

Integration of IBM Aspera sync with IBM Spectrum Scale

Nils Haustein, STSM, EMEA Storage Competence Center



IBM Spectrum Scale Strategy Days 2019



Agenda

Introduction to IBM Aspera sync

Integration with IBM Spectrum Scale

Differentiation with AFM

Challenges with TCP and alternative technologies

Distance degrades conditions on all networks

- Latency (or Round Trip Times) increases
- Packet loss increases
- Fast networks are just as prone to degradation

TCP performance degrades severely with distance

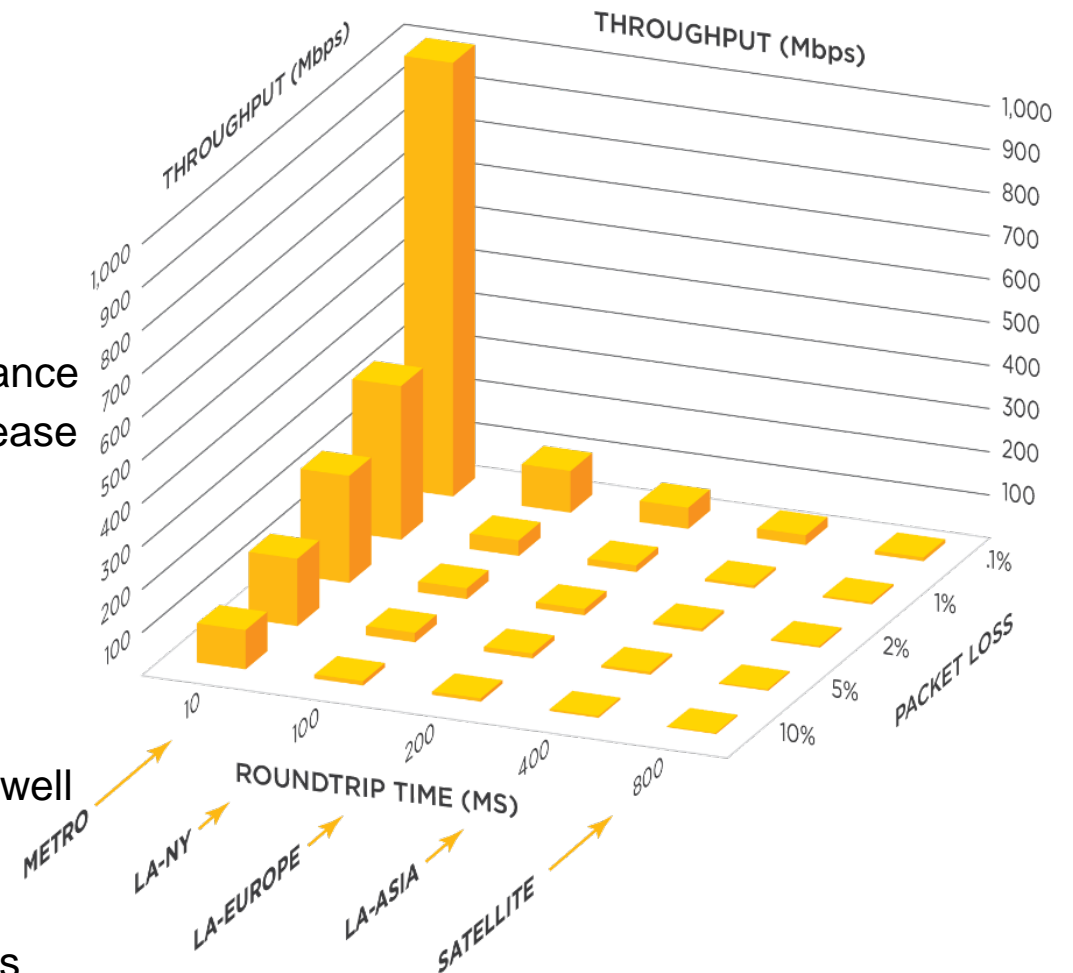
- TCP was designed for LANs and does not perform well over distance
- Throughput bottlenecks are severe as latency & packet loss increase

TCP does not scale with bandwidth

- TCP designed for low bandwidth
- Adding more bandwidth does not improve throughput

Alternative technologies

- TCP-based - Network latency & packet loss must be low to work well
- UDP blasters – Inefficient use of bandwidth leads to congestion
- Modified TCP – Does not scale well on high-speed networks
- Data caching - Inappropriate for many large file transfer workflows
- Data compression - Time consuming & impractical for some file types



Note: Table displays throughput degradation of TCP transfers on a 1Gbps network as estimated round trip time and packet loss increases with distance.

