Nils Haustein | Executive IT Specialist | EMEA Storage Competence Center

IBM Spectrum Protect

Spectrum Protect for Space Management and Spectrum Protect Scalable Backup for Spectrum Scale

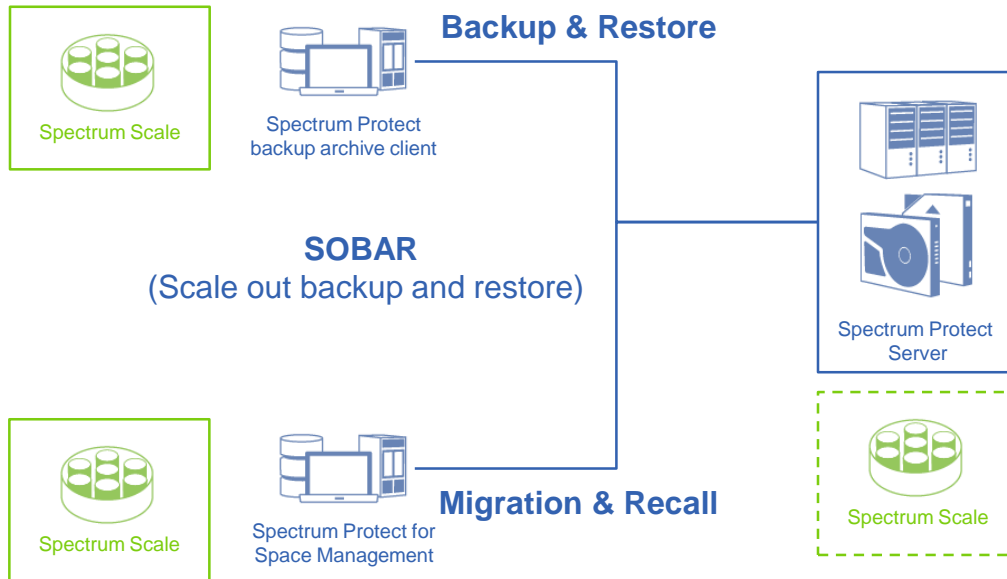**By Dominic Mueller-Wicke**
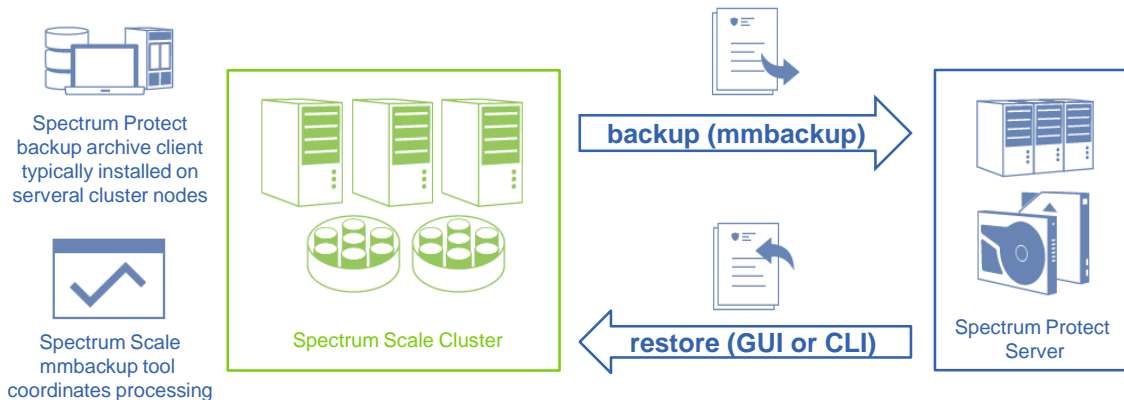**(IBM Spectrum Protect Development)**

