

# AFM Enhancements and Use Cases

Spectrum Scale German User Meeting 2022  
Cologne, Germany – October 19-20, 2022

[mdietz@de.ibm.com](mailto:mdietz@de.ibm.com) (Spectrum Scale RAS Architect)

[kekarmar@in.ibm.com](mailto:kekarmar@in.ibm.com) (Spectrum Scale Architect)



# Disclaimer

This information is provided on an "AS IS" basis without warranty of any kind, express or implied, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose. Some jurisdictions do not allow disclaimers of express or implied warranties in certain transactions; therefore, this statement may not apply to you.

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. The development, release, and timing of any future features or functionality described for our products remains at our sole discretion.

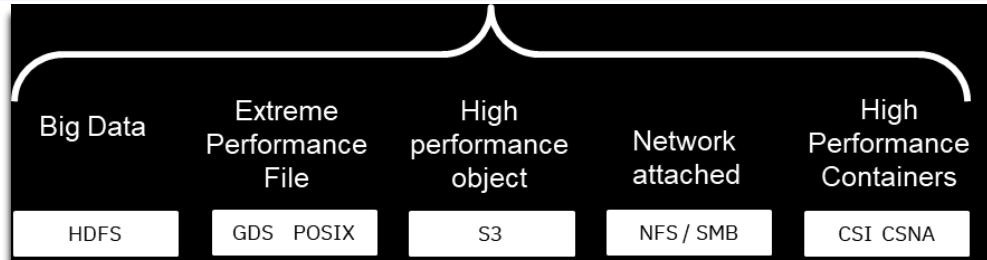
Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve results similar to those stated here.

IBM reserves the right to change product specifications and offerings at any time without notice. This publication could include technical inaccuracies or typographical errors. References herein to IBM products and services do not imply that IBM intends to make them available in all countries.

# IBM's Global Data Platform for File & Object Data



## 1 Data Access Services



## 2 Data Caching Services

### Global Data Platform (powered by Spectrum Scale)



Investment protection

Object Storage

File Storage

NextGen workloads



## 3 Data Management Services

## 4 Data Security Services

- Identify
- Protect
- Detect
- Respond
- Recover

# Global access to data with active file management (AFM)

## Access to multiple sources provides investment protection, global access to data and faster results

- **Investment protection** with an open ecosystem of storage options leveraging multi-vendor and multi-cloud resources
- **Increase application agility** accessing data from edge to core to cloud by bringing more data to applications wherever they are deployed
- **Quickly scale your data** from resources you choose with performance you require
- **Faster access to remote data** by transparently caching remote data locally when needed

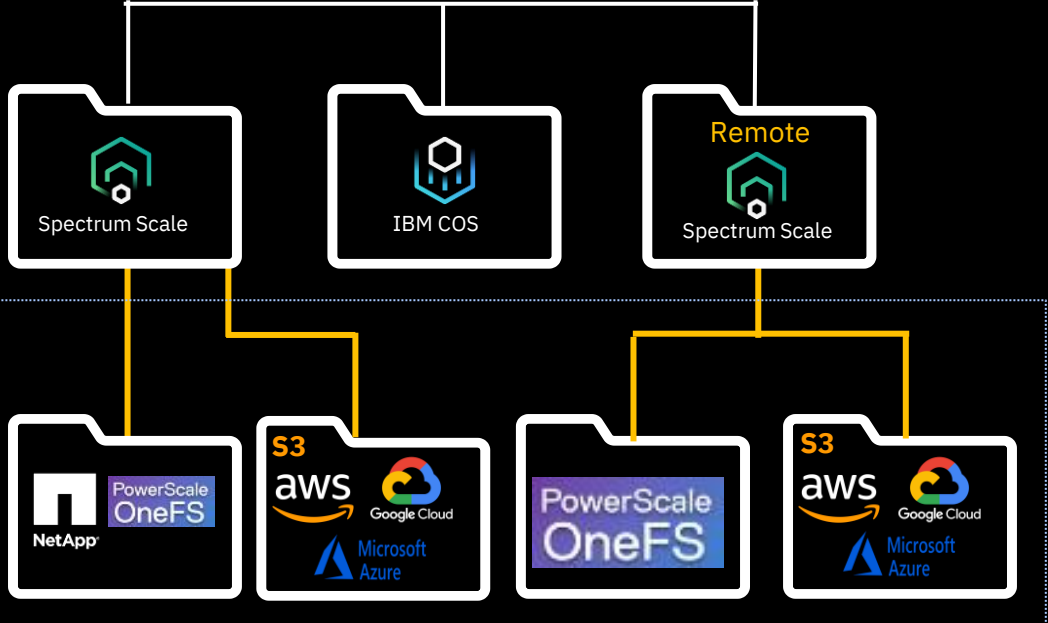
- Turns remote file and object data into active capacity (open ecosystem)
- Masks wide-area network latencies and outages by transparently caching data locally
- Individual files in the file set can be cached ?