IBM Aspera – High-performance Data Transport (Fasp®)

Maximum transfer speed

- Optimal end-to-end throughput efficiency
- Transfer performance scales with bandwidth independent of transfer distance and resilient to packet loss

Congestion avoidance and policy control

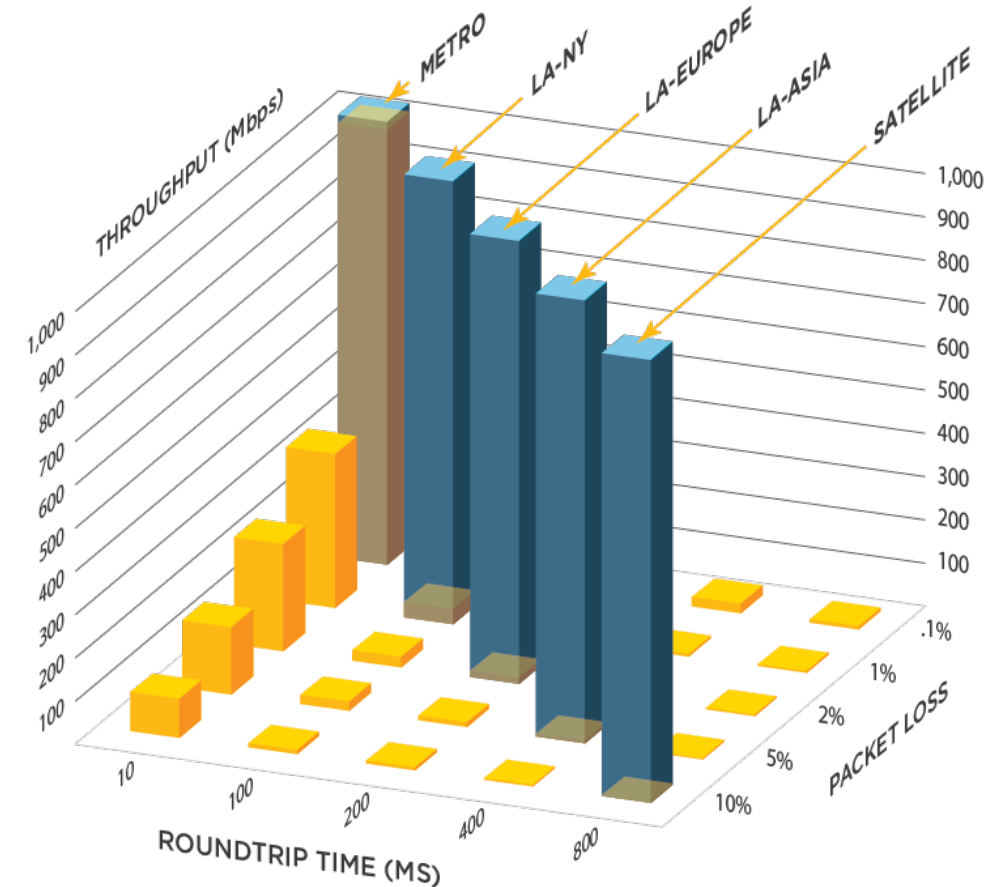
- Automatic, full utilization of available bandwidth (fair play)
- On-the-fly prioritization of transfers
- Set caps on bandwidth allocation for transfers

Uncompromising security and reliability

- Secure, SSH user/endpoint authentication
- AES-128 to 256 cryptography, FIPS 140-2 compliant
- Automatic resume of partial or failed transfers

Scalable management, monitoring and control

- Support highly concurrent transfers
- Real-time progress, performance and bandwidth utilization
- Detailed transfer history, logging, and manifest



Note: Throughput shown for a 1Gbps network. Demonstrates that bandwidth utilization for FASP transfers is immune to distance-induced latency and minimally impacted by packet loss.

