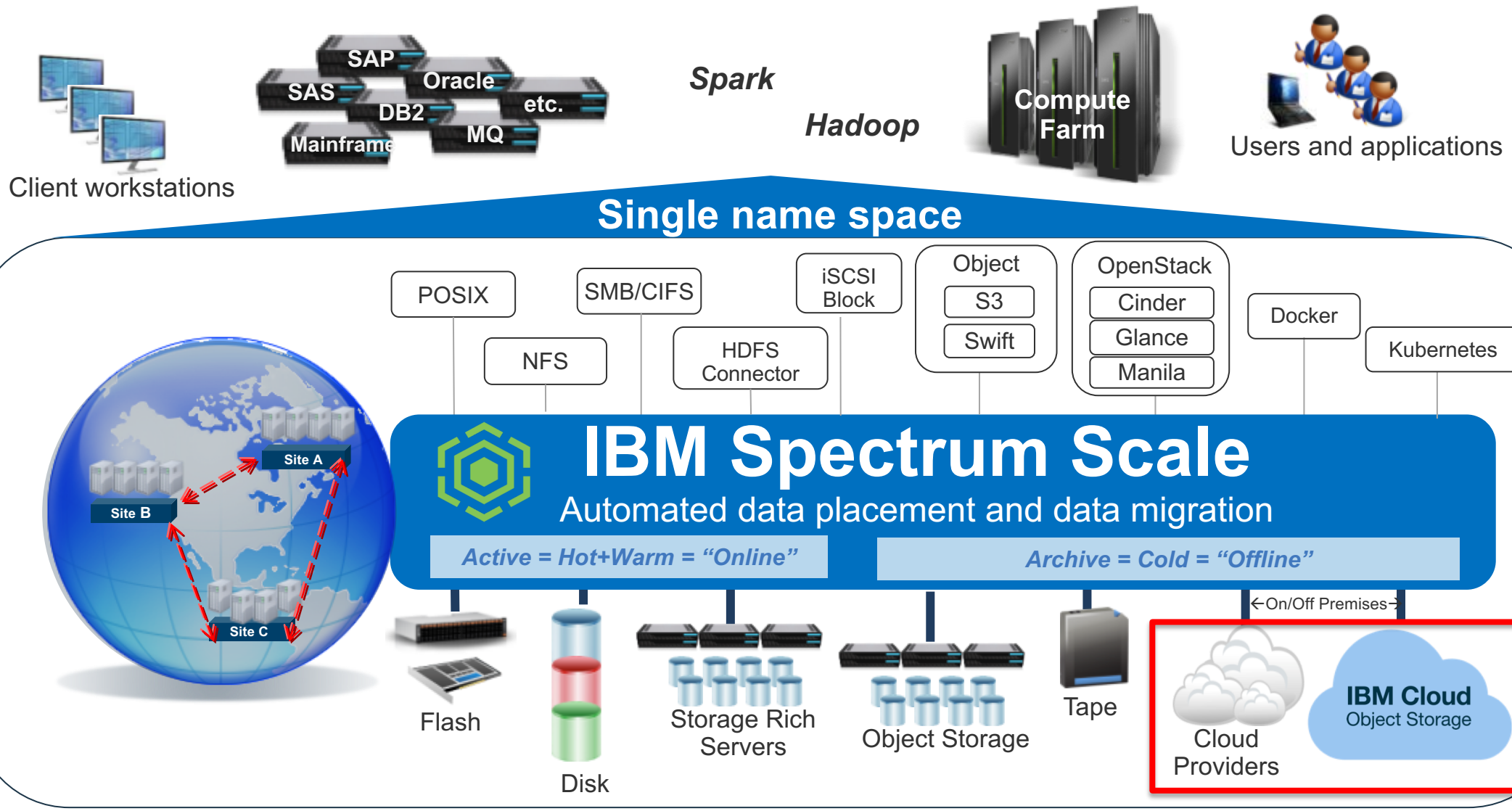
Policy driven data movement

- Managed data placement or migration of cold data



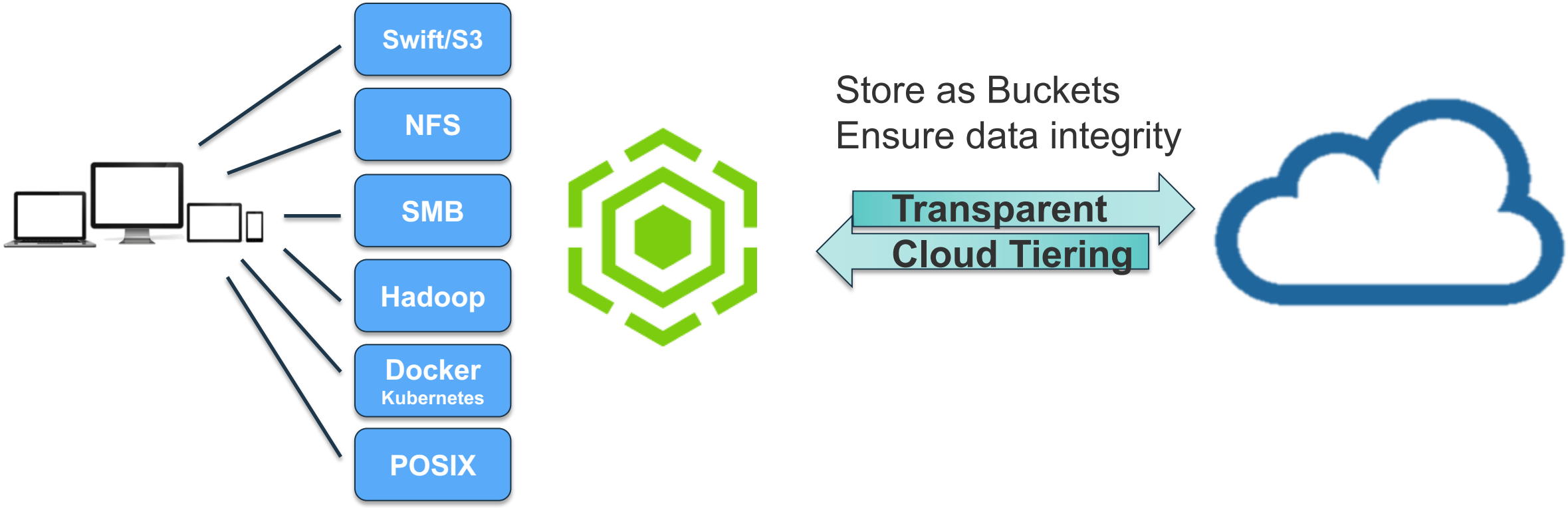


IBM Spectrum Scale – scalable file and object store



Spectrum Scale Transparent Cloud Tiering Introduction

Spectrum Scale Transparent Cloud Tiering (TCT)