IBM

Spectrum Protect – Spectrum Scale Integration

Overview

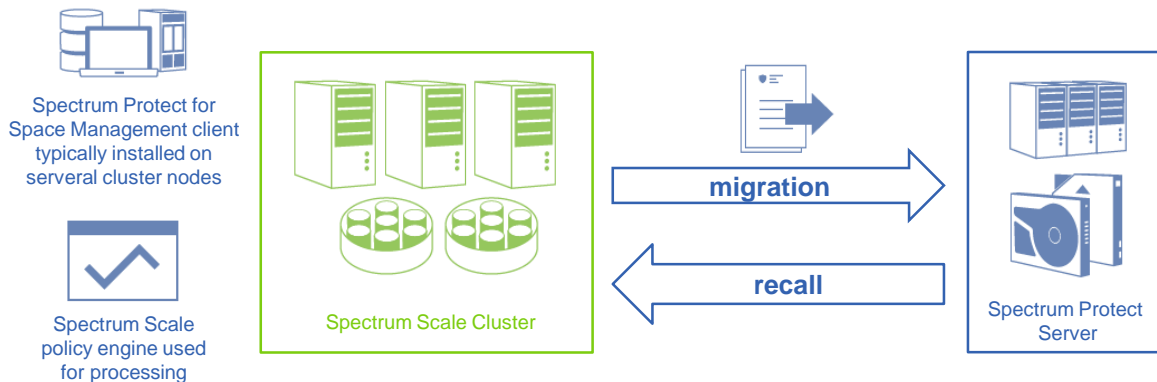# Spectrum Protect / Spectrum Scale Integration Overview

Scalable Backup And Restore

Spectrum Protect
backup archive client
typically installed on
serveral cluster nodes

Spectrum Scale
mmbackup tool
coordinates processing

Spectrum Scale Cluster

backup (mmbackup)

restore (GUI or CLI)
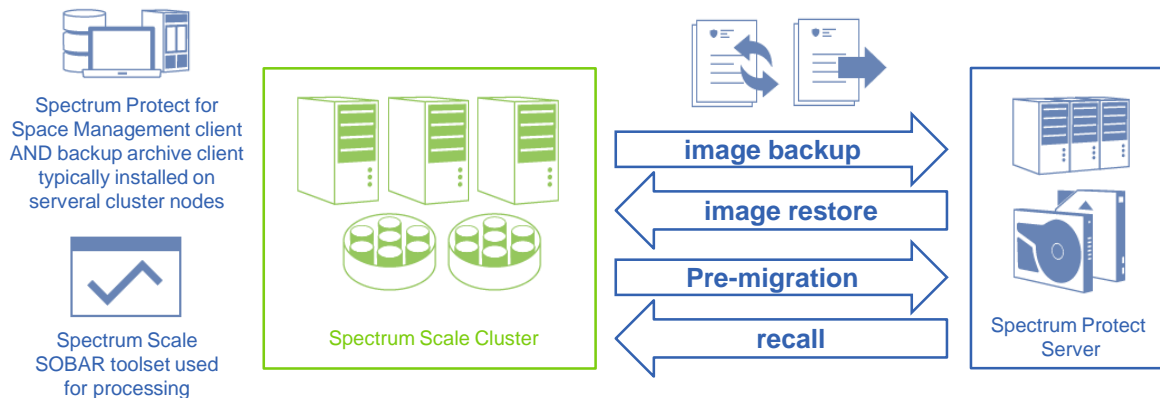
Spectrum Protect
Server

- Massive parallel filesystem backup processing
  - Clusterwide distribution of backup processing
  - Uses policy engine to identify files
  - Uses Spectrum Protect backup archive client under the cover
- Spectrum Protect restore (CLI or GUI) can be used to restore files

Hierarchical Storage Management

Spectrum Protect for Space Management client typically installed on serveral cluster nodes

Spectrum Scale policy engine used for processing

Spectrum Scale Cluster

migration

recall

Spectrum Protect Server

- Massive parallel migration processing to efficiently reduce file system occupation
  - Use of policy engine to identify files to be migrated
  - Cluster wide distributed parallel recall of files in file specific mode
  - Tape optimized recall for mass recalls in batch mode
- Close integration with Spectrum Protect Backup Archive Client
  - Migration requires backup
  - Inline backup for migrated files