Spectrum Scale



Spectrum Fusion



Access any File/Object storage

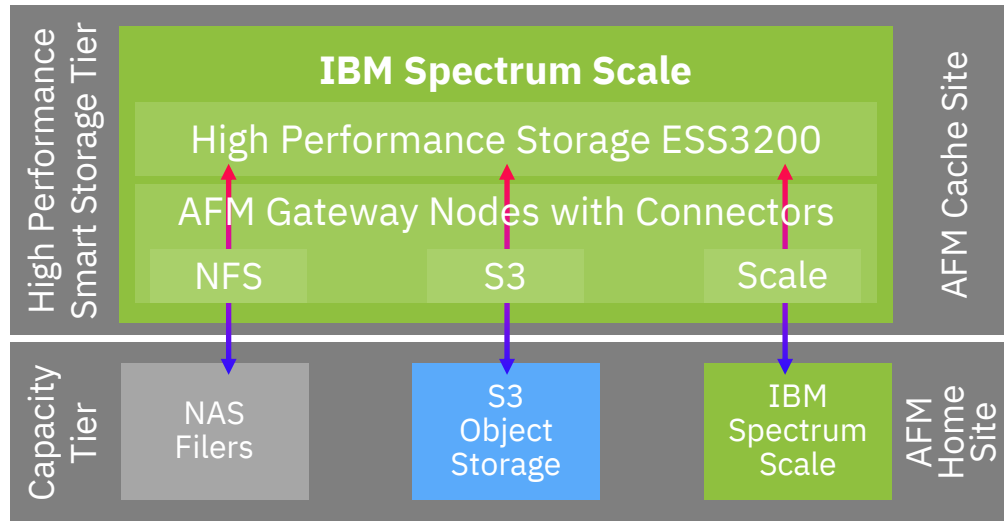
# Data Caching Services (Active File Management) Use cases



Data Virtualization	Data Collaboration	Data Resilience	Hybrid cloud / Bursting
<ul style="list-style-type: none"> <li>• <b>Integrate legacy file and object data</b> stores into single file system to breakdown legacy data silos</li> <li>• <b>Migrate data to new storage</b> or continue to use legacy stores</li> <li>• <b>Create a High-Performance Tier</b> for analytics for legacy data with transparent data access</li> </ul>	<ul style="list-style-type: none"> <li>• <b>Geo-distributed collaboration</b> on data transparently shared between data centers, the cloud and edge sites</li> <li>• <b>Coalesce data to a home site</b> from the edge and redistribute it to all sites</li> </ul>	<ul style="list-style-type: none"> <li>• Provide an <b>asynchronous Disaster Recovery solution</b> for business continuity over WAN distances</li> <li>• Supports analytics and archival access to passive data</li> </ul>	<ul style="list-style-type: none"> <li>• Dynamically increase computation resources in the cloud and <b>optimally make required data available for Cloud bursting</b></li> <li>• Process data consolidated on S3 Cloud Storage on with <b>high performance tier in the Cloud</b> Compute Cluster</li> <li>• <b>Archive data to S3 Object</b> storage</li> </ul>
<p>Public Cloud Services</p> <p>Use case:</p> <p>Enables end user service to upload large amount of data via Object interface that can be analysed on high performance file system</p>	<p>Research / University</p> <p>Use case:</p> <p>Generate 100's of TB per day across multiple silos, leveraged AFM to provide common namespace with transparent multiprotocol data access</p>	<p>Multinational financial services</p> <p>Use case:</p> <p>Disaster Recovery, retention and compliance data with FileNet and ESS</p>	<p>Research Biopharmaceutical</p> <p>Use case:</p> <p>Multi site / public cloud bursting for collaboration</p>