Transparent Data

Metadata is managed by Spectrum Scale

Cloud appears as external storage pool
Auto-tiering & migration

IBM

Protocol mapping vs. Transparent Cloud Tiering

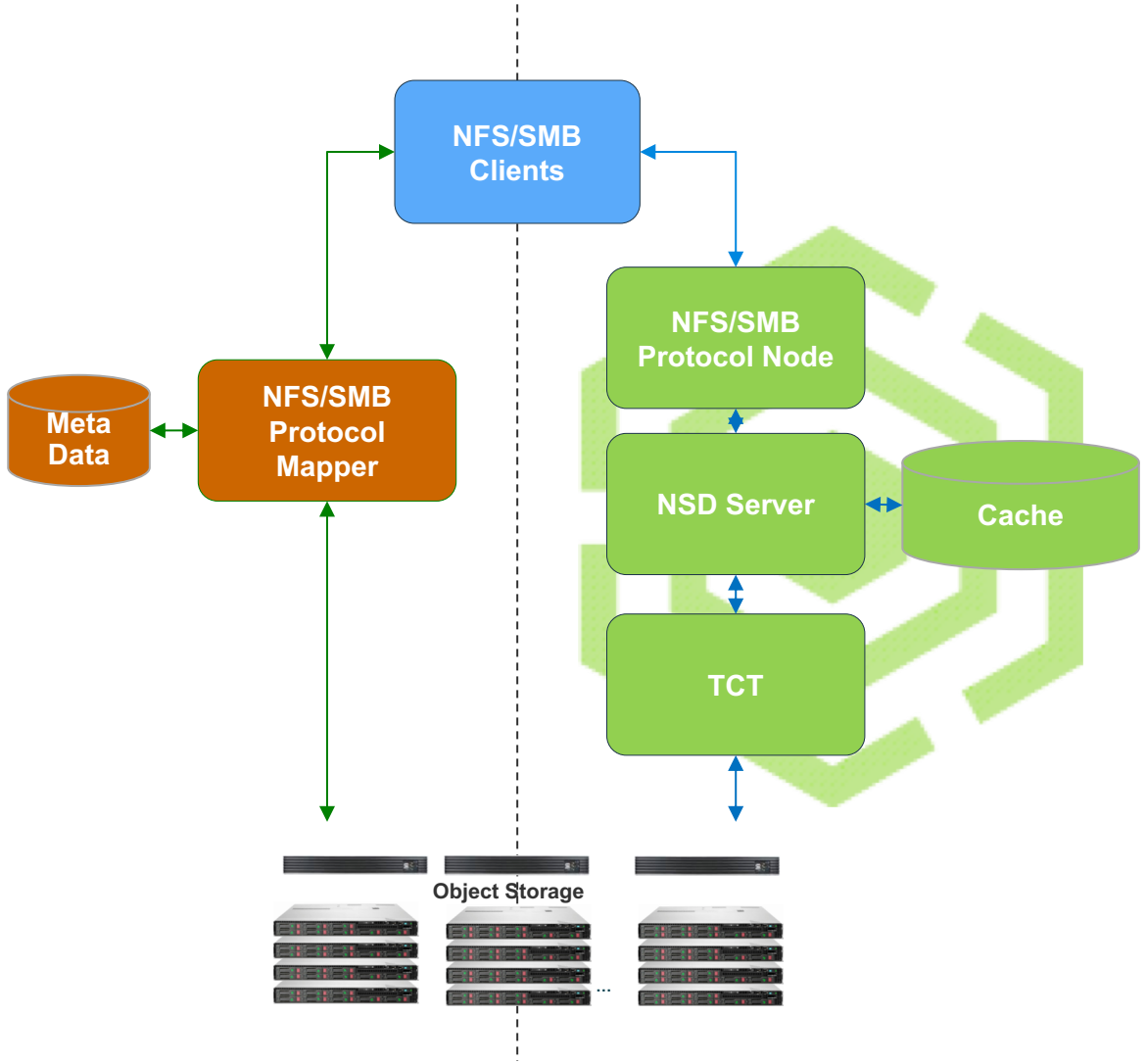
Protocol Mapping

Pros

- Simple Architecture
- No cache storage

Cons

- All writes&reads have to be mapped
- No cache for high-performance read/write
- Often vendor mix required



Transparent Cloud Tiering

Pros

- High performance write operations
- High performance read for cached data
- Public cloud-ready

Cons

- More complex architecture
- Cache storage required
- Increased response time for non-cached data

IBM

Spectrum Scale TCT value

- Extend Spectrum Scale to private or public cloud
  - Open object store interfaces drive new Data Economics
  - Public Cloud:
    - IBM Cloud Object Storage
    - IBM Bluemix Object Storage
    - Amazon S3
  - Private Cloud or on-premises:
    - IBM Cloud Object Storage
    - OpenStack Swift
- Transparent to end-users using Spectrum Scale
- Enhancing Object Storage choice
  - Spectrum Scale as High-Performance, unified file&object
  - IBM Cloud Object Storage as cost-optimized, scalable object storage

Recommended Environment for Spectrum Scale TCT

- Requires Spectrum Scale Advanced / Data Management
- Available upon request, prerequisites are
  - Appropriate file sizes (>1MB, the larger, the better)
  - Stable data for long-term retention
    - Files likely to be updated for deleted (reads are ok)
  - Sufficient Object Storage capacity
  - Appropriate network connectivity to cloud storage, load balancing

Spectrum Scale TCT capabilities (4.2.1)

- Up to 4 TCT nodes as part of Cloud Service node class
  - Simultaneous data transfer
  - High-availability
- Encryption by default
  - Local key store and ISKLM support
- Policy-based data movement and lifecycle
- Transparent recall
- Built-in integrity checking

Spectrum Scale TCT capabilities (4.2.2)

- Support for up to 4 node classes for additional scaling
- Cloud Data sharing
  - Data import/export to/from cloud, preserves data and allows native access on both ends
  - Export can be policy-based, import requires object list
- Windows explorer thumbnail support
  - Keep portion of migrated data to support thumbnails
- Spectrum Scale GUI integration
  - Health state
  - High-level operations/throughput monitoring

Spectrum Scale TCT capabilities (4.2.3)

- WAN Proxy support
  - For outbound connections in proxy-controlled environments
- Premigration
  - Migrated files stay resident
  - Quick migrate when required
  - Fast access to small files
- WORM / SnapLock Support
  - Using Spectrum Scale Immutability and IBM Cloud Object Storage „Locked Vault“ features
  - Provides SEC-17a compliant WORM storage solution leveraging on-premises Cloud Object Storage

Spectrum Scale TCT capabilities (5.0.0)

- **File set support**
  - File set support allows for more granular management of data movement to object storage