IBM Aspera sync – performance breakthrough



MOVING A 10GB FILE				
	Network Bandwidth	Across US	US - Europe	US - Asia
Legacy Transport	100 Mbps	10-20 Hours	15-20 Hours	Impractical
	1 Gbps			
	10 Gbps			
Aspera FASP®	100 Mbps	14 Min	14 Min	14 Min
	1 Gbps	1.4 Min	1.4 Min	1.4 Min
	10 Gbps	8.4 Sec	8.4 Sec	8.4 Sec

Distance Independent

FASP transfer speeds remain constant as transfer distances increase while FTP speeds dramatically decrease

Predictable & Reliable

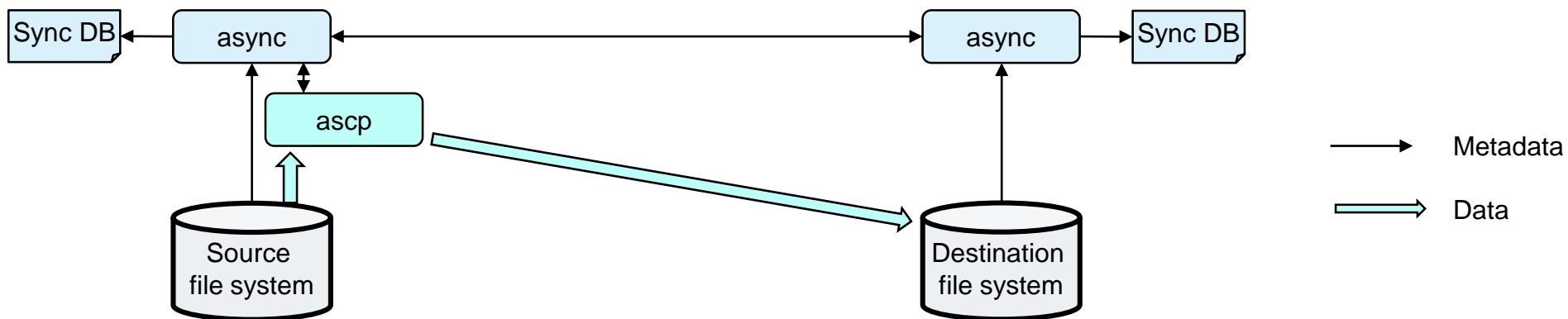
FASP transfer times decrease linearly as bandwidth increases. FTP transfer times don't improve with more bandwidth

Versatile

Supports massive file sizes (500 GB+) as easily as very large sets (millions) of small files

IBM Aspera Sync architecture

- Aspera sync includes two tools:
 - Async: file synchronization program that keeps source and destination in sync
 - Ascp: backend copy program using the patented Aspera transfer technology
- Async crawls the file system, identifies files for synchronization and keeps track of file metadata (time stamps, ACL, checksum, etc.)
 - File metadata is stored in local Sync DB (there is one local DB)
- Ascp is invoked by async to copy specific files from the source to the destination
 - Ascp does not keep track of metadata, just copies files as requested



Example: async command line

- Synchronize files from source directory */fs1* to the destination directory */target/fs1*

```
# async -N sync_fs1 -d /fs1 -r user@remotehost:/target/fs1
--preserve-xattrs=native -i ~/.ssh/id_rsa --preserve-uid --preserve-gid
--preserve-access-time [--preserve-time --create-dir]
```

- Some important options:

-N session-name	Session name (should be the same for a given replication pair)
--preserve-xattrs=native	transfers extended attributes (required for Spectrum Scale)
--preserve-uid	preserve user ID
--preserve-gid	preserve group ID
--preserve-access-time	preserve last access time stamp (important for immutable files)
--preserve-time	preserve time stamps for mtime
--create-dir	create the destination directory if this does not exist

- Session name for a given replication relation shall always be the same
 - Identifies the tables in the Sync DB storing metadata for files already been copied

Example: ascp command line

— Copy source directory */fs1* to the destination directory */target/fs1*

```
# ascp -l 1G -d -p --preserve-xattrs=native --preserve-file-owner-gid  
--preserve-file-owner-uid -i ~/.ssh/id_rsa /fs1/ user@remotehost:/target/fs1/
```

— Some important options:

--preserve-file-owner-uid	preserve user ID
--preserve-file-owner-gid	preserve group ID
--preserve-xattrs=native	preserve extended attributes (required for Spectrum Scale)
-p	preserve time stamps (important for immutable files)
-d	create destination directory if this does not exist