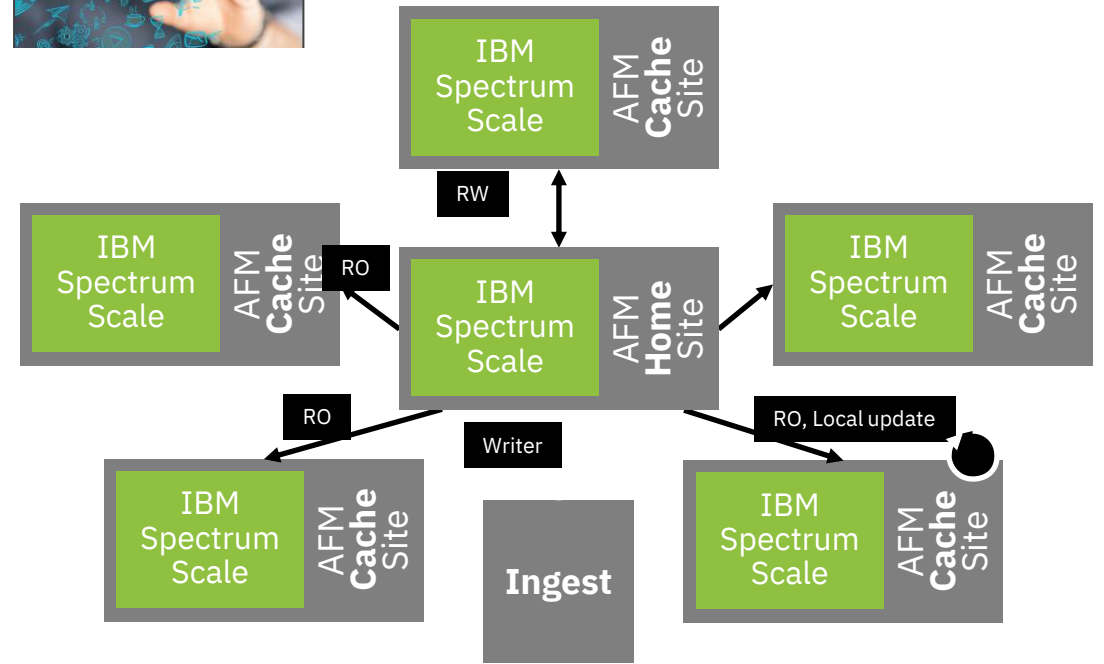
# Data Caching Services (AFM) Use case details

## Data Virtualization



- Vertical caching
- Common namespace across isolated data silos in legacy 3rd party data stores
- Transparent access to all data regardless of silos
- Scale-out Posix performance
- Data export via NFS, SMB, HDFS, Object
- Can be used to seamlessly migrate data to new storage

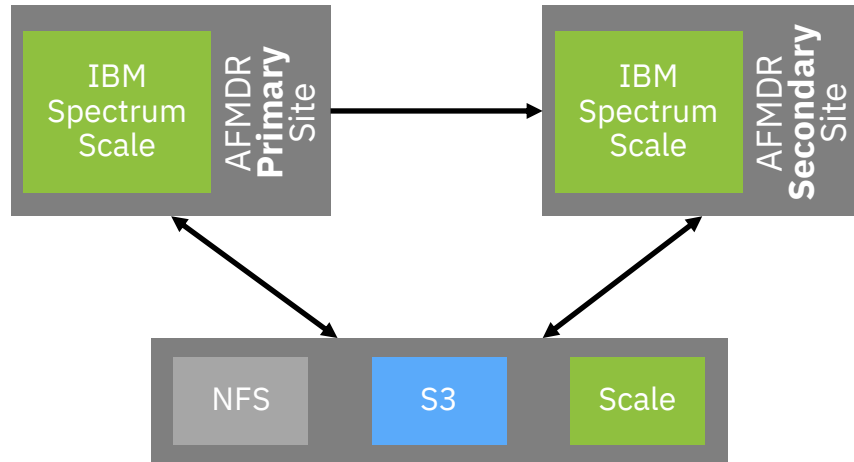
## Data Collaboration



- Consistent cache provides a single source of truth with no stale data copies
- Horizontal caching
- Bi-direction traffic from Edge to Center
- Eventually Consistent data cache
- Transparent on-demand data access and transfer
- Policy driven data prefetch and eviction