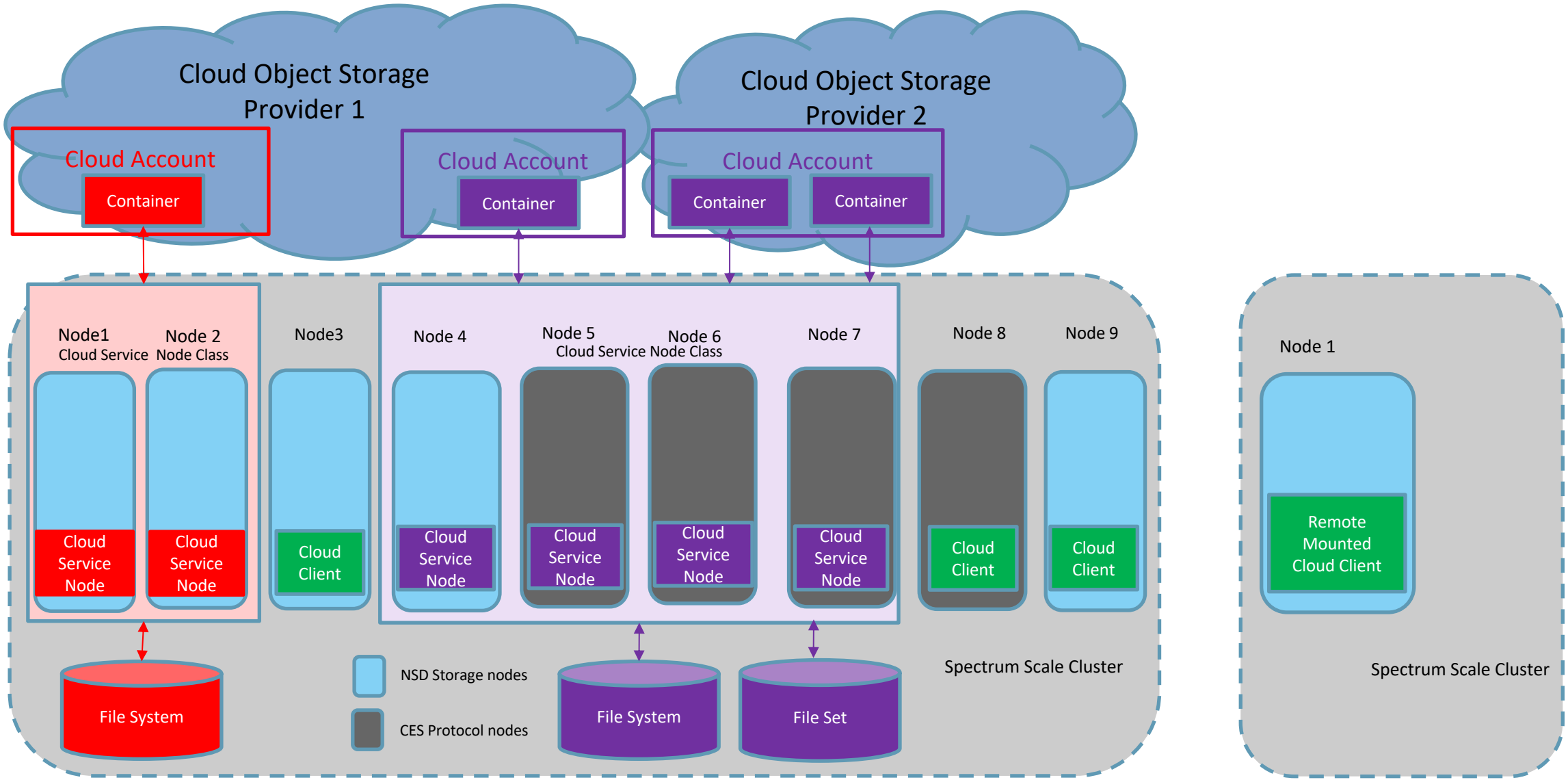
- **Multiple file system support**
  - The limit of 1 file system per Cloud Service Node group has been lifted
  - Scaling target: 128 file sets/file systems per Cloud Services Node Group

- **Multiple cloud accounts support**
  - Each filesystem or file set can support up to two Cloud Accounts

- **High Scalability using Container spill-over**
  - better support of large filesystems or file sets.
  - Scaling target: 200 containers per node group with each container spilling over at 100 million files

- **Remote mounted Client Support**
  - Remotely mounted clients can use TCT services on multiple remote clusters for transparent recall
  - More flexible ESS (separate cluster) support

Spectrum Scale 5.0.0 Multi-Cloud TCT

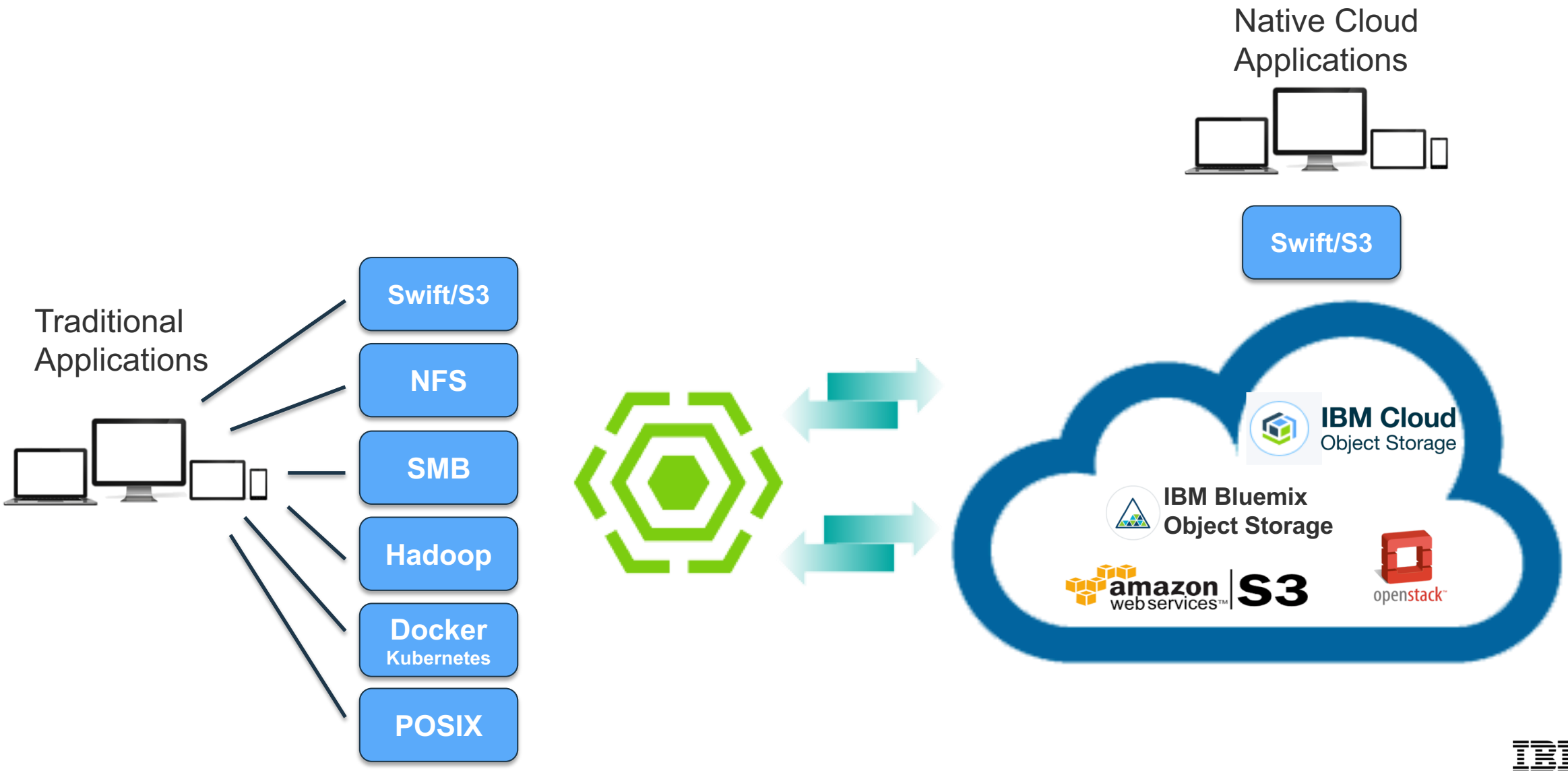


Spectrum Scale TCT capabilities (5.0.1)

- **Greatly improved large file latency performance**
  - Expanded parallel threading to cover multi-threading within large individual files
- **Automated maintenance scheduling**
  - Most background maintenance activities are now automated
- **Scalable TCT DR service restoration**
  - SOBAR based back-ups allow restore service at around 1 billion files/day
  - Background automation to be added, outlook is next Scale release
- **Ongoing service and support improvements**
  - Added latency metrics
  - No retries on errors that won't recover
  - Improved error messaging
  - Bottleneck detection

Cloud Data Sharing

Cloud Data Sharing



IBM Spectrum Scale Cloud Data Sharing (cont.)