— The ascp tool has no notion of sessions, just copies the files give in the command line

— In general there is no need to use ascp instead of async

- Async has the intelligence to select files that need to be copied

— Both tools allow file lists as input (requires a certain format)

Agenda

Introduction to IBM Aspera sync

Integration with IBM Spectrum Scale

Differentiation with AFM

IBM Aspera sync integrates IBM Spectrum Scale

- Why is the integration with Spectrum Scale special?
 - Spectrum Scale stores ACL and immutability flags in GPFS Extended Attributes
 - These GPFS Extended Attributes are not accessible through the POSIX interface
 - Requires the GPFS API call (`gpfs_fgetattr()` and `gpfs_fputattr()`)

- IBM Aspera sync version 3.9 and above fully integrate with IBM Spectrum Scale
 - Supports synchronizing GPFS Extended Attributes including ACL and immutability flags
 - Uses `async` and `ascp` to copy the files and attributes

- Integration options with Spectrum Scale:
 - Use `async` standalone
 - Use `async` or `ascp` with EXTERNAL LIST policies

Using async standalone

- Async crawls through the file system, identifies files that need to be synched
 - Identified files are copied using ascp
 - Async keeps track of the file metadata using source and destination Sync DB
 - Async command can be scheduled (see [Automation Whitepaper](#))

- Important async command line options

<code>-N session-name</code>	Session name must be constant
<code>--preserve-xattrs=native</code>	transfers GPFS Extended Attributes
<code>--preserve-uid --preserve-gid:</code>	preserve user ID, preserve group ID
<code>--preserve-access-time --preserve-time</code>	preserve time stamps

- Async can only run on one Spectrum Scale node

- Concurrent access to Sync DB from different node for the same session name may cause errors
- Async however can be configured to run multiple transfer threads on the single node

- If crawling the file system takes too long consider using the Spectrum Scale policy

IBM Spectrum Scale policy engine

— Policy engine is a powerful tool to identify files according to certain selection criteria

— Selection criteria are encoded in policy rules

- For example to identify files that have been created or modified after a certain time stamp:

```
RULE EXTERNAL LIST 'files' EXEC ''  
RULE 'asperaRule' LIST 'files' WHERE  
MODIFICATION_TIME >= TIMESTAMP("2019-03-20 14:00:00") OR  
CHANGE_TIME >= TIMESTAMP("2019-03-20 14:00:00")
```

— Policy engine is invoked with the policy file:

```
# mmapplypolicy /fs1 -P policyfile -f /tmp/async
```

- Identified files are stored in file /tmp/async.list.files
- Format of the file list:

```
48900 1741777473 0 -- /fs1/file1
```

— File list can be used as input to async or ascp (requires some conversion)

Integrating async with Spectrum Scale policy engine

- Files identified by the policy engine can be directly passed to an external script (run-async.sh)

```
RULE EXTERNAL LIST 'files' EXEC `/usr/local/bin/run-async.sh'  
RULE 'asperaRule' LIST 'files' WHERE  
MODIFICATION_TIME >= TIMESTAMP("2019-03-20 14:00:00") OR  
CHANGE_TIME >= TIMESTAMP("2019-03-20 14:00:00")
```

- To invoke the policy engine with external script use:

```
# mmapplypolicy /fs1 -P policyfile -N asperanode -m 1 -B 1000 --single-instance
```

- This will store the file names matching the policy in file lists and pass these to the script
 - It will pack 1000 files in a file list (-B)
 - It will start 1 run-async.sh process (-m)
 - The single async process can only run on one node, asperanode (-N)
 - The run-async.sh script will now invoke async with the file list
- The external script needs to be implemented to invoke async with the file list

The external script

— External script is invoked by the policy engine with the following parameters:

1. Parameter: string describing the operation (TEST and LIST)
 - TEST is invoked first and allows the script to perform some checking
 - LIST will pass the file lists in the second parameter

2. Parameter: depends on 1. Parameter.
 - If the first parameter is LIST then the second parameter is the name of the file list
 - If the first parameter is TEST then the second parameter is the name of the file system.