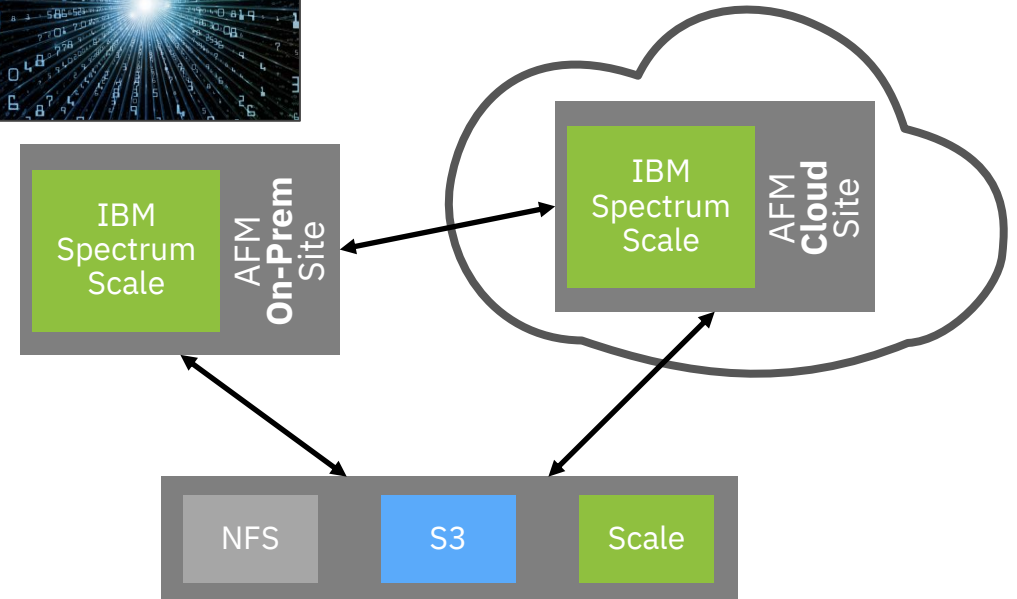
# Data Caching Services (AFM) Use case details

## Data Resilience



- Active-Passive DR over WAN or Cloud
- Designed for high latency and asynchronous DR
- Hot standby failover to DR site
- Automatic fallback data reconciliation
- Read-only access / analytics to all data at passive site

