IBM Global Data Platform
Data Tiering with IBM Storage Scale
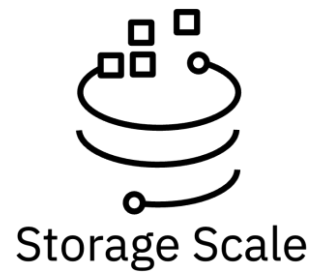
IBM

Max Huber

Storage Technical Specialist – Software-Defined Storage
Max.Huber@ibm.com
+49-172-2594998

Disclaimer

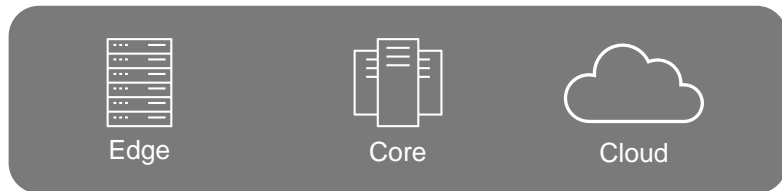# A Global Data Platform for Scale-Out File & Object Data

What's driving the need for a global data platform?

Data centric application development

- Example: The emergence of AI/ML use cases
  - Data hungry apps and GPUs, need access to more data, faster
  - As new applications and use cases roll out, data silos occur
  - Need unified and consistent approach to accessing data throughout AI/ML Pipeline – in both native object and file storage repositories

- Data fabric initiatives with requirements to provide consistent services across diverse infrastructure

The diverse IT infrastructure options available

- Many choices, From edge to core data center to public cloud

- Containers to simplify hybrid cloud infrastructure choices

- Drives the need for a single source of truth across diverse infrastructure that facilitates secure access while eliminating data redundancy and inconsistencies.

HPC | AI / ML | Analytics | Enterprise | Containers | Backup / Archive

Object Storage

File Storage

Edge | Core | Cloud

IBM Storage Scale
Delivering a High-Performance Global Data Platform for Unstructured Data
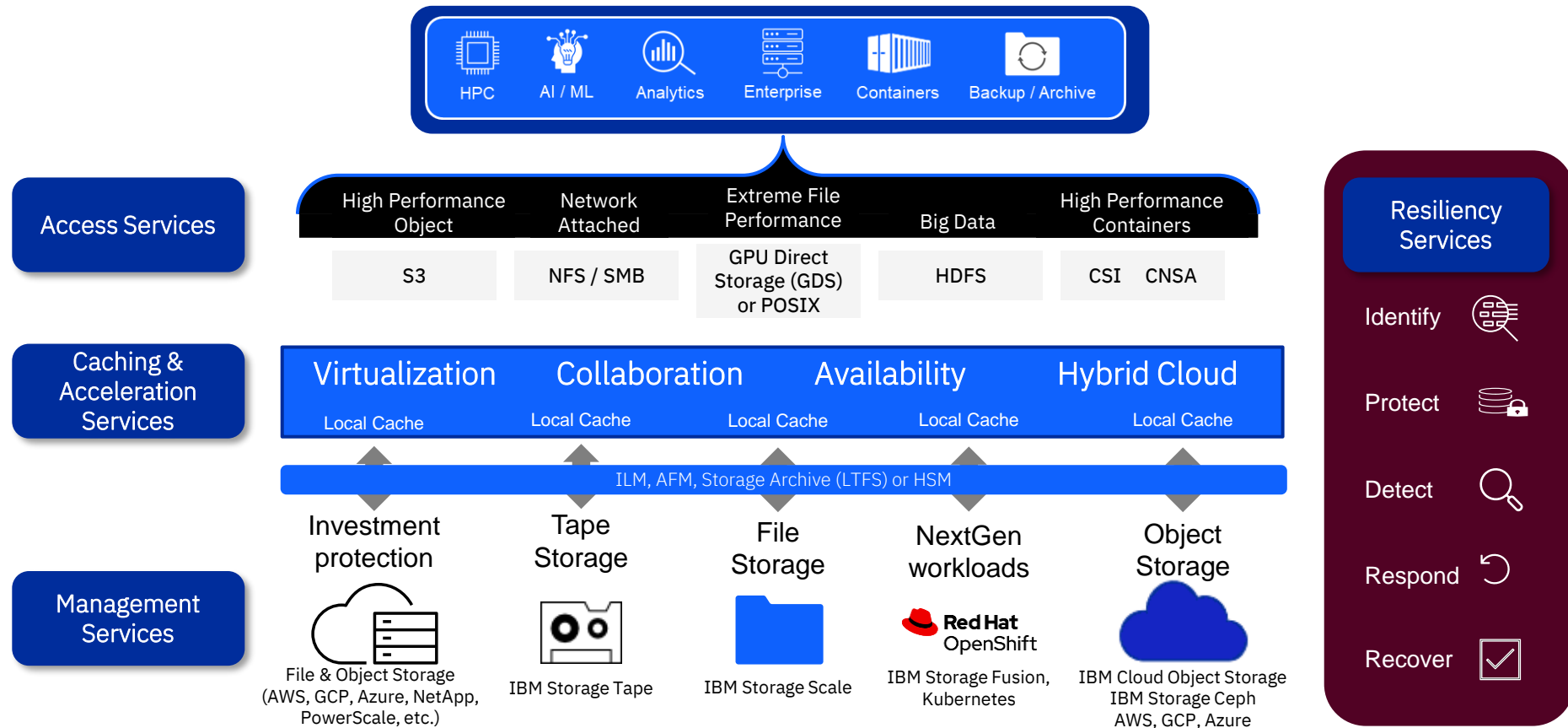
Store everywhere. Run anywhere.

