AFM Deep Dive

Storage Scale UK User Group Meeting
London, UK – June 27-28, 2023

Venkat Puvvada (vpuvvada@in.ibm.com)

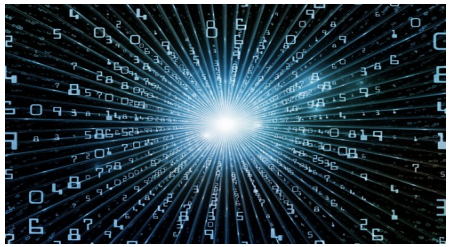
Disclaimer

Storage Scale

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.

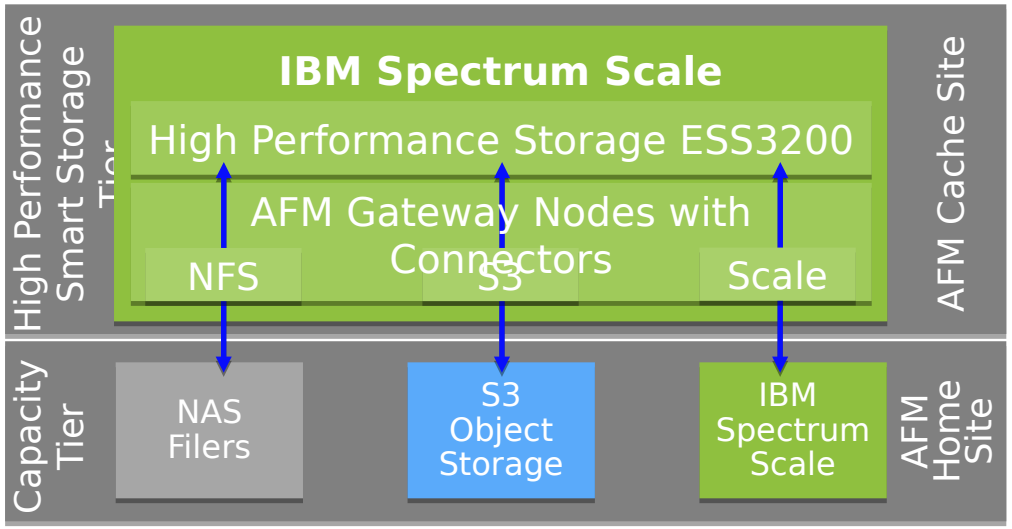
IBM reserves the right to change product specifications and offerings at any time without notice. This publication could include technical inaccuracies or typographical errors. References herein to IBM products and services do not imply that IBM intends to make them available in all countries.

Data Caching Services (Active File Management) Use cases

			
Data Virtualization	**Data Collaboration**	**Data Resilience**	**Hybrid cloud / Bursting**
• **Integrate legacy file and object data** stores into single file system to breakdown legacy data silos • **Migrate data to new storage** or continue to use legacy stores • **Create a High-Performance Tier** for analytics for legacy data with transparent data access	• **Geo-distributed collaboration** on data transparently shared between data centers, the cloud and edge sites • **Coalesce data to a home site** from the edge and redistribute it to all sites	• Provide an **asynchronous Disaster Recovery solution** for business continuity over WAN distances • Supports analytics and archival access to passive data	• Dynamically increase computation resources in the cloud and **optimally make required data available for Cloud bursting** • Process data consolidated on S3 Cloud Storage on with **high performance tier in the Cloud** Compute Cluster • **Archive data to S3 Object** storage
Public Cloud Services Use case: Enables end user service to upload large amount of data via Object interface that can be analysed on high performance file system	Research / University Use case: Generate 100's of TB per day across multiple silos, leveraged AFM to provide common namespace with transparent multiprotocol data access	Multinational financial services Use case: Disaster Recovery, retention and compliance data with FileNet and ESS	Research Biopharmaceutical Use case: Multi site / public cloud bursting for collaboration