## Hybrid cloud / Bursting

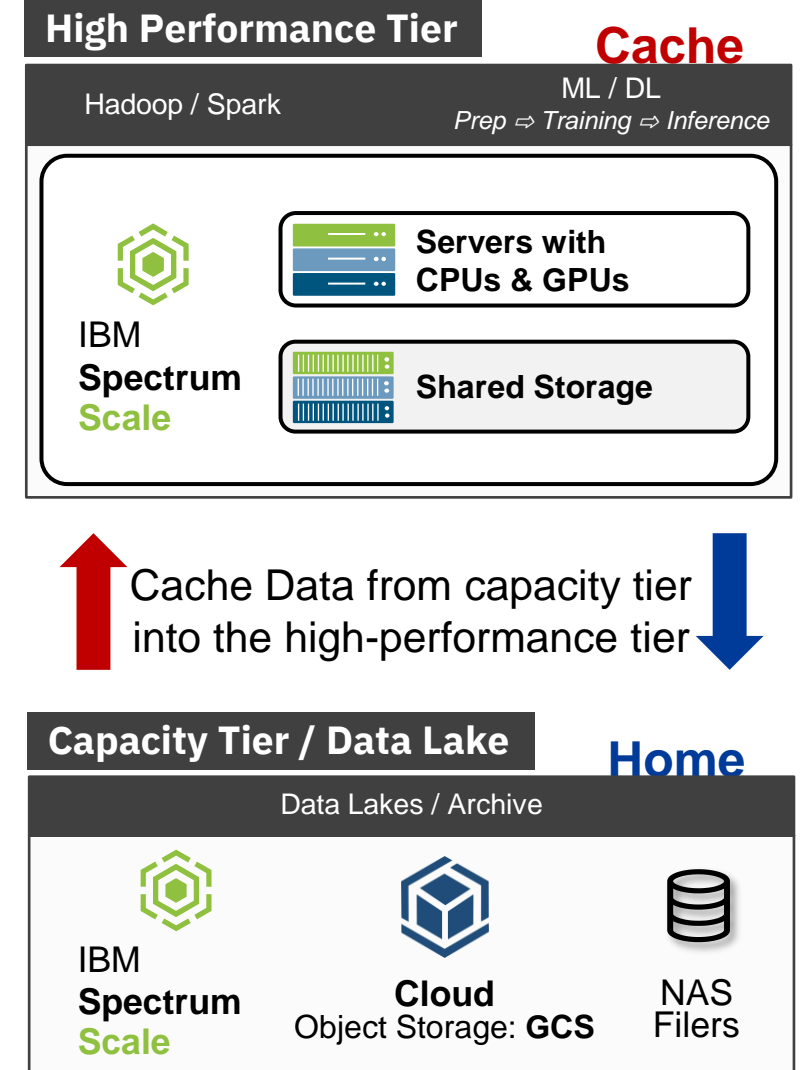


- Rapidly expand compute resources to cloud or data centers
- Common file system creates a single namespace across all locations
- Transparent access to data
- Cost effective way to increase compute on existing data
- Analytic results automatically pushed to home site

# Additions to S3 Object Storage support

Continued additions of Cloud Object Storages environments

- IBM Cloud Object Storage [5.1.0](#)
- Amazon S3 [5.1.0](#)
- Microsoft Azure Blob storage using S3 Gateway [5.1.3](#)
  - AFM doesn't support Microsoft Azure native API
  - Requires intermediate S3 gateway such as Minio
  - Two ways to deploy S3 Gateway (Minio)
    - Deploy as fully managed service in Microsoft Azure market place
    - Deploy Minio on-prem and configure to communicate to Azure Blob
- Google Cloud Platform [5.1.4](#)
  - Use `-gcs` option while configuring AFM to Google Cloud Storage relationship
  - Creation of bucket not supported via AFM. Ensure bucket exists on GCS
- Seagate Lyve Cloud Object Storage [5.1.5](#)
  - Lyve cloud APIs are almost similar with S3 API







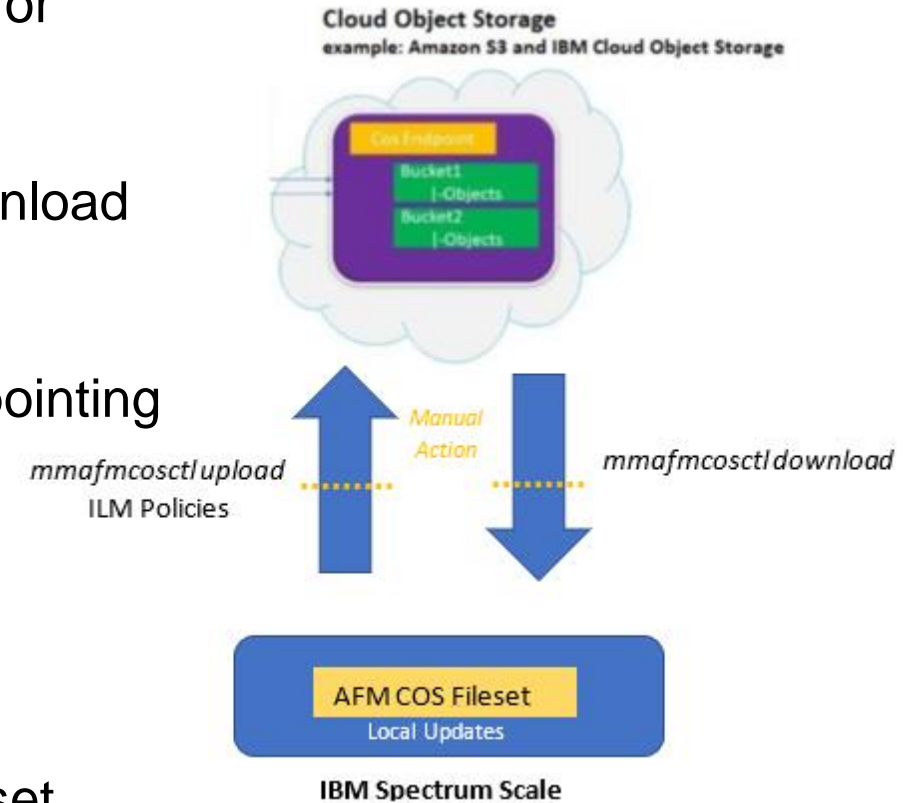
## More AFM S3 Object storage enhancements