3. Parameter (optional): passes the options encoded in the EXTERNAL LIST rule

— External script has to process the parameters and invoke async:

```
case $1 in
  (TEST)  check if file system in $2 is mounted & exit;;
  (LIST)  convert the file list to the format used by async
          run async with the file list
          exit with return code
esac
```

Using async vs ascp

- Instead of async, ascp can be used in the external script
 - Ascp has less overhead and might be faster
 - Ascp does not update the Sync DB, so subsequent use of async would transfer all files again
 - Ascp could be run on multiple cluster nodes using the policy engine (-N nodes)

- To remove deleted files from the destination ascp must be run with a special options:
 - `--delete-before-transfer`
 - `--overwrite={diff|diff+older}`
 - This assures that files that failed transfer before are transferred now

Considerations and guidance

- Use async standalone if the file system crawling does not impact performance
 - This method ensures that all files are synched and file deletions are propagated

- When using the integration with the policy engine then run async standalone periodically
 - Use the same session name
 - Assures that files that may have failed the transfer are transferred
 - Assures that files that are deleted on the source are deleted on the destination

- Integrating async with the policy engine allows to exclude migrated files and prevent recall storms

- Async also copies immutability flags for immutable files
 - Immutable filesets have to be created manually on the destination prior to the synchronization

- Async can use a snapshot on the source

Agenda

Introduction to IBM Aspera sync

Integration with IBM Spectrum Scale

Differentiation with AFM

AFM key characteristics

- AFM is integrated with IBM Spectrum Scale Data Management Edition license
 - Configuration, operation and monitoring is included in IBM Spectrum Scale
- AFM tracks changed blocks and does not necessarily transfer the entire file again
 - Transfers extended attributes except of immutability attributes
 - Compressed files on the source can be compressed on the target
- AFM can be configured for high availability with two or more AFM gateways
 - Tolerates network outages and automatically resumes the replication operation if the network is available.
- AFM recovers network outages automatically
- AFM DR has integrated processes and tools to facilitate failover and failback
 - Manages snapshot on source and destination, providing a consistent recovery point
- AFM DR has limited support with storage tiering to tape or cloud
- AFM requires fileset as source and target directories

Aspera sync characteristic

- Aspera sync is made for long distance data transfer
 - Guarantees performance regardless of transfer distance, file size or volume, and network conditions
- Aspera sync supports the transfer of extended attributes including immutability attributes
 - Uses checksum to assure file integrity
- Aspera sync can be integrated with storage tiering to tape or cloud
 - Integration with Spectrum Scale policy engine allows to exclude migrated files
- Aspera sync does not require filesets on the source and the target directory.

- Aspera sync does not create snapshots and does not provide consistent recovery points
- Aspera sync standalone crawls through the file system
- Aspera sync is not cluster aware and can only run on one cluster node
- Aspera sync sessions have to be scheduled and monitored by the administrator
- Aspera sync does not track changed blocks and transfers the file entirely again
- Aspera sync does not recover network outages automatically, session has to be manually restarted