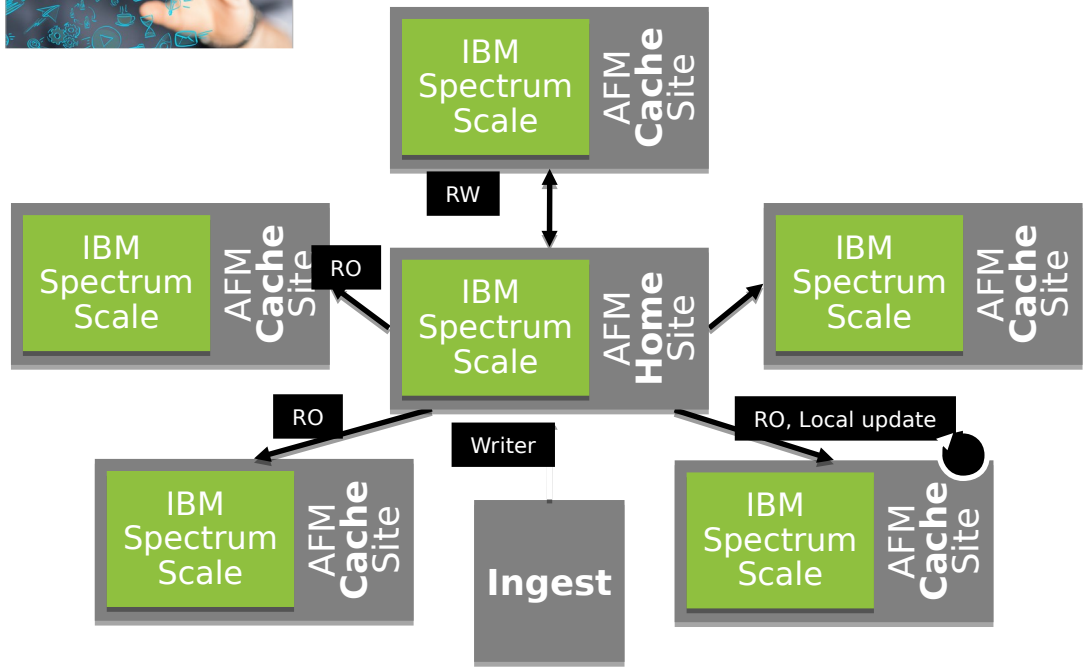
Data Caching Services (AFM) Use case details

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

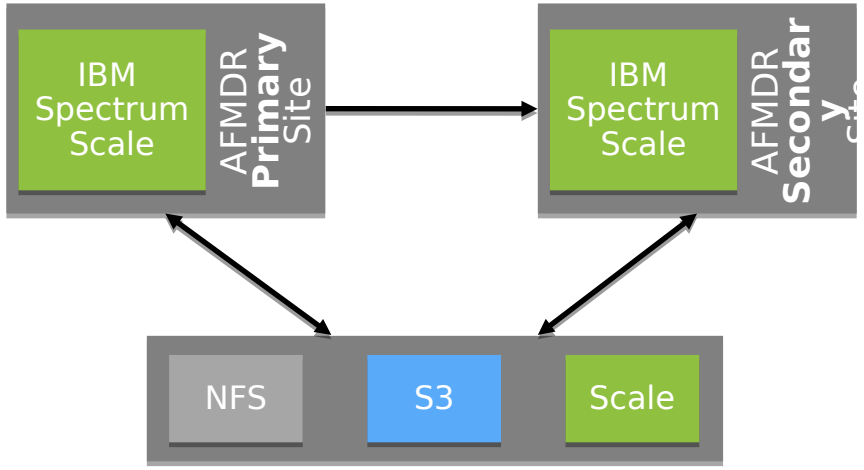
Data Collaboration



- 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

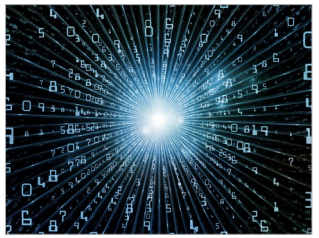
Data Caching Services (AFM) Use case details