- Unified Scale-out Data Lake

  - Access data using multiple protocols

  - High-performance concurrent access with integrity

  - Analytics on demand

  - Single management plane

  - Cluster replication and global namespace

  - Enterprise storage features across file, object, and HDFS

- Global collaboration with Active File Management

  - Filesystem caching and single namespace view across multiple geographically distributed remote sites

  - Extend collaborative workflows

  - Mitigate network bottleneck with advanced routing

  - Flexible configuration with writer and read-only sites

  - Disaster recovery for enterprise resiliency
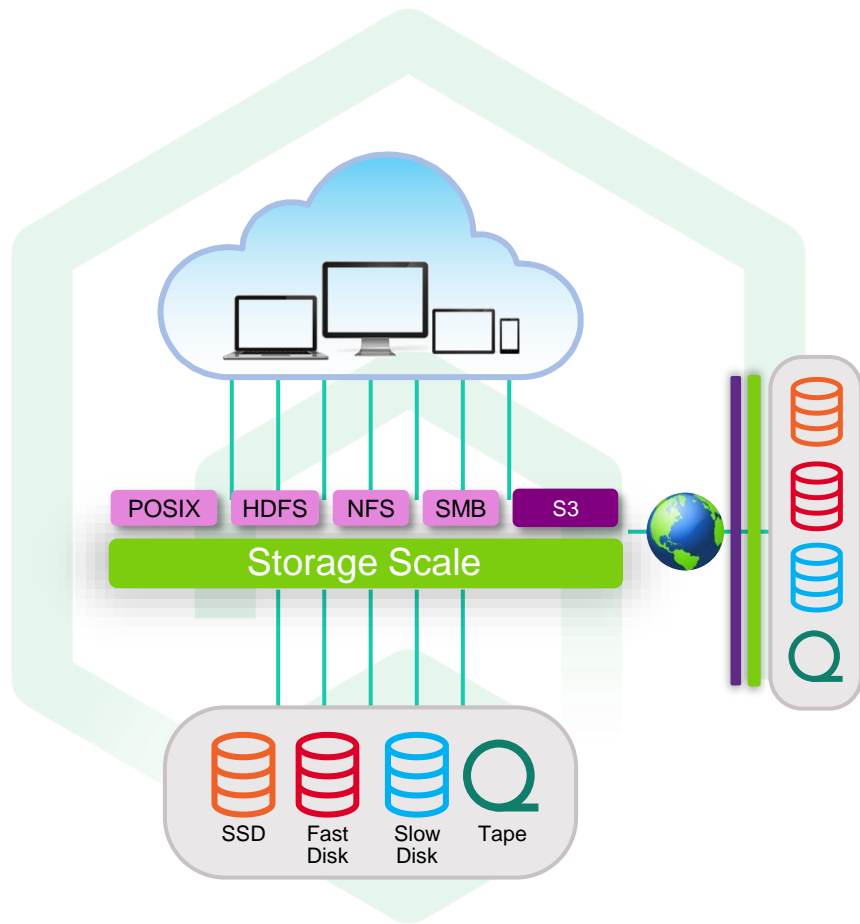
Improve data economics

- Information Lifecycle Management with intelligent automatic tiering of data

  - Policy engine for fast metadata scans

  - Automated data movement based on policy

  - Movement among multiple types of storage: Flash, SSD, HDD, external tape, object storage, and cloud

  - Reserves high speed storage for work in progress, moves everything else transparently to lower cost storage tiers

  - Visibility / access to content regardless of storage tiers

- Enterprise scale

  - Seamless expansion and upgrades

- Policy driven compression increases effective capacity
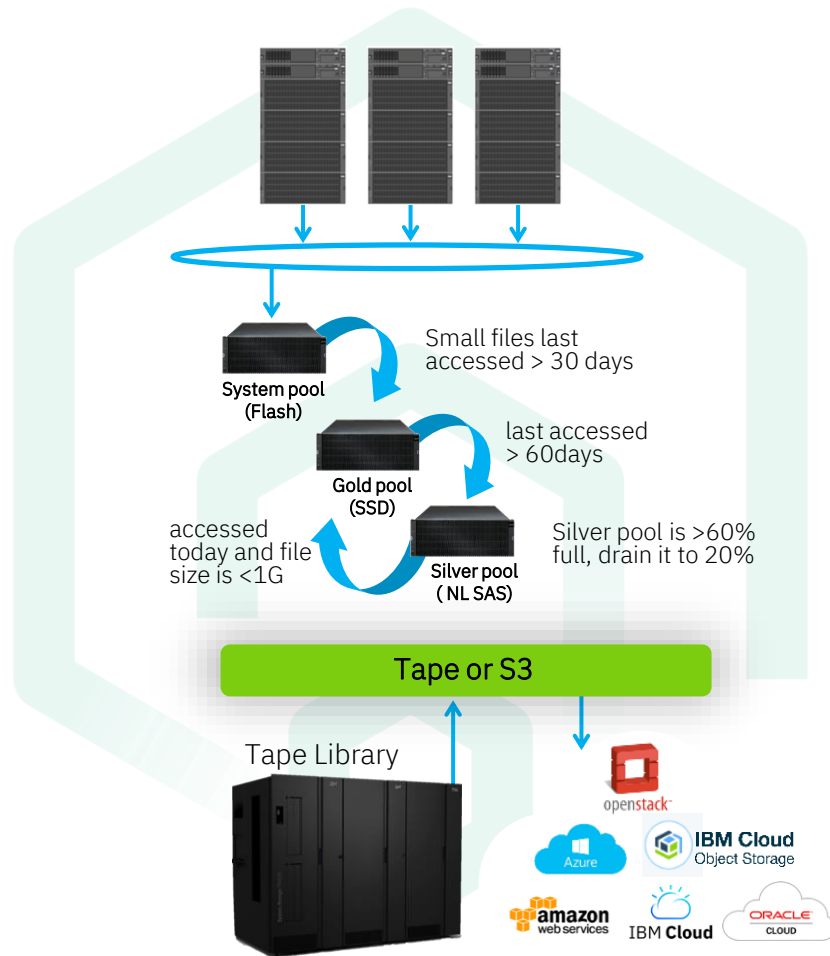
IBM Storage Scale Information Lifecycle Management

- Incorporates file and object into active capacity

- Building blocks that can expand to 8YB

- Incorporate tape or cloud into archive tier

- Use built in policy tools or granular policy engine with IBM Storage Discover for AI workflows