Scale out Backup and Recovery

Spectrum Protect for
Space Management client
AND backup archive client
typically installed on
serveral cluster nodes

Spectrum Scale
SOBAR toolset used
for processing

Spectrum Scale Cluster

image backup

image restore

Pre-migration

recall

Spectrum Protect
Server

- Combines backup and HSM function
  - Backup is used to backup configuration data and metadata (image dump)
  - HSM is used to premigrate or migrate files continuously
- Provides ultra-fast recovery for entire file system
  - Restore and apply configuration data and metadata
  - Recall files transparently and in tape optimized manner

Scale Out Backup And Restore (SOBAR)

Level of file and file system recovery using IBM Spectrum Protect

Small number of files (Full Restore):
➢ dsmc restore –restormigstate=no

Huge number of files or full directory trees (Stub Restore):
➢ dsmc restore –restoremigstate=yes
or
➢ dsmc restore –dirsonly
➢ dsmmigundelete

Full filesystem recovery (SOBAR Restore):
➢ SOBAR

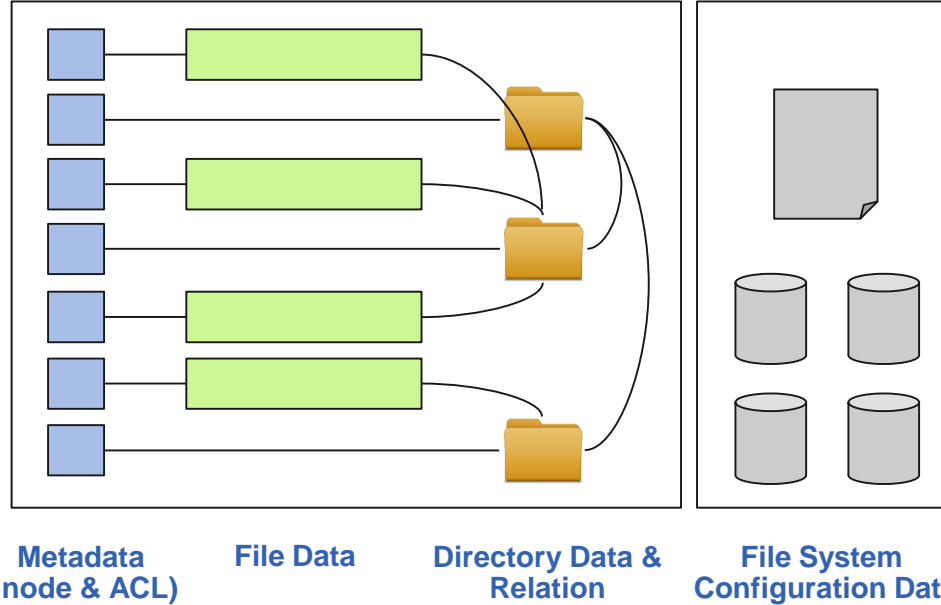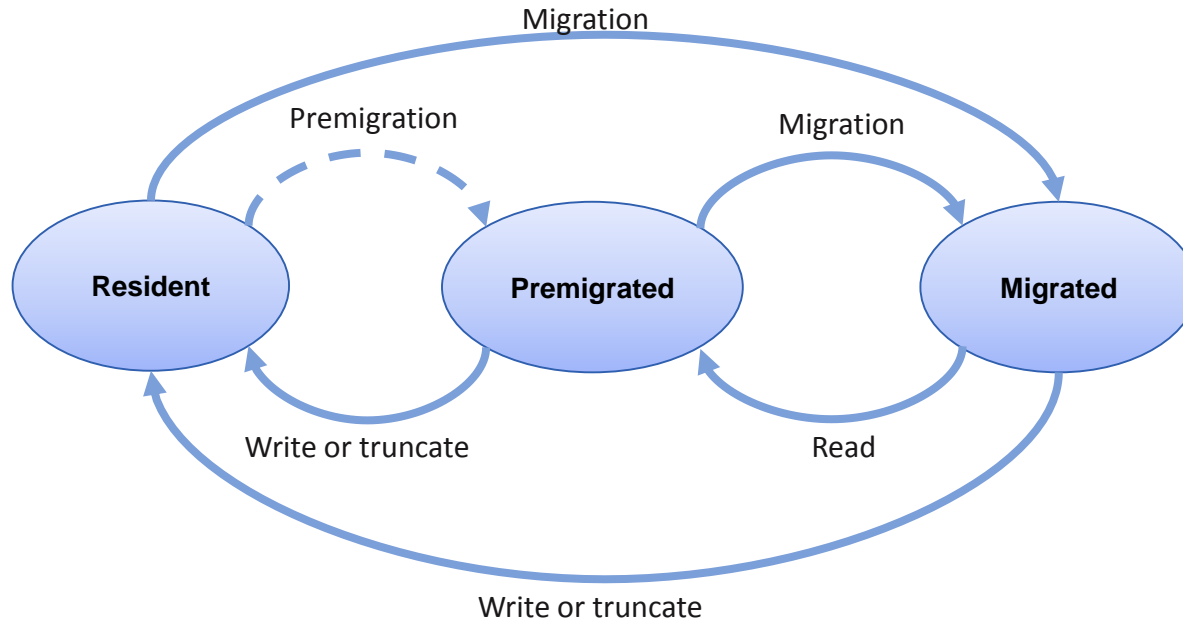# Level of file and file system recovery using IBM Spectrum Protect