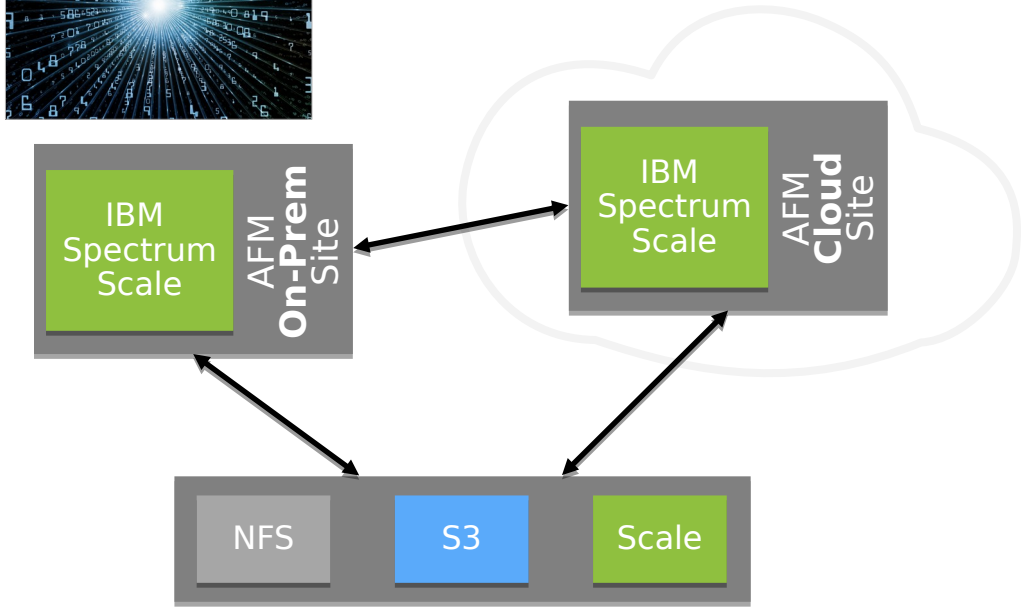
Data Resilience



- 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





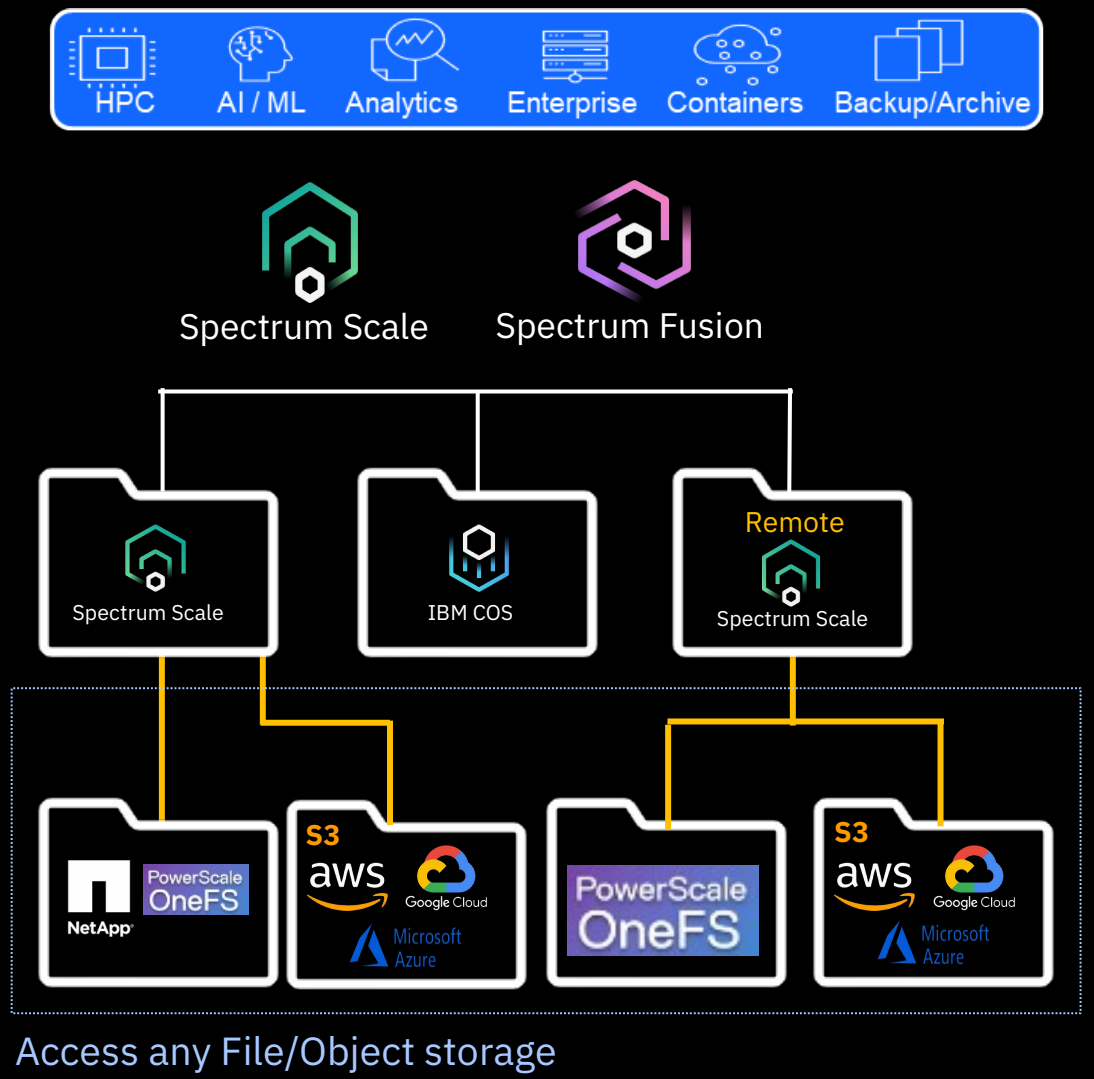
- Rapidly expand compute resources to cloud or data centers
- Common file system creates a single namespace across all locations
- Transparent access to data
- Cost effective way to increase compute on existing data
- Analytic results automatically pushed to home site