High performance

IBM Storage Scale performance or ESS 3500 / 6000

Scale on Cloud

Active capacity

IBM Cloud Object Storage System

Storage Scale on Cloud

IBM COS in IBM Cloud

IBM Storage Scale capacity or ESS 3500 / 6000 capacity

Archive capacity

IBM Storage Scale archive to Cloud

Tape
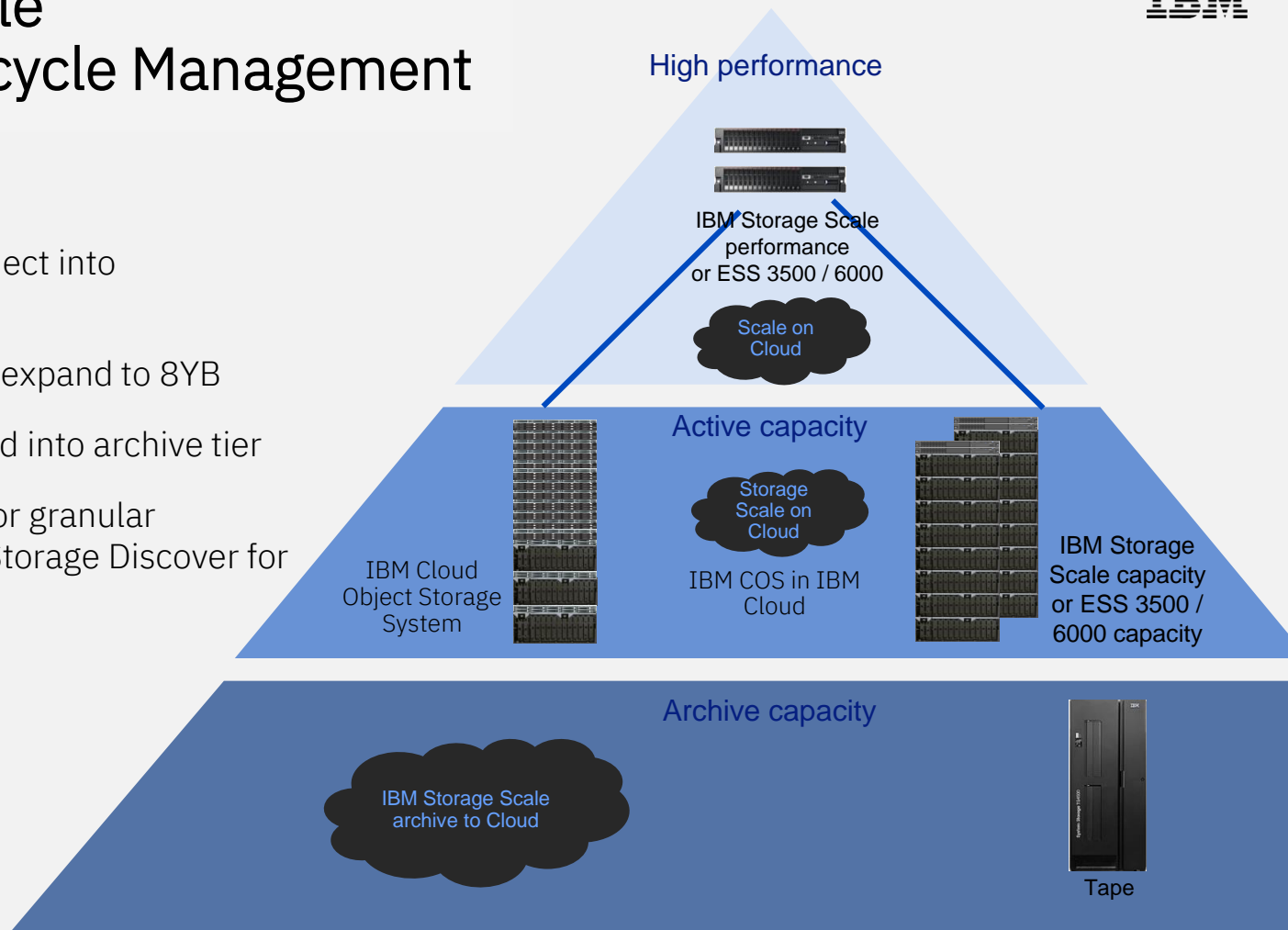
What is the Policy Engine?

Parallel high performance scan can run at > 10 million files/minute per node with low overhead

- Scan used to identify files needed to be managed or physically moved

- Traditional 'walk the directory tree' speeds are in the tens of thousands of files per minute

- That means performing management of the storage requires IO-intensive, time-intensive 'walking the directory trees' time

High performance scan engine is a unique value

- The larger the file system, the more value the Scale scan engine provides

Today's major new requirement how can we scan the file systems fast enough in order to identify files that must be:

- Migrated to another storage pool
- Migrated to external pool or cloud
- Propagated to remote sites
- Compressed
- Decompressed
- Encrypted
- Backed up
- Restored
- Deleted
- Any other storage management requirements

The three major phases of a policy run – high level overview

Phase one: Selecting candidate files

- all the files within the specified GPFS file system device, or below the input path name, are scanned

- The attributes of each file are read from the file's GPFS inode structure

Phase two: Choosing and scheduling files

- some or all of the candidate files are chosen.

Phase three: Migrating and deleting files

- the candidate files that were chosen and scheduled by the second phase are

  - Migrated

  - Deleted

  - Compressed

  - Etc.

each according to its applicable rule

Phase one:	**Phase two:**	**Phase three:**
Selecting candidate files	Choosing and scheduling files `mmapplypolicy -I prepare`	Migrating, deleting or processing files `mmapplypolicy -I yes` (yes = default)

File placement policy

- File placement policies are used to automatically place newly created files in a specific storage pool

- Installed file placement policy is required for file systems with more than one storage pool