- Support of STS token for AFM to cloud object fileset
- Support for expiring and automatically refreshing cloud storage keys
- Tunable to manage key expiration using 'afmObjKeyExpiration'
- Support of creating and upload objects for empty directories in AFM to cloud object storage.
- Support of marking files and directories as local in AFM to cloud object storage fileset.
  - `#mmafmctl fs setlocal -j AFMtoCOS --path /ibm0/fs/AFMtoCOS/file1`
- Support of adding user defined prefix in AFM to cloud object storage fileset.
  - `#mmafmcosconfig fs1 afmbktprefix1 --endpoint https://region@endpoint --object-fs \  
--xattr--prefix dir1 --bucket bkt1 --acis--mode sw`
- Support of replicating more than 2K metadata in AFM to cloud object storage fileset.



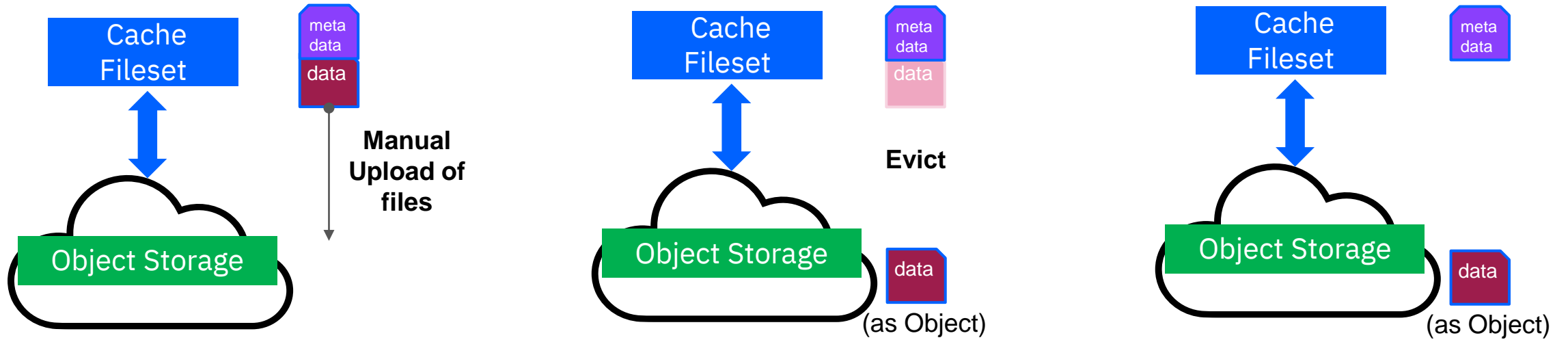
# AFM for archive use case with Manual Update (mu) mode

- Supports manual upload/ download of objects using ILM policies or object list and avoids automatic upload/ download
- All changes are local. Manual action is required to upload or download files to cloud object storage
- Metadata is refreshed only once when the MU fileset is created pointing to non-empty bucket
- Manual control over deletion from cache
- Manual deletion from Cloud Object Storage
- An independent fileset can be converted to a MU mode AFM fileset
- Auto removal of files/ objects from Cache and Object Storage using fileset parameter 'afmMuAutoRemove'
- Specify policy to be used with `mmafmcostl reconcile -policy` command





# AFM for archive use case with Manual Update (mu) mode





# AFM Cloud Object Storage Operation modes

## ObjectFS mode

- Behaves like normal AFM modes fileset.
- Objects are downloaded on read or on access
- IW and SW modes push files to cloud object storage  
RO, LU, IW automatically pulls objects from the cloud object storage and stores as files.