Global access to data with active file management (AFM)

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

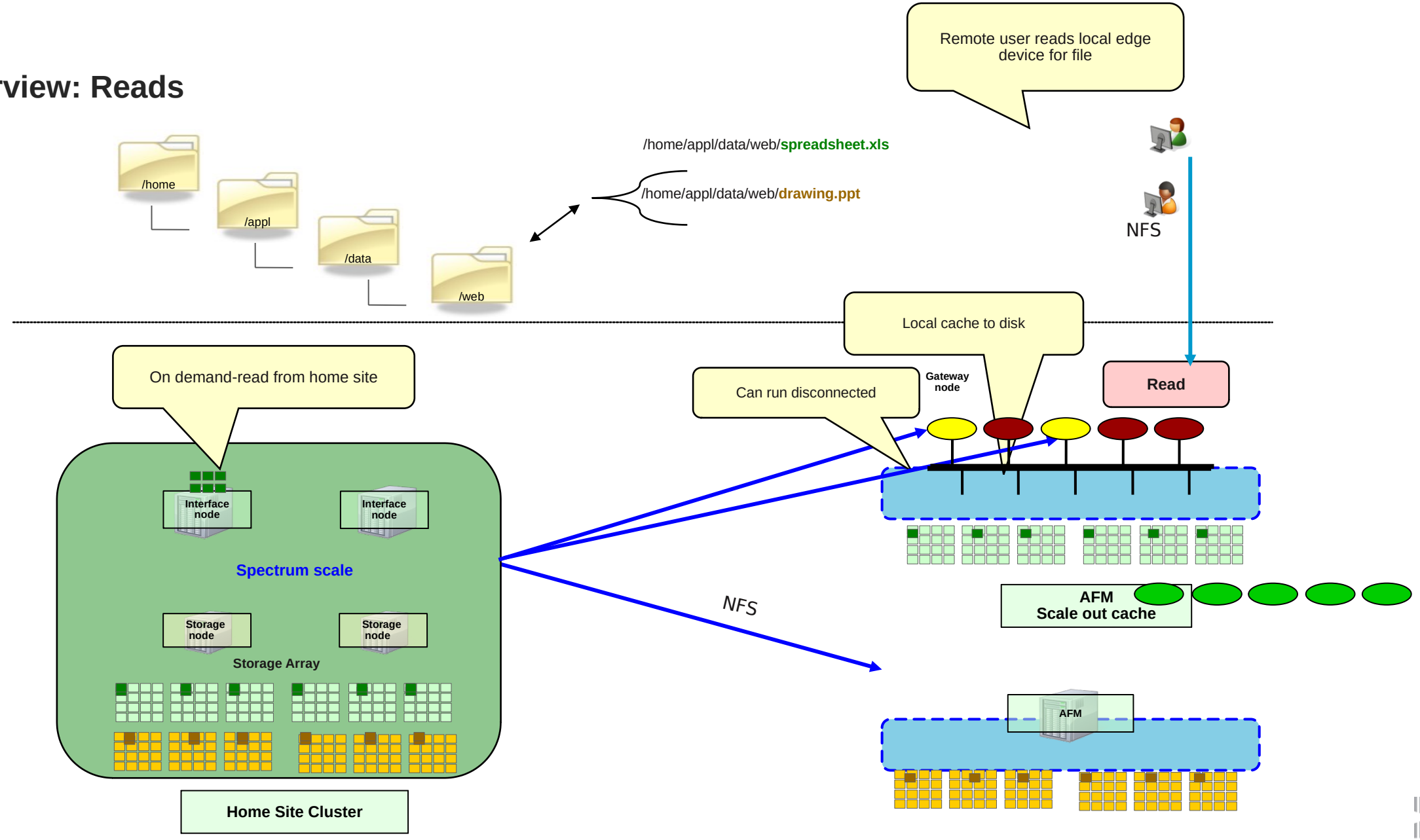
- **Investment protection** with an open ecosystem of storage options leveraging multi-vendor and multi-cloud resources
- **Increase application agility** accessing data from edge to core to cloud by bringing more data to applications wherever they are deployed
- **Quickly scale your data** from resources 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 ?