Guidance

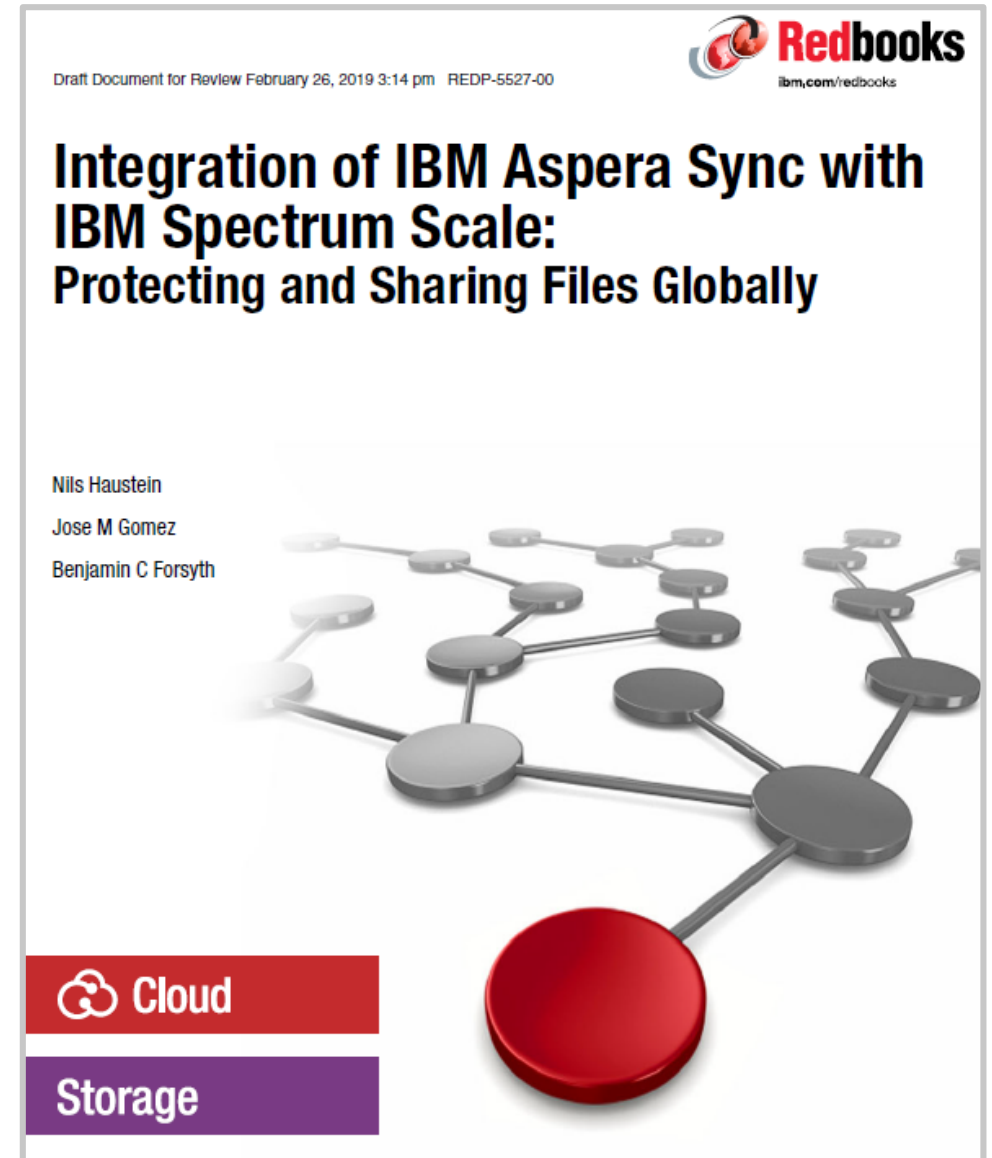
- Using AFM is simpler, because it is fully integrated with Spectrum Scale

- Use Aspera sync when AFM is not appropriate
 - Transfer of immutable files
 - Integration with storage tiering to tape or cloud
 - Long distance file transfer
 - For high workloads on the source, where AFM may get out of sync with changed block tracking
 - Tool to copy files and their attributes from one Spectrum Scale file system to another

Learn more

- New IBM Redpaper describes the integration of IBM Aspera sync with IBM Spectrum Scale
 - Describes Aspera sync tools (async and ascp) and important command line options
 - Describes the integration of Aspera sync with Spectrum Scale policy engine
 - Describes use cases
 - Differentiates the solution from Spectrum Scale AFM

<http://www.redbooks.ibm.com/abstracts/redp5527.html>



Thank You

References

Redpaper: Integration of IBM Aspera with IBM Spectrum Scale.

<http://www.redbooks.ibm.com/abstracts/redp5527.html>

Aspera sync command reference version 3.9 (Linux):

https://download.asperasoft.com/download/docs/entsrv/3.9.0/cs_admin_linux/webhelp/index.html#dita/sync.html

Aspera ascp command reference version 3.9 (Linux):

https://download.asperasoft.com/download/docs/entsrv/3.9.0/cs_admin_linux/webhelp/index.html#dita/ascp_2.html

Github repository for Spectrum Scale policy integration

<https://github.com/IBMRedbooks/Integration-of-IBM-Aspera-Sync-with-IBM-Spectrum-Scale>

Spectrum Scale mmapplypolicy command:

https://www.ibm.com/support/knowledgecenter/STXKQY_5.0.1/com.ibm.spectrum.scale.v5r01.doc/bl1adm_mmapplypolicy.htm

Practical guide for using Spectrum Scale policies

<http://w3-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP102642>

Notices and disclaimers

- © 2018 International Business Machines Corporation. No part of this document may be reproduced or transmitted in any form without written permission from IBM.
- **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 per 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 follows any law.

Notices 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 about 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 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 and [names of other referenced IBM products and services used in the presentation] 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.
- .