- IBM Cloud Object Storage serves as
  - Central data repository
  - Data distribution point
  - Data is accessible for native cloud applications through IBM COS S3 API
- Sharing can occur between object storage and many Spectrum Scale Clusters
- Shared Data may originate in either the cloud or Spectrum Scale
- Data may be moved or copied between Spectrum Scale clusters and object storage
- Efficient sharing mode: Special stubbed move available, so metdata only is stored on Spectrum Scale with data imported transparently when a file is requested
- One or more Spectrum Scale clusters can push to the cloud
- Spectrum Scale can pull data on demand

Real-world Use Cases

Autonomous driving measurement data

Goal:
Effectively storing and accessing huge amounts of unstructured measurement data and videos

Requirements:
NFS native data access or NFS via appliance

HTTP RESTful S3 API and SWIFT API access

20-40TB ingest per day

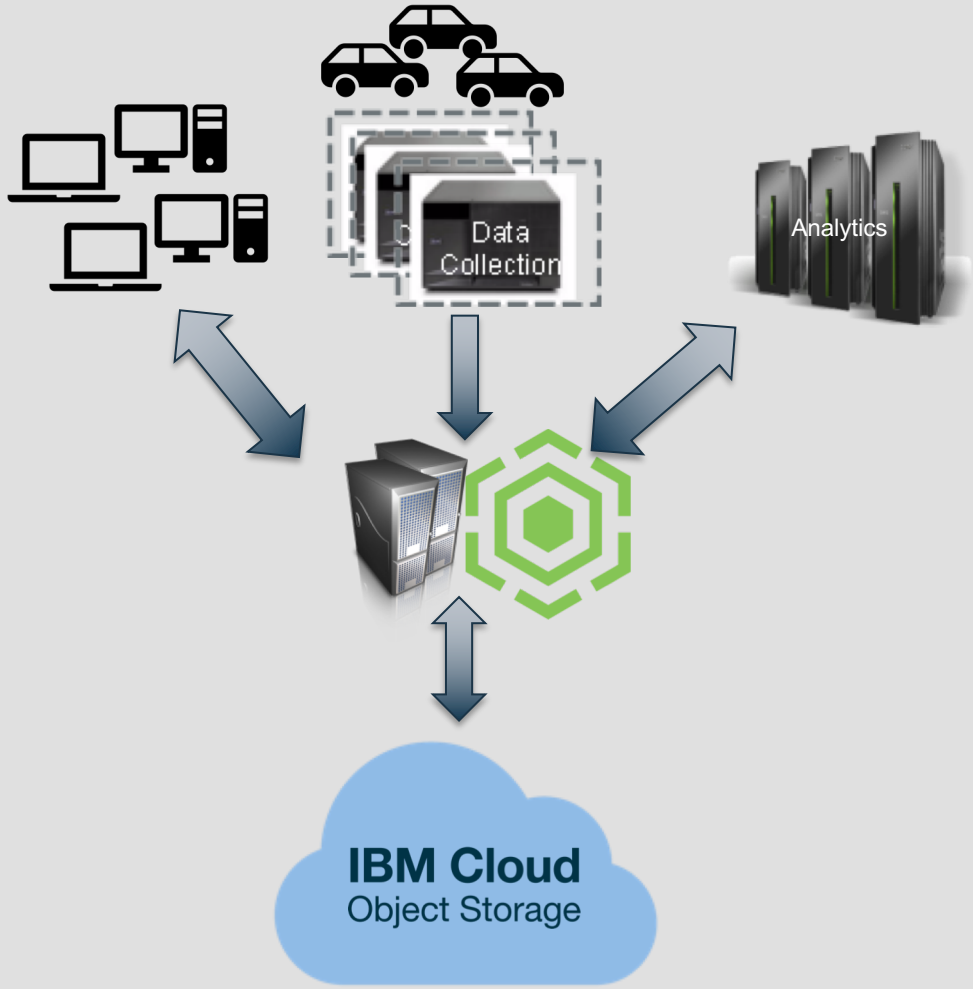
> 100 Million Objects, 10MB average object size

> 1000 Clients or client groups

Multi-tenancy, authorization and authentication

Encryption
- Data-in-transition encryption (at least HTTPS)
- Data-at-rest encryption (encrypted disks)

Interface to Analytics



Autonomous driving measurement data (cont.)

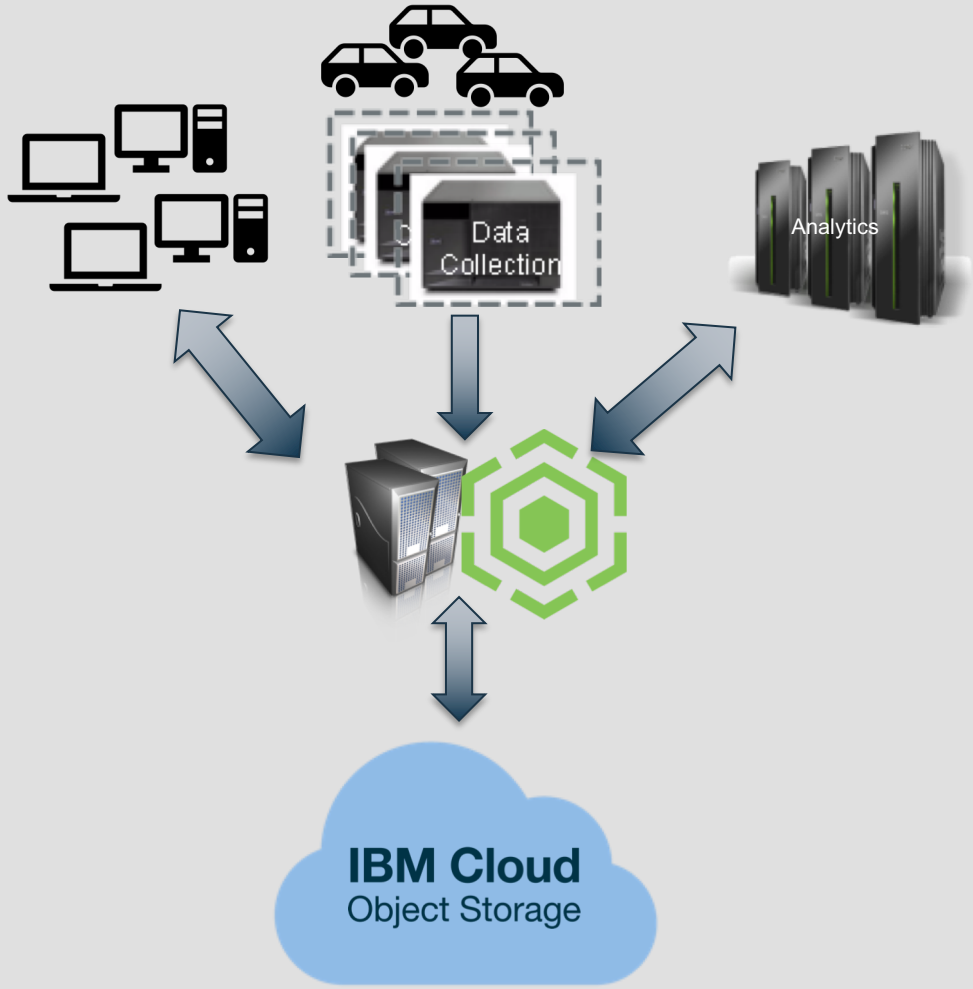
Insights:

Required 800MB/s write throughput can be delivered by using ESS GL4

End-to-end encryption using HTTPS, Spectrum Scale at-rest and TCT Encryption with ISKLM, potentially NAS over IPSEC/VPN

Tenant isolation through multiple TCT node sets that map 1:1 to separate file systems

Separate NAS and TCT nodes for optimized resource usage



Bodycam data

Goal:

Providing archive for bodycam data with file access (unstructured, video)

Requirements:

8-hour shift generates approx. 20GB of video data per camera

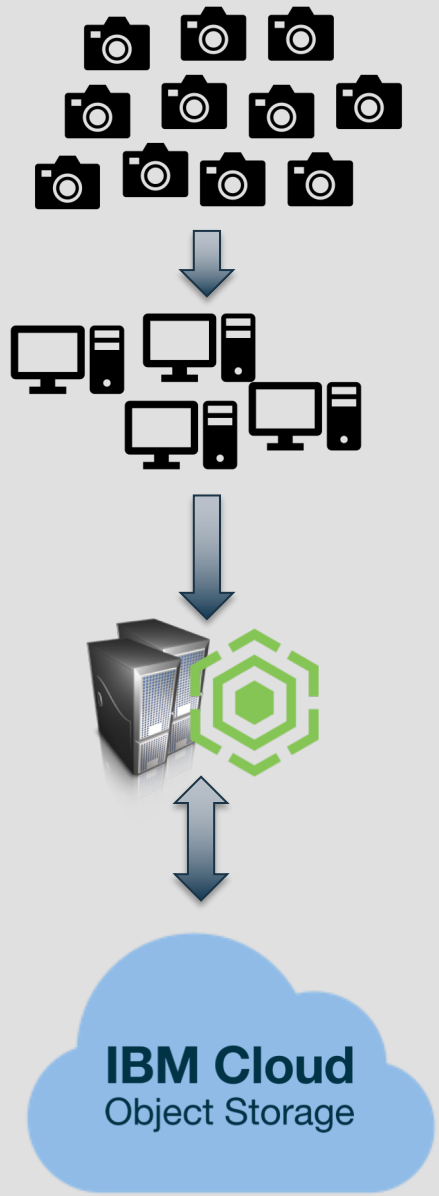
Legal requirement to keep the data for years
- Varies by country/case

NAS protocols for data ingest

Secure authorization and authentication, encryption

Insights:

Spectrum Scale stretched cluster combined with multi-site on-premises COS turned out to be perfect match



Backup and file access private Cloud

Goal:

Transition to backup and file access cloud environment

Requirements:

Move away from traditional backup
- expensive, maintenance-intensive and slow for recovery operations according to the client

Standardize and harmonize NAS and backup services

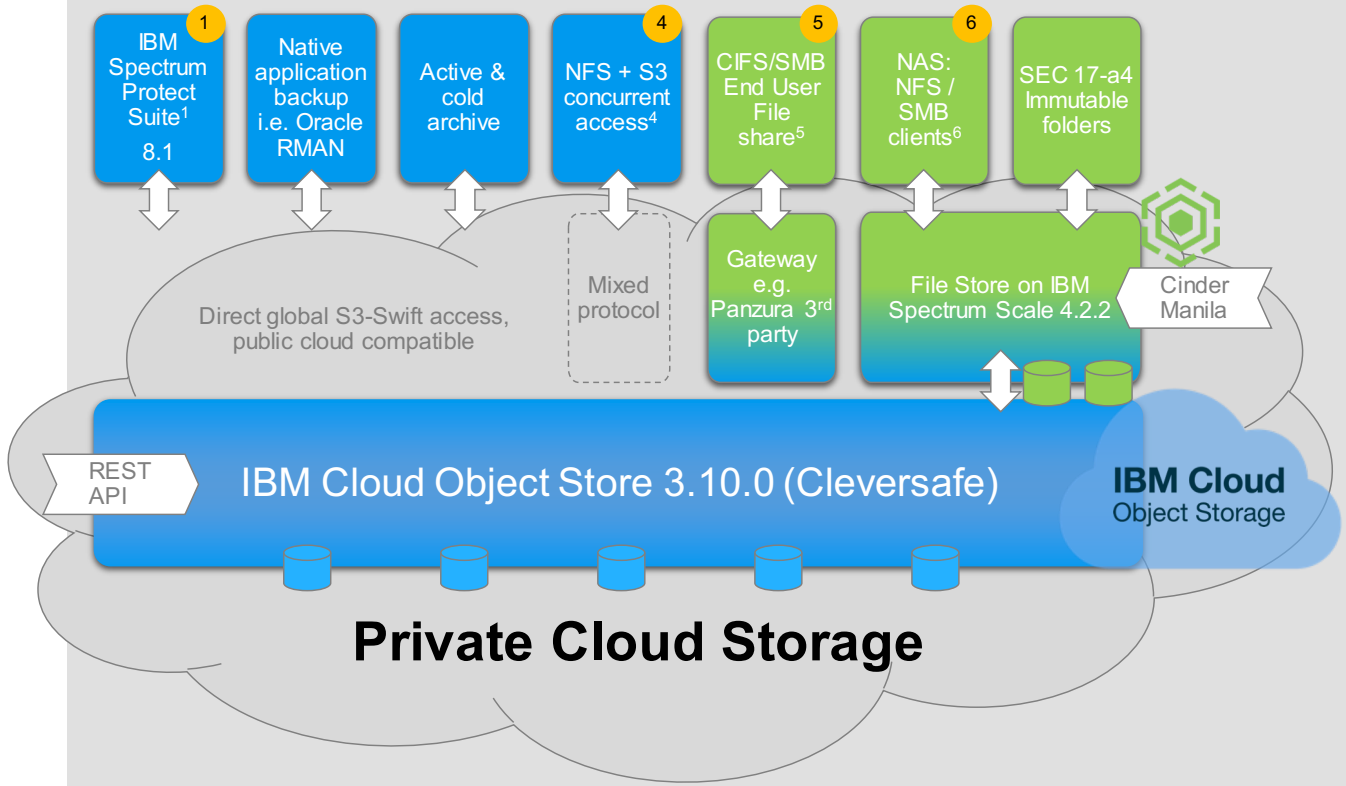
Immutability for compliance requirements

Option for native Object protocol access

Capability to utilize public cloud

Insights:

Scale and COS combination strongly supports client roadmap towards a storage cloud



StaaS provider leveraging public cloud

Goal:

Leverage IBM COS as archival tier behind Spectrum Scale as part of Storage-as-a-Service offering

Requirements:

Capability to utilize public cloud

Insights:

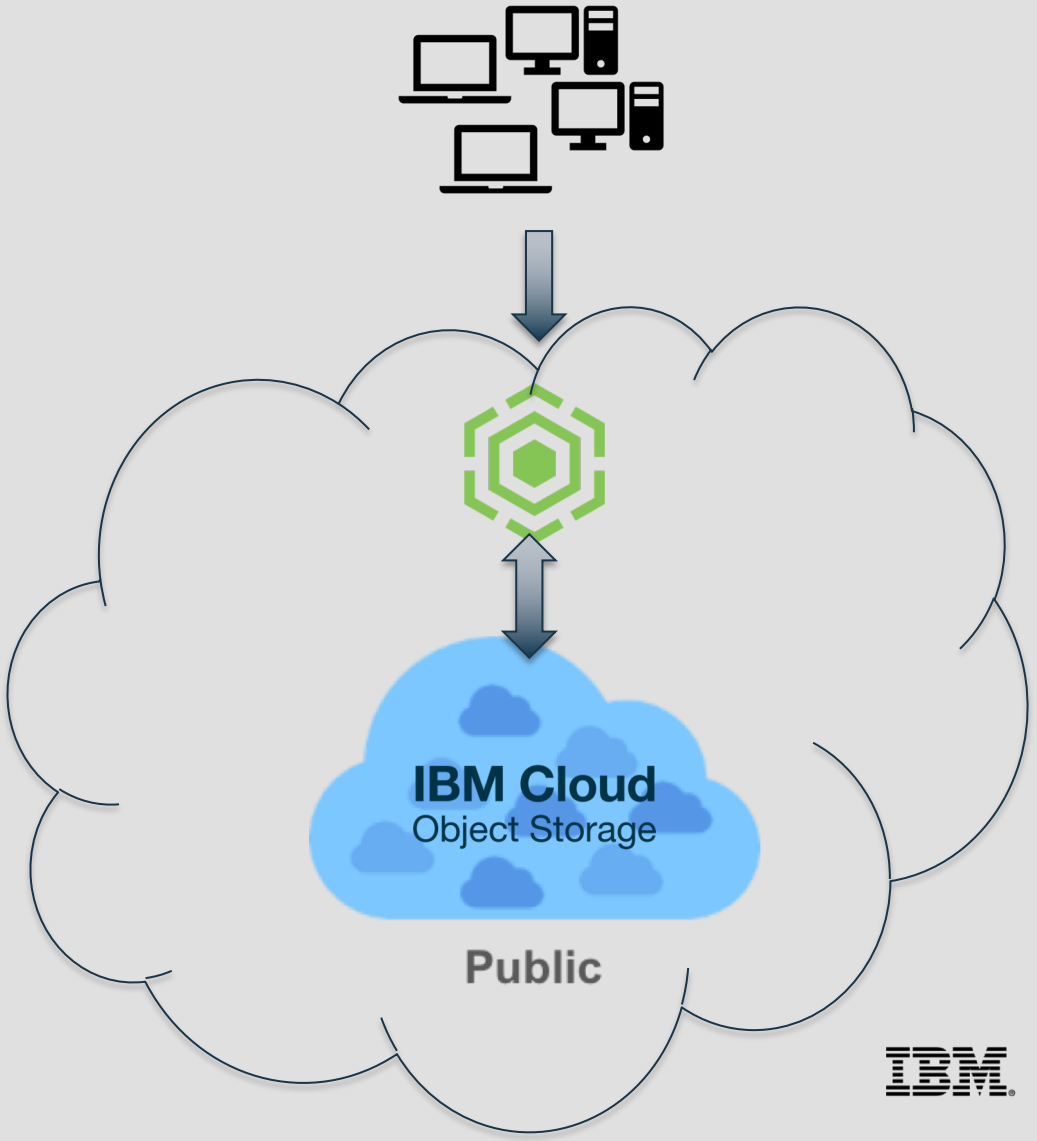
Apart from public cloud GB/month storage cost, need to take transfer costs into account

- COS on IBM Cloud: ~ 1ct per 1000 PUTs, similar per 10000 GETs

Every migrate/recall causes at least two PUT/GET requests (data and metadata)

Migrate uses multi-part upload (defaults to 100MB part size)

- Providers treat a single part as “single PUT”



Log File Archiving

Goal:

Archive log data, freeing up space on the primary Spectrum Scale Tier

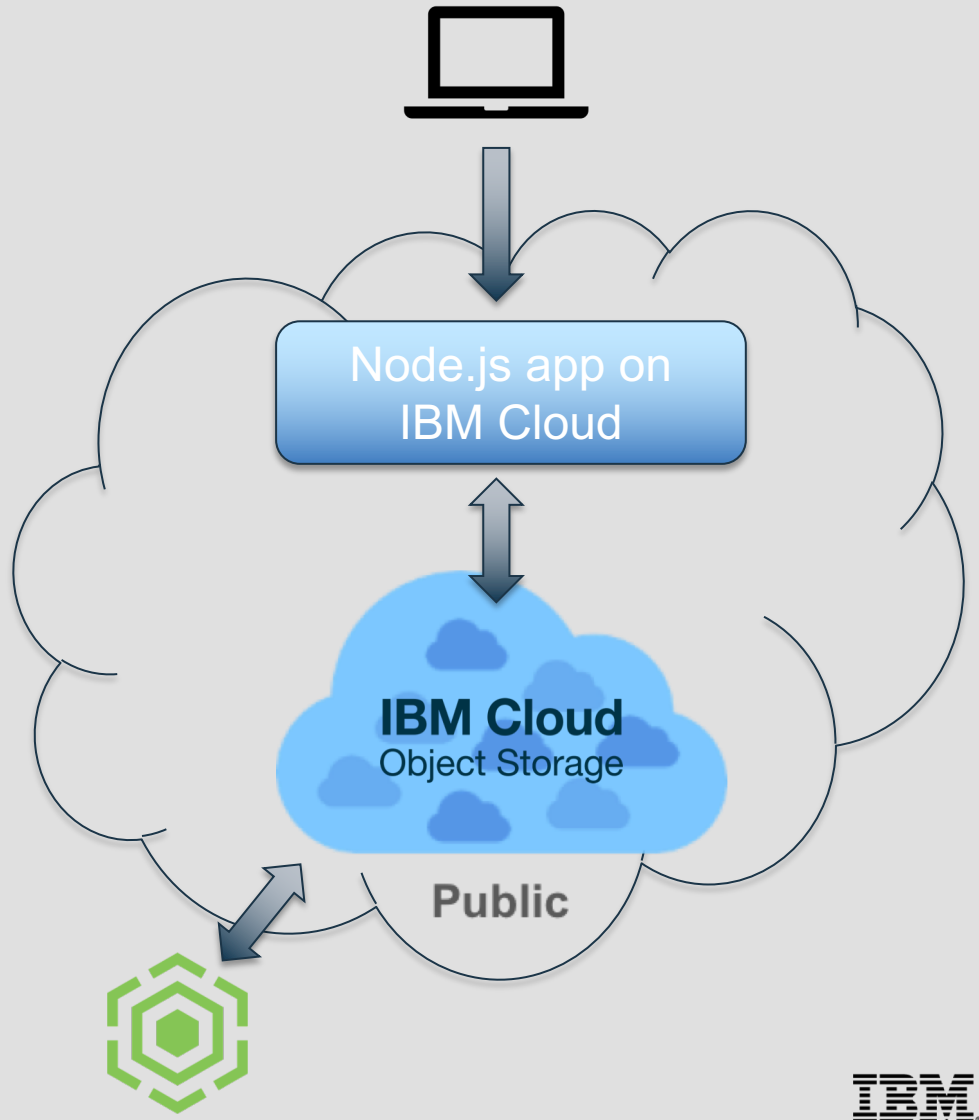
Requirements:

No stub files on Spectrum Scale namespace → Cloud Tiering is not an option

Insights:

Export log data through Cloud Data Sharing

Re-import for analysis or provide a cloud application for data access



Video surveillance „anti-use-case“

Goal:

Build System fast enough to handle numerous parallel data streams while cost-efficient to keep data for some period

Requirements:

NFS ingest

600 cameras, 9Gbps sustained ingest

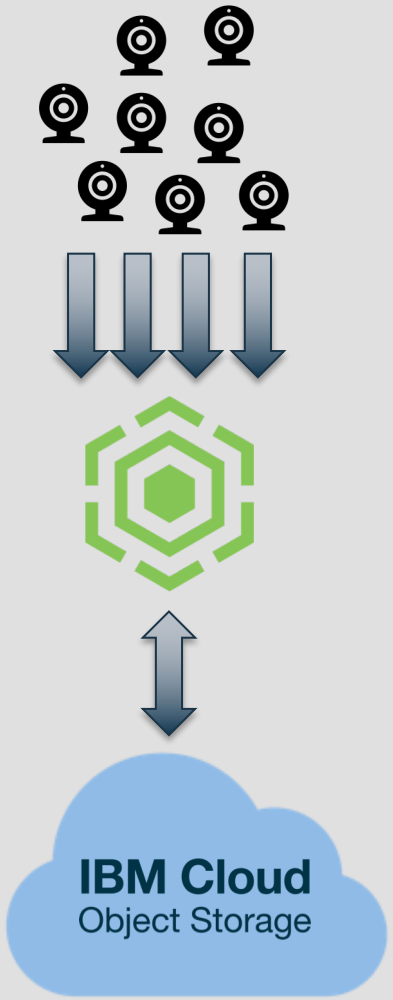
Average file size 5MB

Insights:

With 2PB COS, data can be kept for 23 days

Reconcile deletions performed using single node/thread only
- as of Spectrum Scale 4.2.2

80TB / 15M files need to be reconciled every day
- This is the “showstopper” → recommended ESS-only solution

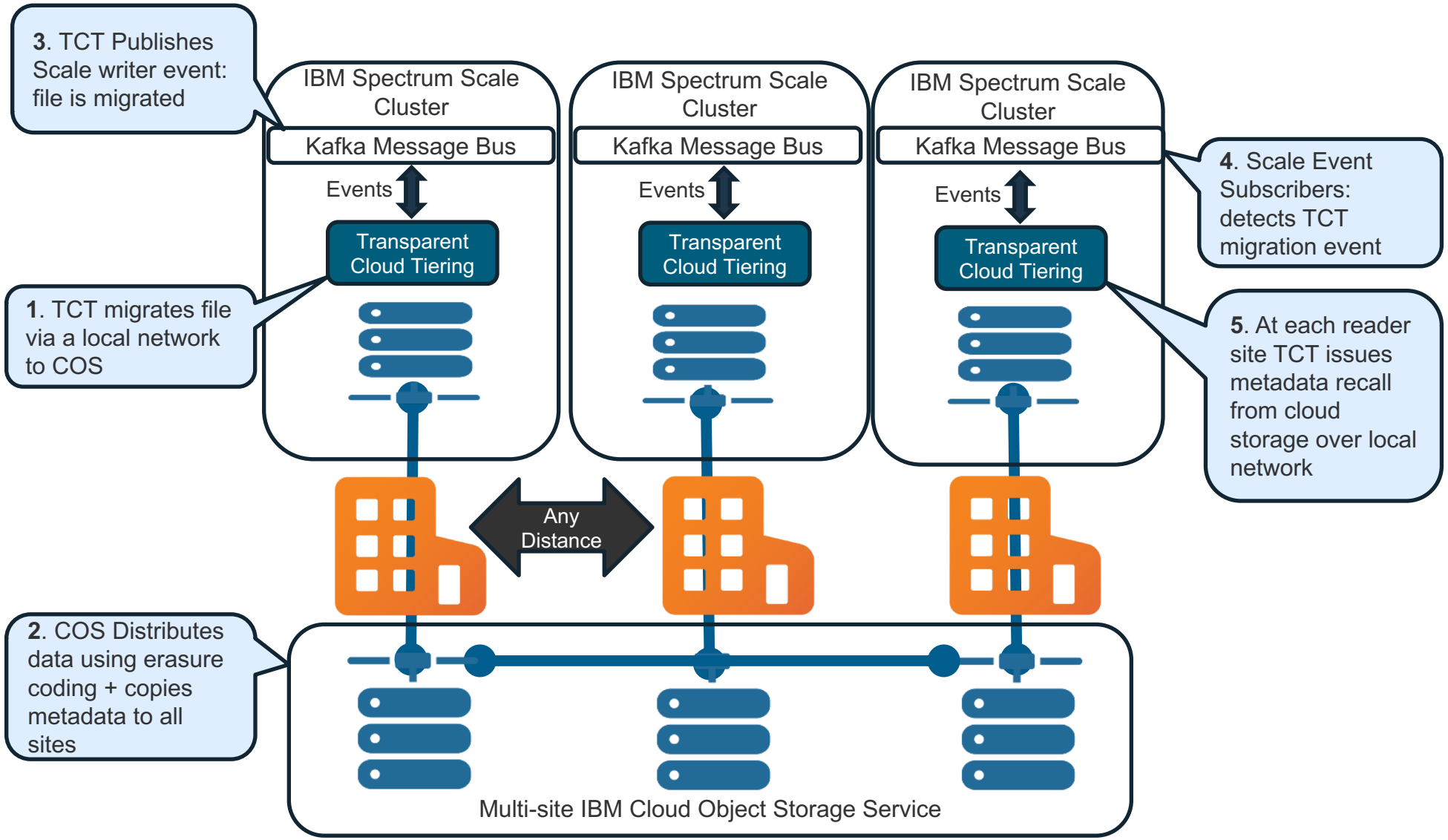


Legal Disclaimer

IBM's statements regarding its plans, directions, and intent are subject to change or withdrawal without notice at IBM's sole discretion. Information regarding potential future products is intended to outline our general product direction and it should not be relied on in making a purchasing decision. The information mentioned regarding potential future products is not a commitment, promise, or legal obligation to deliver any material, code or functionality. Information about potential future products may not be incorporated into any contract. The development, release, and timing of any future features or functionality described for our products remains at our sole discretion.

Upcoming TCT Feature: Support for Multi-site use case with TCT Multi-cluster support

Multi-cluster Support Diagram



IBM

TCT Multi-cluster Tiering Service Design Overview

Tiering Data Overview – TCT data tiered to object storage in one cluster is available for use on other clusters that could be at the same site or a site at any distance that has access to the shared object storage

- **Sharing Paradigm:** One cluster as writer, other clusters as readers all keep a copy of directory and file meta-data / namespace
  - Sharing is enforced by TCT and can also be enforced by vault/container ACLs on object storage
  - Generally, readers pull in the data only on demand
  - Multi-cluster writes possible with application level coordination
- **Sharing Granularity: I**s at the level of a Cloud Service (file set granularity coupled with an associated migration ILM policy)
- **Sharing Consistency**: Name space consistency is gated primarily by inter-cluster network speed and reliability
  - Directory consistency is based on file path names and parent directory ACLs so file and directory moves are not well supported
- **Sharing Security:** Inter-cluster artifacts are encrypted – events, meta-data, data (inter-cluster federated SKLMs recommended)
- **Adding readers and writers:** Allows adding new readers to a service or a new writer service at any time
  - This feature to add a reader also supports synching up a reader that has been down for a long time (past the event horizon)
- **Inter-cluster / cluster down failure tolerance:** Routine recovery for down times of a few days (Kafka event horizon time)
- **Disaster Recovery:** Fail-over support allows a reader site to take over as a writer
  - There is no failback mechanism but the original writer can be made a reader and then again take over as writer via a 2nd failover
- **Management Paradigm:** All configuration, management, and monitoring is by cluster – no centralized source