## ObjectOnly mode

- Default for object operation mode
- No on-demand refresh on read
- Need to manually download metadata/data from COS.
- Objects are uploaded automatically (IW and SW)
- Avoids frequent trips and reduce network contention by selective download/uploads.

	ObjectFS	ObjectOnly
Read Only (RO)	Upload - NA Download – On access (Auto)	Upload - NA Download - On demand
Local Update (LU)	Upload – NA (only On demand) Download – On access (Auto)	Upload – NA (only On demand) Download - On demand
Single Writer (SW)	Upload - Auto Download - On access / On demand	Upload - Auto Download – On demand
Independent Writer (IW)	Upload - Auto Download – On access (Auto)	Upload - Auto Download - On demand
Manual Update (MU)	Upload - On demand Download - On demand	

# Important AFM S3 Object configuration parameters

- **afmMUAutoRemove**
  - Enabling this for a MU fileset will queue file remove operations to the COS automatically
- **afmObjectDirectoryObj**
  - Enabled at filesset level to support synchronisation of all directories with/without object with COS
- **afmObjectFastReaddir**
  - Enabled at the filesset level to skip the fetching of extended attributes and acl from COS for readdir operations leading to faster listing operations
- **afmObjKeyExpiration**
  - Specifies the COS key expiration timeout value which allows to reload the access/secret keys after the defined timeout value. Default is 36000 seconds. Set at the cluster level
- **afmParallelReadThreshold**
  - Defines the threshold beyond which parallel reads become effective and is enabled by default . Default is 1024MB.
- **afmParallelReadChunkSize**
  - Defines the minimum chunk size of the read that needs to be distributed among the gateway nodes during parallel read. Default is 128MB . A value of zero disables parallel reads
- **afmPrefetchThreshold**
  - Controls partial file caching. Default is 0 which caches the entire file by pulling all the blocks when 3 blocks are read at the cache. Values are in the range 1-100 and specified the percentage of file that must be read to cache the whole file. A value of 100 disables full file prefetching.
- **afmAsyncDelay**
  - Time by which the asynchronous operations are delayed to home



# AFM DR enhancements

## ***AFM resync version 2 improves replication on heavily stressed systems***

- Updates to message queuing to improve AFM resync and recovery
- Lower memory usage on gateway nodes with faster replication
- Faster recovery and resync after gateway node failures
- Improved role reversal in AFM-DR
- See 'AFM resync version 2' in the knowledge center, and the `afmResyncVer2` parameter in `mmchfileset` to activate this feature

## ***Enable support of 60 mins RPO for AFM-DR fileset***

- Provides reduced RPO of 60 minutes instead of 720 minutes
  - Fileset level snapshots now do not require file system level IO freeze enabling smaller RPO
- Default RPO is still kept disabled
- Admin can configure to set the RPO according to the workload.
- Note: With high workload, `afmRPO` might not triggered properly. Admin might see RPO miss event if RPO was not triggered due to file system quiesce issue. This can be avoided by setting RPO value based on workload.



# AFM network related enhancements

## ***NFSv4 support for AFM replication***

- Allows NFSv4 to be used as the underlying AFM protocol for data transfer
- Supports NFSv4 ACLs on third-party storage to support data migration

## ***Enable support of IPV6 protocol for AFM, AFM-DR & AFM to cloud object fileset***

- AFM fileset can use IPV4, GPFS(Multi-cluster), http/https, IPV6 protocols to synchronize data.
- IPv6 support is added for AFM, AFM-DR, AFM to Cloud Object Storage filesets.
- To use Parallel Data Transfer for AFM fileset, user can create mapping using IPV6
- Only IPv6 (cache) to IPv6 (home) communication is supported.
- Tested with kernel NFS, Ganesha/CES nfs, GPFS (Multi-cluster), http/https which is pointing to IPV6 target.
- Gateway Node and Target node should be able to resolve/talk on IPv6 addresses.
- A Cache and Home Cluster needs to be configured using IPV6 parameters before configuring AFM fileset.
- Hostname/IP should be using IPV6 address.
- While creating AFM fileset, target IPv6 addresses must be mentioned in the square brackets