- Example

```
[root@gpfs51 ~]# mmlspool fs1
Storage pools in file system at '/fs1':
Name                   Id   BlkSize Data Meta ...
system                  0     4 MB  yes  yes ...
data                65537     4 MB  yes   no ...
```

```
[root@gpfs51 ~]# cat policy/default.pol
RULE 'PROFILE' SET POOL 'system'
     WHERE FILESET_NAME LIKE 'profiles'
RULE 'MMBACKUP' SET POOL 'system'
     WHERE FILESET_NAME LIKE 'mmbackuptmp'
RULE 'DEFAULT' SET POOL 'data'
```

```
[root@gpfs51 ~]# mmchpolicy fs1 policy/default.pol
Validated policy 'default.pol': Parsed 3 policy rules.
Policy `default.pol' installed and broadcast to all nodes.
```

```
[root@gpfs51 ~]# mmlspolicy fs1 -L
RULE 'PROFILE' SET POOL 'system'
     WHERE FILESET_NAME LIKE 'profiles'
RULE 'MMBACKUP' SET POOL 'system'
     WHERE FILESET_NAME LIKE 'mmbackuptmp'
RULE 'DEFAULT' SET POOL 'data'
```
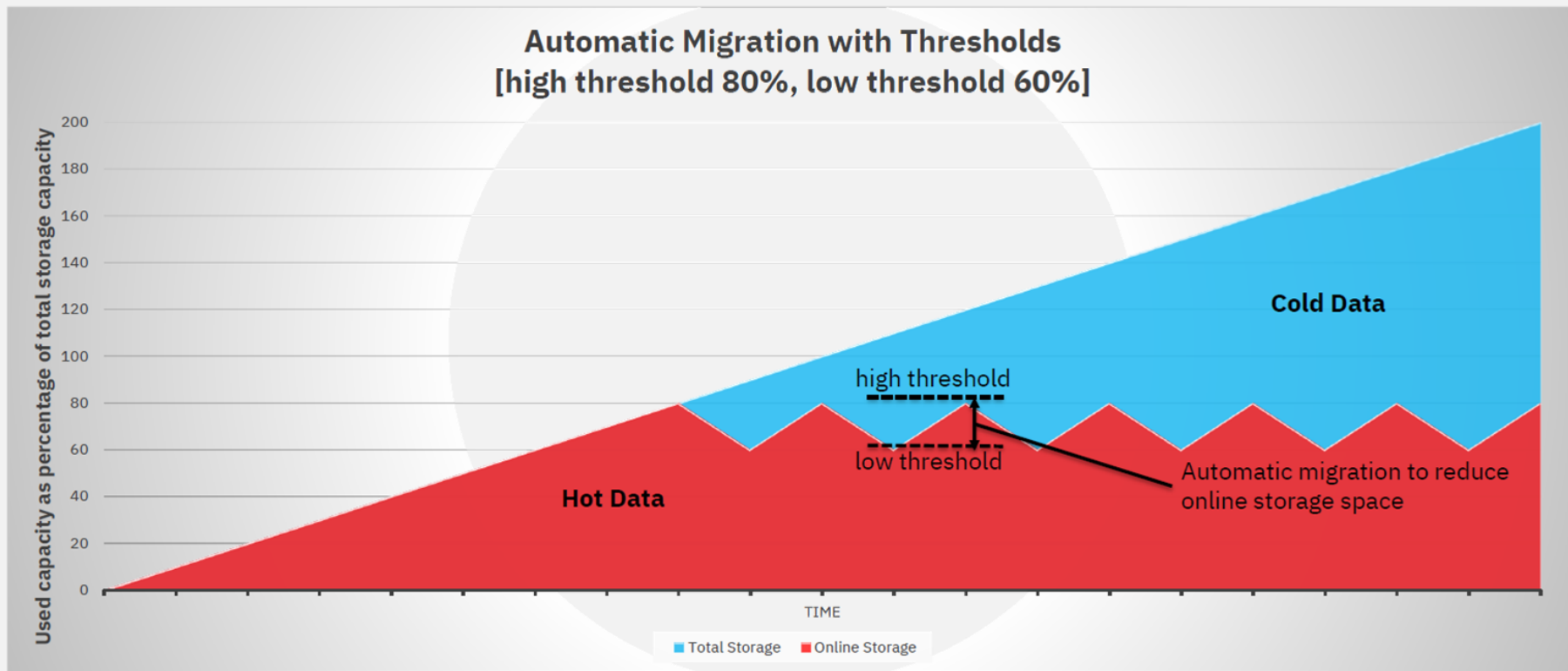
File migration policy - example

Migrate all pictures from storage pool 'system' to storage pool 'data' when the following conditions apply:

- Create date of the file is older than 7 days and younger than 720 days
- File location can be in the file set "user" or in the file set "pool"
- File extension must be jpg (not case sensitive)
- Start migration if file system utilization is larger than 60%
- Stop migration if file system utilization of 80% reached
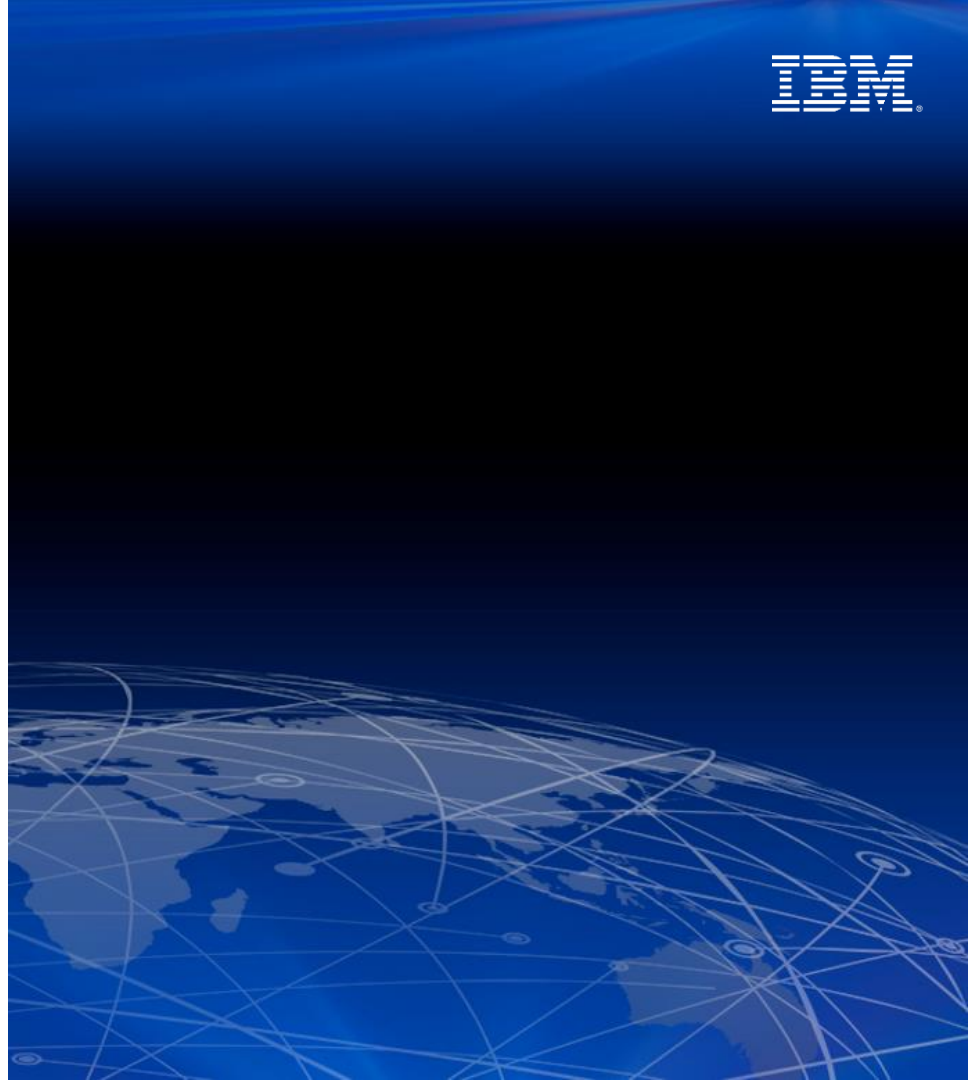- Migrate largest files first

```
[root@gpfs51 ~]# cat policy/ilm_migrate_pics_to_data2.pol
RULE 'move_pics_to_data' MIGRATE FROM POOL 'system' TO POOL 'data'
THRESHOLD (80,60)
WEIGHT (FILE_SIZE)
WHERE ((FILESET_NAME LIKE 'user') OR (FILESET_NAME LIKE 'pool'))
AND (lower(NAME) LIKE '%.jpg' )
AND (DAYS(CURRENT_TIMESTAMP) - DAYS(CREATION_TIME)) > 7
AND (DAYS(CURRENT_TIMESTAMP) - DAYS(CREATION_TIME)) < 720
```

# Managing continuous data growth with lifecycle policies

Automatic Migration with Thresholds
[high threshold 80%, low threshold 60%]

Cold Data

high threshold

low threshold

Automatic migration to reduce online storage space

Hot Data

Used capacity as percentage of total storage capacity

TIME

■ Total Storage   ■ Online Storage

Active file management (AFM)

- Enables data sharing across clusters, even if the network is unreliable or has high latency.

- Changes managed per independent fileset. Results in modular, scalable architecture capable of supporting billions of files and peta bytes of data.

- AFM Gateway nodes handle the sync. Data is eventual consistent.

- "Relationships" can be generated to NFS, Scale and S3 storage systems
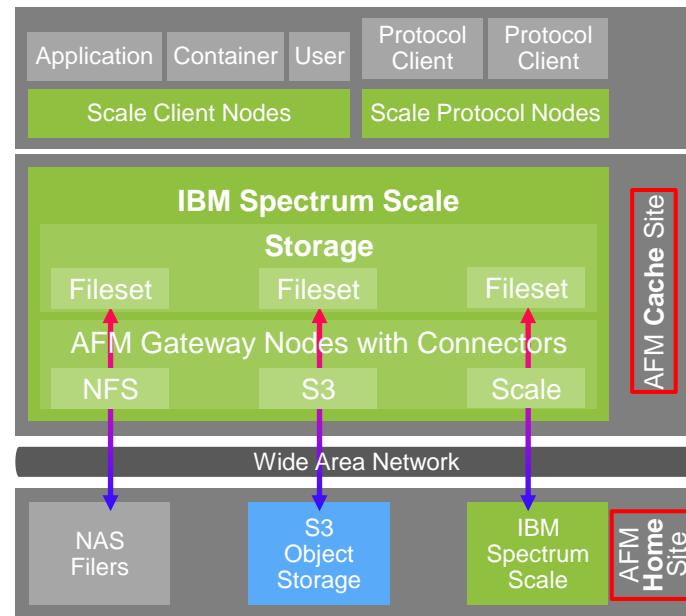
Active file management (AFM) collaboration

- Incorporates file and object into active capacity

- Masks wide-area network latencies and outages by transparently caching data locally

- A multinode, consistent cache of data transparently integrated with the file system users an applications

- Unifies object storage from S3 buckets or NFS data from non-IBM storage file sets such as PowerScale or Netapp