TCT Multi-cluster Tiering Service Performance and Scaling

- **Sharing Scale:** Sharing between up to 12 clusters, with tiering file limits per cluster no different than single site tiering
- **Writer Sharing performance:** Writer performance is gated by current migration performance (no special gates for multi-site)
- **Reader Sharing performance:** Reader performance roughly on par with writer performance (order of 1000 FPS / Node Class)
- **Inter-cluster / cluster down failure recovery performance:** Recovery rate on par with reader/writer performance
- **Reader Initialization / Synchronization Performance:** Initialization and recovery from outages past the event horizon happens at a faster rate than reader/writer performance (target is 5-10 times faster) since the inodes are packaged together from the writer and sent directly -- no need to go through object storage
- **Disaster Recovery:** Fail-over support is manual and takes a few minutes

TCT Multi-cluster Sharing Service Overview

This document describes support for TCT multi-cluster Sharing in Spectrum Scale

Sharing Data Overview – TCT data shared with object storage in one cluster is available for use on other clusters

- **Sharing Paradigm:** One cluster as writer, others as readers all must keep a common directory path and the file meta-data
  - Sharing for Object Storage originated data is manual to all Scale readers (no object storage event support at this time)
  - Sharing is enforced by TCT and can also be enforced by vault/container ACLs on object storage
  - Generally, readers pull in the data only on demand
  - Multi-cluster writes possible with application level coordination
- **Sharing Granularity:** For Scale originated data sharing granularity is at the level of a Cloud Service (file set granularity coupled with an associated export ILM policy)
  - Object Storage originated data s imported manually so can be at any desired level of granularity
- **Sharing Consistency**: Scale originated data has event-driven name space consistency and is gated primarily by inter-cluster network speed and reliability (object storage originated data is kept consistent manually)
- **Sharing Security:** Inter-cluster events are encrypted (inter-cluster federated SKLMs recommended)
  - file meta-data, data are encrypted if sharing Scale-to-Scale – wire encryption only if sharing from object storage
- **Adding readers and writers:** Allows adding new readers to a service or a new writer service at any time
  - Manifest files can be used to synchronize new readers with earlier exported data
- **Inter-cluster / cluster down failure tolerance:** Routine recovery for down times of a few days (Kafka event horizon time)
- **Disaster Recovery:** Fail-over support to allows a reader site to take over as a writer
- **Management Paradigm:** All configuration, management, and monitoring is by cluster – no centralized configuration source

IBM

TCT Multi-cluster Sharing Service Performance and Scaling

- **Sharing Scale:** Sharing between up to 12 clusters, with sharing file limits per cluster no different than single site sharing
- **Writer Sharing performance:** Writer performance is gated by current export performance (no special gates for multi-site)
- **Reader Sharing performance:** Reader performance roughly on par with writer performance (order of 1000 FPS / Node Class)
- **Inter-cluster / cluster down failure recovery performance:** Recovery rate on par with reader/writer performance
- **Reader Initialization / Synchronization Performance:** Initialization and recovery past the event horizon depend on the manifest with manifest import performance is on part with reader and writer performance
- **Disaster Recovery:** Fail-over support is manual and takes a few minutes

Summary and call to action

Consider Spectrum Scale Transparent Cloud Tiering for matching use cases

- Active Archive
- Analytics

Extend Spectrum Scale to public cloud or IBM Cloud Object Storage System

- Transparent to end-users using Spectrum Scale
- Enhancing Object Storage choice
  - Spectrum Scale as High-Performance, unified file and object storage
  - IBM Cloud Object Storage as cost-optimized, scalable object storage

We need your feedback for the multi-site support!

Thank you!

Notice and disclaimers

- **U.S. Government Users Restricted Rights — use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM.**

- Information in these presentations (including information relating to products that have not yet been announced by IBM) has been reviewed for accuracy as of the date of initial publication and could include unintentional technical or typographical errors. IBM shall have no responsibility to update this information. **This document is distributed "as is" without any warranty, either express or implied. In no event shall IBM be liable for any damage arising from the use of this information, including but not limited to, loss of data, business interruption, loss of profit or loss of opportunity.** IBM products and services are warranted according to the terms and conditions of the agreements under which they are provided.

- IBM products are manufactured from new parts or new and used parts. In some cases, a product may not be new and may have been previously installed. Regardless, our warranty terms apply."

- **Any statements regarding IBM's future direction, intent or product plans are subject to change or withdrawal without notice.**

- Performance data contained herein was generally obtained in a controlled, isolated environments. Customer examples are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual performance, cost, savings or other results in other operating environments may vary.

- References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business.

- Workshops, sessions and associated materials may have been prepared by independent session speakers, and do not necessarily reflect the views of IBM. All materials and discussions are provided for informational purposes only, and are neither intended to, nor shall constitute legal or other guidance or advice to any individual participant or their specific situation.

- It is the customer's responsibility to insure its own compliance with legal requirements and to obtain advice of competent legal counsel as to the identification and interpretation of any relevant laws and regulatory requirements that may affect the customer's business and any actions the customer may need to take to comply with such laws. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the customer is in compliance with any law.

Notice and disclaimers (continued)

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products in connection with this publication and cannot confirm the accuracy of 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. IBM does not warrant the quality of any third-party products, or the ability of any such third-party products to interoperate with IBM's products. **IBM expressly disclaims all warranties, expressed or implied, including but not limited to, the implied warranties of merchantability and fitness for a particular**, **purpose.**

The provision of the information contained herein is not intended to, and does not, grant any right or license under any IBM patents, copyrights, trademarks or other intellectual property right.

IBM, the IBM logo, ibm.com, AIX, BigInsights, Bluemix, CICS, Easy Tier, FlashCopy, FlashSystem, GDPS, GPFS, Guardium, HyperSwap, IBM Cloud Managed Services, IBM Elastic Storage, IBM FlashCore, IBM FlashSystem, IBM MobileFirst, IBM Power Systems, IBM PureSystems, IBM Spectrum, IBM Spectrum Accelerate, IBM Spectrum Archive, IBM Spectrum Control, IBM Spectrum Protect, IBM Spectrum Scale, IBM Spectrum Storage, IBM Spectrum Virtualize, IBM Watson, IBM z Systems, IBM z13, IMS, InfoSphere, Linear Tape File System, OMEGAMON, OpenPower, Parallel Sysplex, Power, POWER, POWER4, POWER7, POWER8, Power Series, Power Systems, Power Systems Software, PowerHA, PowerLinux, PowerVM, PureApplica- tion, RACF, Real-time Compression, Redbooks, RMF, SPSS, Storwize, Symphony, SystemMirror, System Storage, Tivoli, WebSphere, XIV, z Systems, z/OS, z/VM, z/VSE, zEnterprise and zSecure are trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at: www.ibm.com/legal/copytrade.shtml.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both. Java and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

Session reference links

Additional material

Redbook „Cloud Object Storage as a Service“
http://www.redbooks.ibm.com/redbooks.nsf/redbookabstracts/sg248385.html?Open

Redpaper „Enabling Hybrid Cloud Storage for IBM Spectrum Scale Using Transparent Cloud Tiering“
http://www.redbooks.ibm.com/redpieces/abstracts/redp5411.html

Redpaper „Cloud Data Sharing with IBM Spectrum Scale
http://www.redbooks.ibm.com/redpieces/abstracts/redp5419.html

Try yourself

Get free IBM Cloud Object Storage trial
https://www.ibm.com/cloud-computing/bluemix/cloud-object-storage

Get free IBM Spectrum Scale trial
http://www-03.ibm.com/systems/storage/spectrum/scale/trial.html