Access any File/Object storage

AFM Overview: Reads

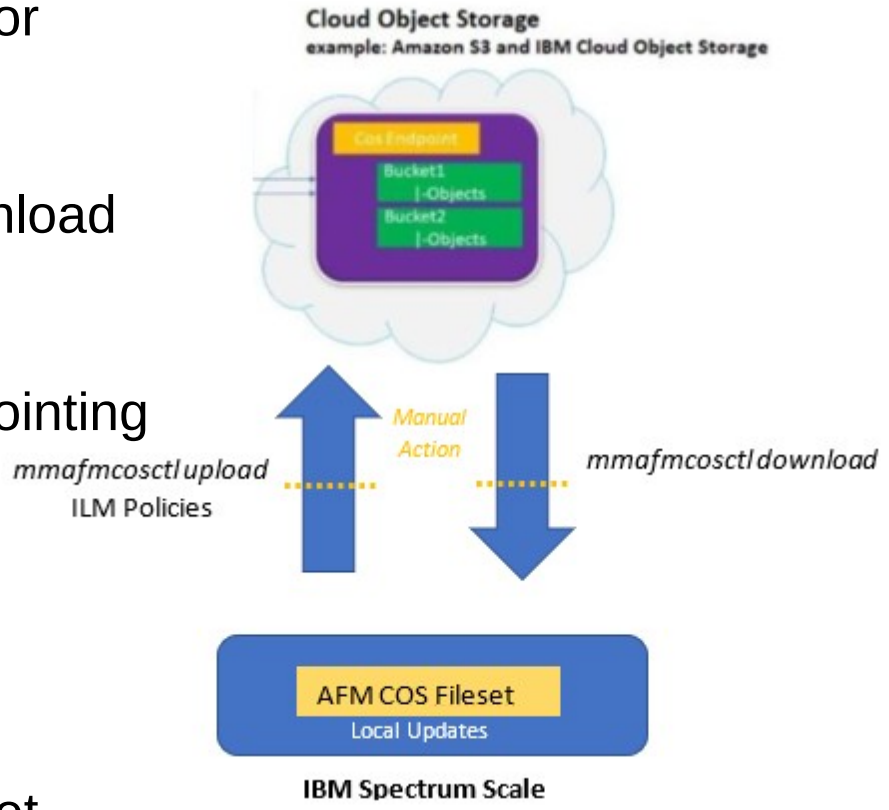


AFM S3 enhancements

- 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 --acls--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 mmafmcosctl reconcile –policy command



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 enhancements

- Allows NFSv4 to be used as the underlying AFM protocol for data transfer
- Supports NFSv4 ACLs on third-party storage to support data migration
- AFM fileset online-disable after the migration (without unlink)
- AFM lookup performance improvements (avoids fileset locks)
- Better handling of the read requests on home migrated files(afmSyncReadMount)
- IPv6 Support
- Prefetch improvements (--skip-dir-list-file and --empty-ptrash options)
- Resync improvements (afmResyncVer2)
- Default hash version is 5 (assign gateway to a fileset)

AFM Tuning (5.1.8, new options)

- AfmAllowReadBypass – Allows small read requests to go home even though full file prefetch is in progress. Default is “no”.
- AfmReadBypassMapSize – Map size of the bypass reads to avoid the lock contention. Default values is 0.
- AfmFastLookup – Enables fast lookup on the AFM filesets without acquiring the fileset locks. Default value is “no”. Not supported on RHEL 9.2 and 8.8.
- AfmObjNumTransferProc – Number of S3 communication process used for the fileset. Default value is 1. Maximum value is 4.
- AfmReadSparseThreshold – If this value is set to 0 on the object enabled filesets, file is uploaded even though file is opened for the writing.

AFM Outlook (intended)

Upcoming AFM improvements

- AFM recovery performance improvements
- Out-of-band metadata population over NSD/GPFS backend
- Immutable snapshots replication
- AFM+COS usability improvements. User exits after command completion and download/upload statistics
- Multi target replication . Replicate the same object to different COS providers or targets.
- Support Ceph as Object backend
- Inline migration from TCT to AFM
- Azure BLOB API support
- Dependent fileset replication
- Block based eviction and range prefetch

Thank you for using IBM Storage Scale!