- Individual files in the file set can be compressed

- Different Caching Modes (Read-Only, Single-Writer, Independent Writer, ..)
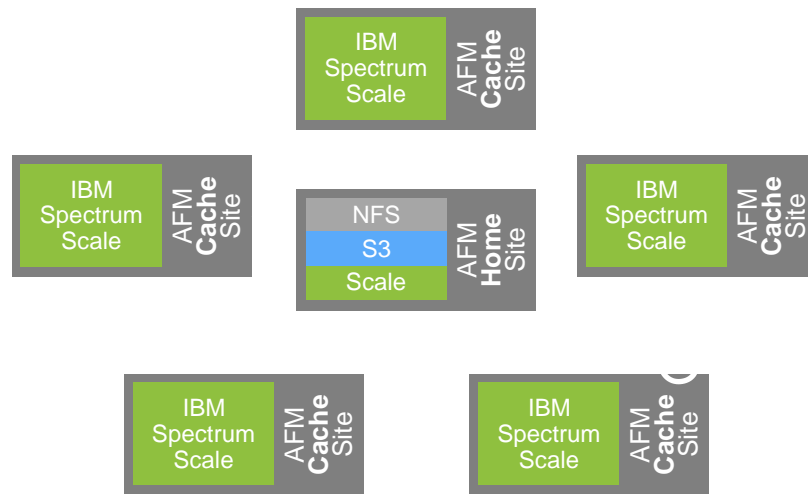
Solve the data and AI puzzle faster with global search, data tagging and orchestration
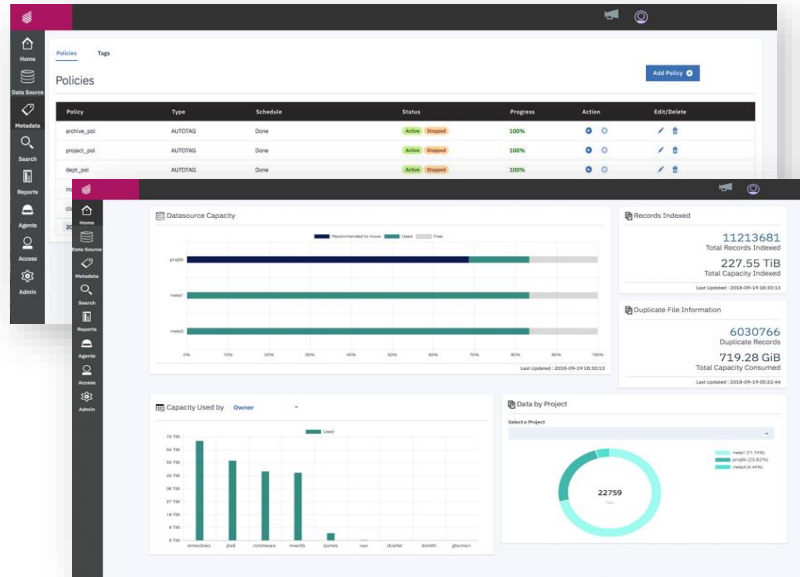
Intelligent insights
Discover security, compliance and governance issues before they become problematic

Faster results
Find data sets faster for AI and ML analysis and place the data at the location or the best performing tier of storage
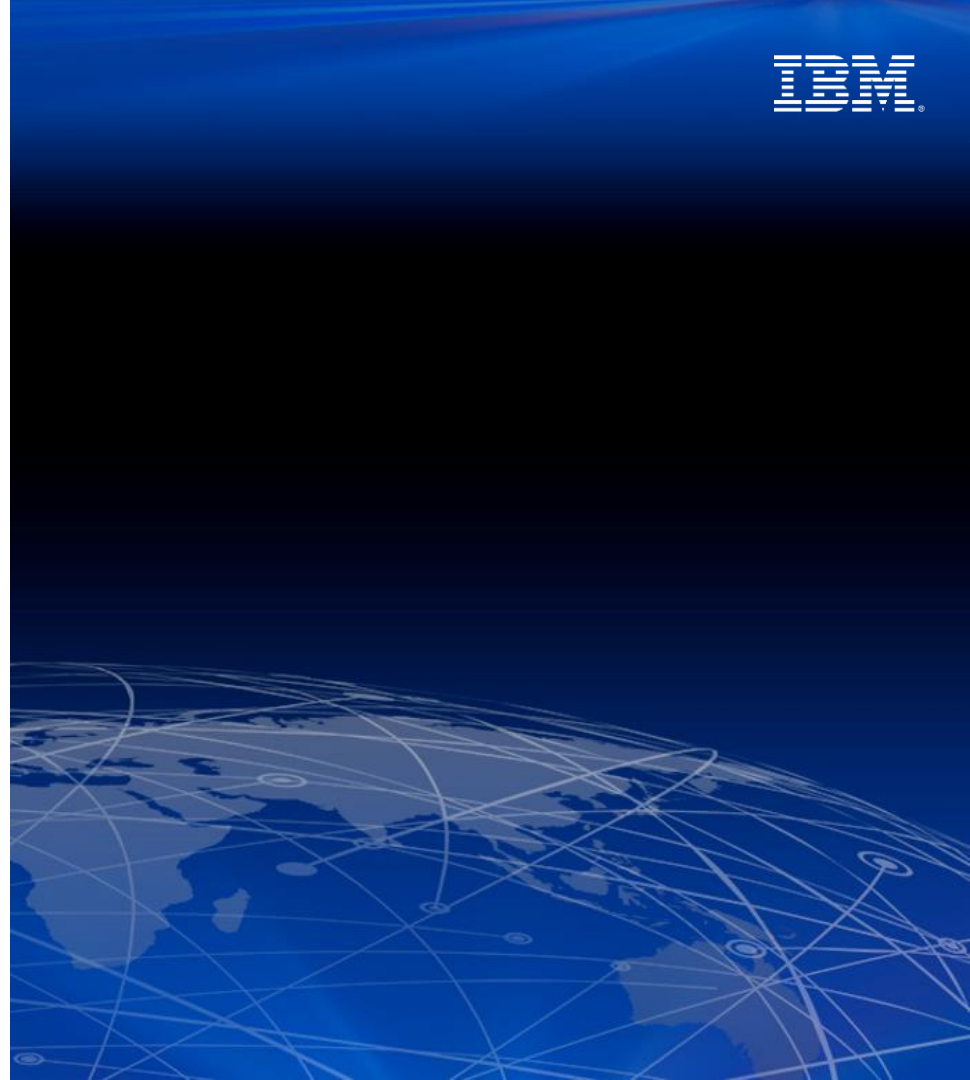
Real time data equals better AI
With continuous real time updates whenever changes to data occur means users can easily find and analyze billions of current objects and files in less than a second.
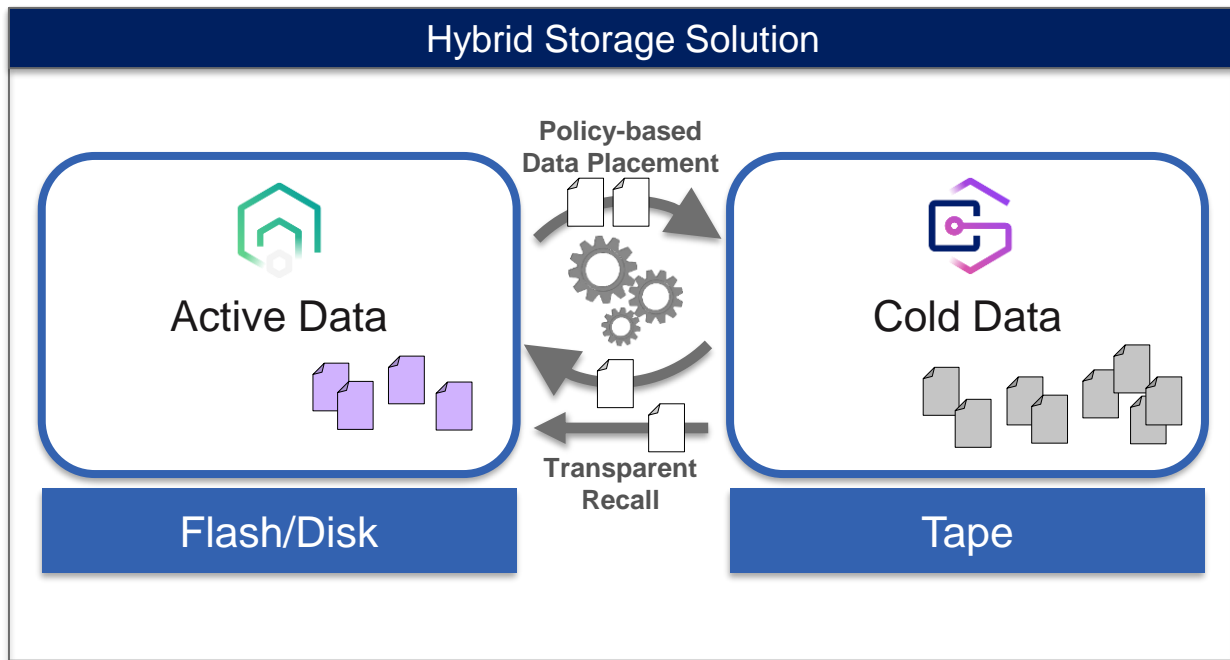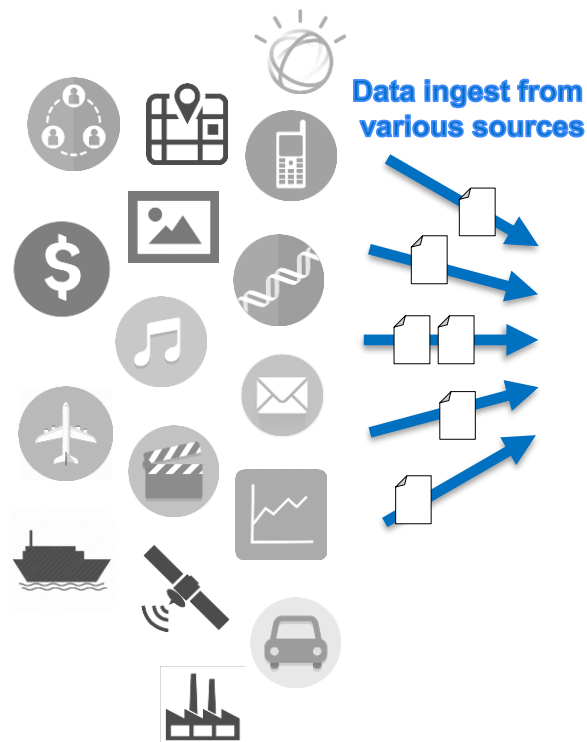
Leverage IBM AI solutions
Link file and object data to IBM Watson solutions and IBM Cloud Pak for Data for deeper insight