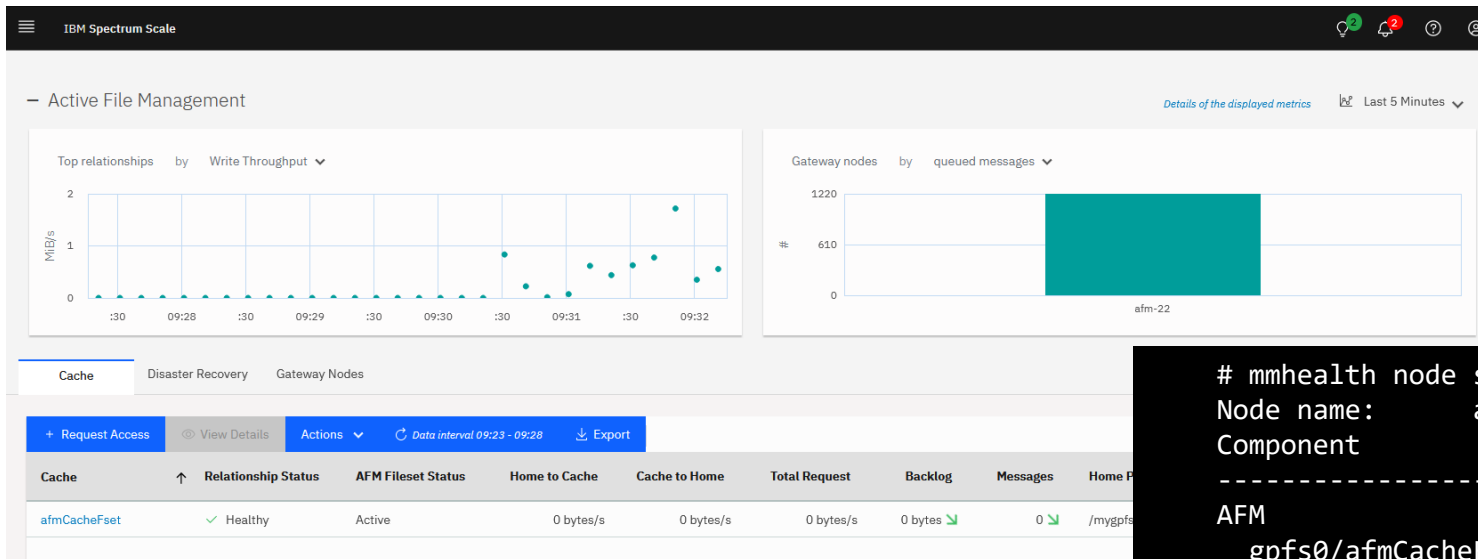
# AFM configuration and monitoring enhancements

## Configuration simplification

- The afmDIO parameter is now set by default based on the backend protocol
- Support eviction of partially cached files (see mmadmctl documentation for details)

## Monitoring enhancements

- Query uncached and dirty files, and new read and write statistics
- Enhanced cloud object upload and download statistics (options `-upload-stats` and `-download-stats`)
- See [mmafmcscctl](#) man page for more details



```
# mmhealth node show AFM -v
Node name:      afm-22
Component      Status          Status Change  Reasons & Notices
-----
AFM            HEALTHY        2022-10-17 03:16:30  -
gpfs0/afmCacheFset HEALTHY        2022-10-14 03:41:05  -
```





# AFM Outlook

## *Upcoming AFM improvements*

- AFM recovery performance improvements
- Disable AFM without requiring to unlink the fileset, migration usecase
- Out-of-band download to improve performance of AFM+COS filesets
- AFM+COS usability improvements. User exits after command completion and download/upload statistics
- Multi target replication (tentative) . Replicate the same object to different COS providers or targets.
- Support Ceph as Object backend

Thank you for using  
IBM Spectrum Scale!