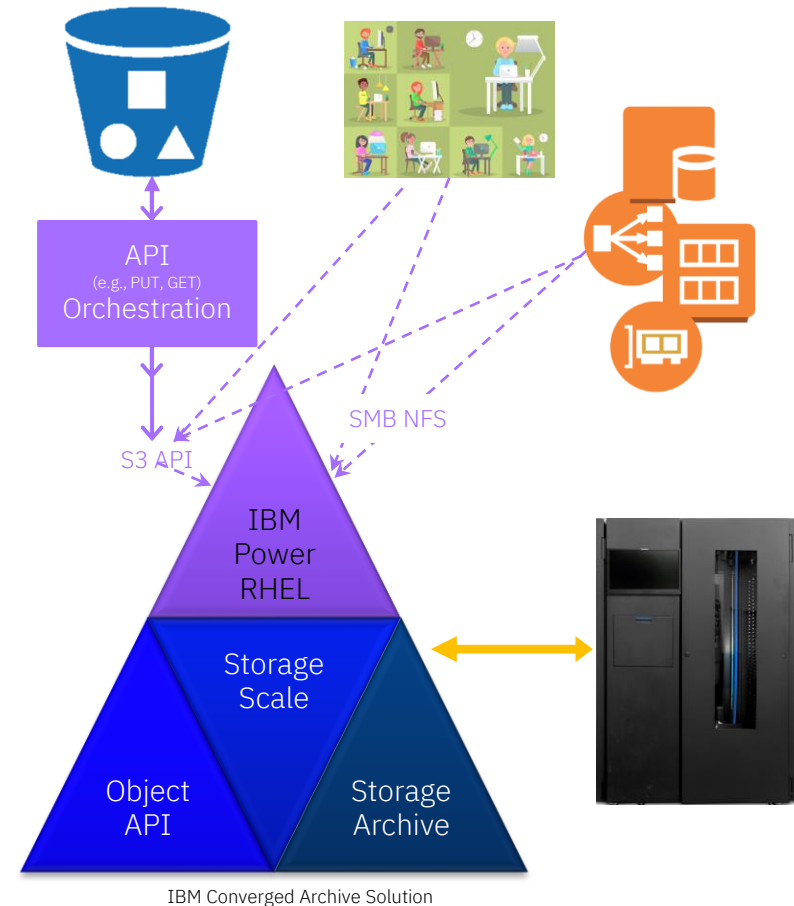# Storage economics: right place, right time with Scale & Archive

Use Case – Archive

→ NAS storage continues to grow
→ Object Storage out of control
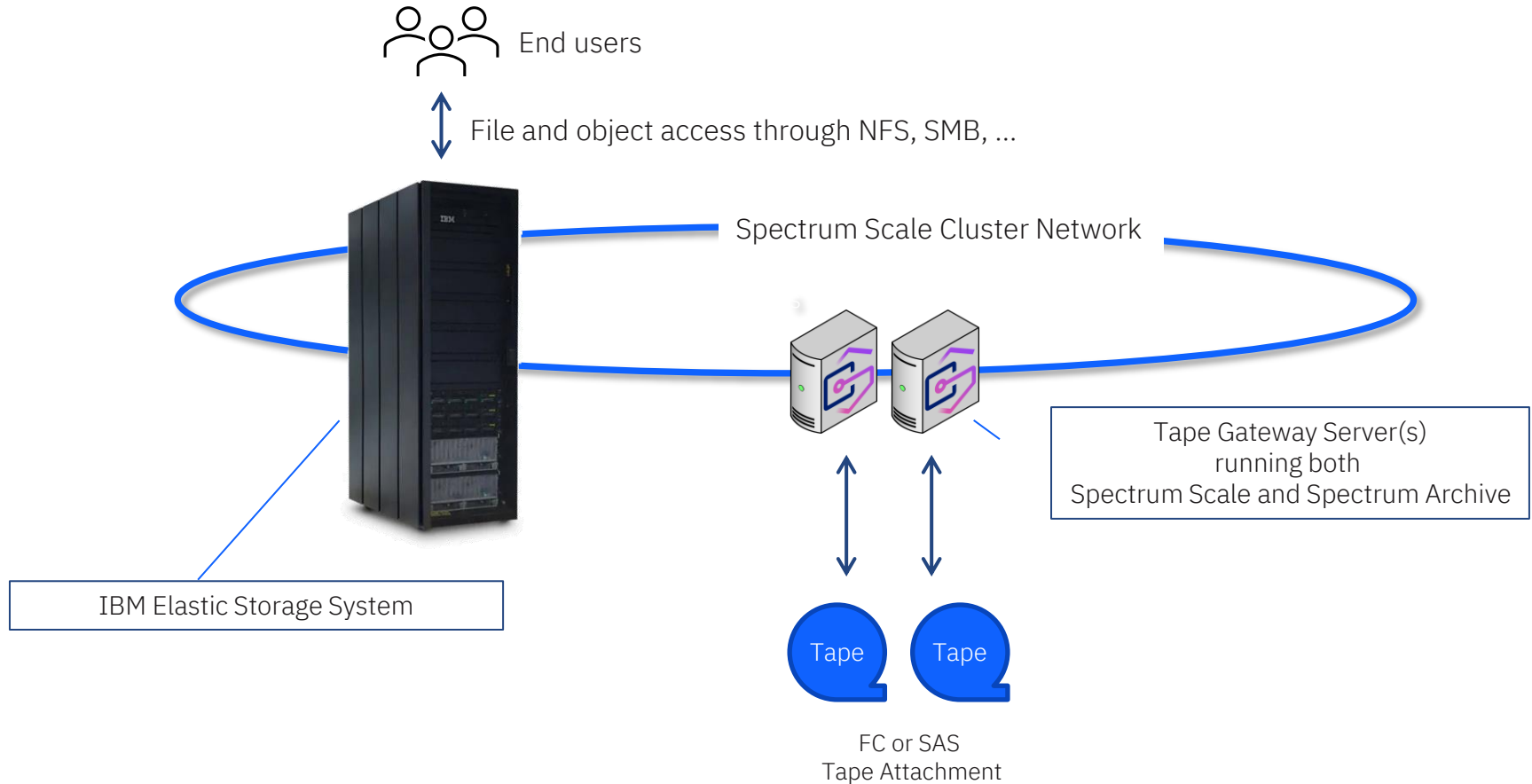→ Silos of data include Analytics

with Storage Archive:

- User access to data from anywhere to anywhere

- Application independent
  - Changes in backend storage do not affect the application
  - Users are not affected by upgrades
  - User data access familiarity for all use cases

- Lower user cost in charge back environments
  - Tape offers internal data control, no charge for more recall
  - User prioritization of recall, no external dependency
  - User transparent tiered storage through ILM

- Infrastructure upgrades are non-disruptive

API
(e.g., PUT, GET)
Orchestration

S3 API

SMB NFS

IBM
Power
RHEL

Storage
Scale

Object
API

Storage
Archive

IBM Converged Archive Solution
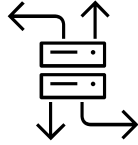
Automated ILM migrates data to Deep Archive with
lowest cost tape in open format LTFS

IBM Storage Scale System with IBM Spectrum Archive EE

End users

File and object access through NFS, SMB, ...

Spectrum Scale Cluster Network

Tape Gateway Server(s)
running both
Spectrum Scale and Spectrum Archive

IBM Elastic Storage System

Tape

Tape

FC or SAS
Tape Attachment

Data & AI
Tiering Storage
# for Data Lakes

Business Challenge

Building the right information infrastructure is crucial for any enterprise. Tiering information is critical for operational constraints, secondary tier on cloud or on tape strategical on it.

Solution

- **Hot Data**, IBM Storage Scale (System)
  File based storage with Object & HDFS support, High End I/O performance, Information Lifecycle Management (ILM), and Sub Micro-seconds access time.

- **Warm Data**, IBM Cloud Object Storage (S3)
  Site fault tolerant, Geo Disperesed and WW scale, Easy to deploy, and Mili-seconds access time

- **Cold Data**, IBM Storage Archive & Tape
  Lowest TCO, Tape ILM target, especially frozen archive, Long term retention and minutes access time, Access as files via LTFS, Reduced floor space requirements and energy consumption and Up to 260PB native capacity in a single tape library

Data & AI
HyperScale
Archival Service
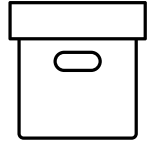
Business Challenge

One of the largest internet and search company challenge was that hyperscale businesses are creating new data every year, and it costs them lots of money to save this data using x86 distributed storage with local disk, while most of the data was accessed and used only to a limited degree.

Solution

Tape near-line storage based on IBM Storage Archive, Storage Scale and Tape Library to save 200PB data. Data migration from machine local disk to tape, saving IT costs with this solution.

Raw Data
Primary storage and primary analytics
Storage Scale (System)
Secondary storage and distribution
Cloud Object Storage
Long term archive
Storage Archive
IBM Tape