Restore Type	Restored Content	Comments
Full Restore	• directories • file data • file metadata (POSIX/ACL/EA)	• full data transfer over network
Stub Restore	• directories • file metadata (POSIX)	• Quick, stubs only • next backup requires recall of files • Does not recreate ACL/EA
SOBAR Restore	• directories • file metadata (POSIX/ACL/EA)	• fast • scalable

Scale Out Backup And Restore – Protected Items

**Metadata (Inode & ACL)**   **File Data**   **Directory Data & Relation**   **File System Configuration Data**

Space Management file states and transitions

Scale Out Backup And Restore – Preparation

Preparation:

- premigrate (or migrate) all file data using Spectrum Protect for Space Management and Spectrum Scale policy engine

- Optional: backup all file and directory data using Spectrum Scale mmbackup and Spectrum Protect Backup Archive Client
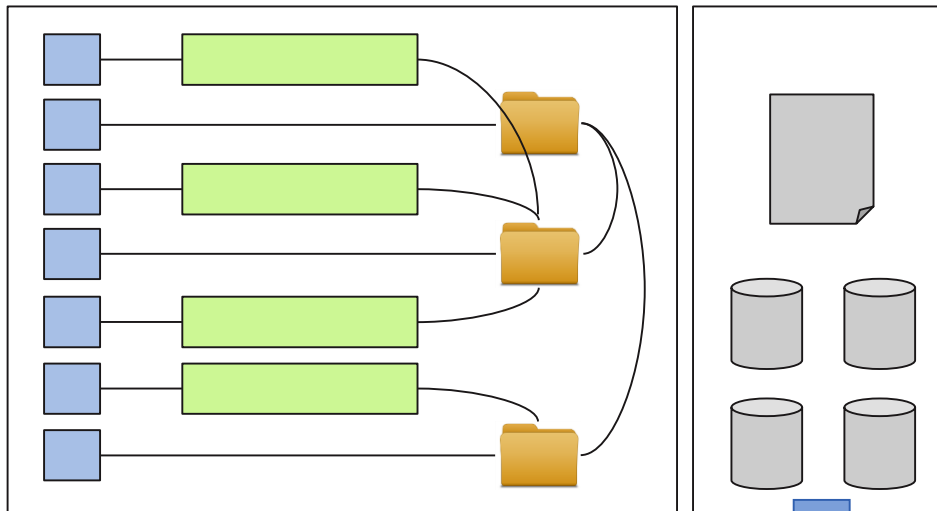

IBM Spectrum Protect Server
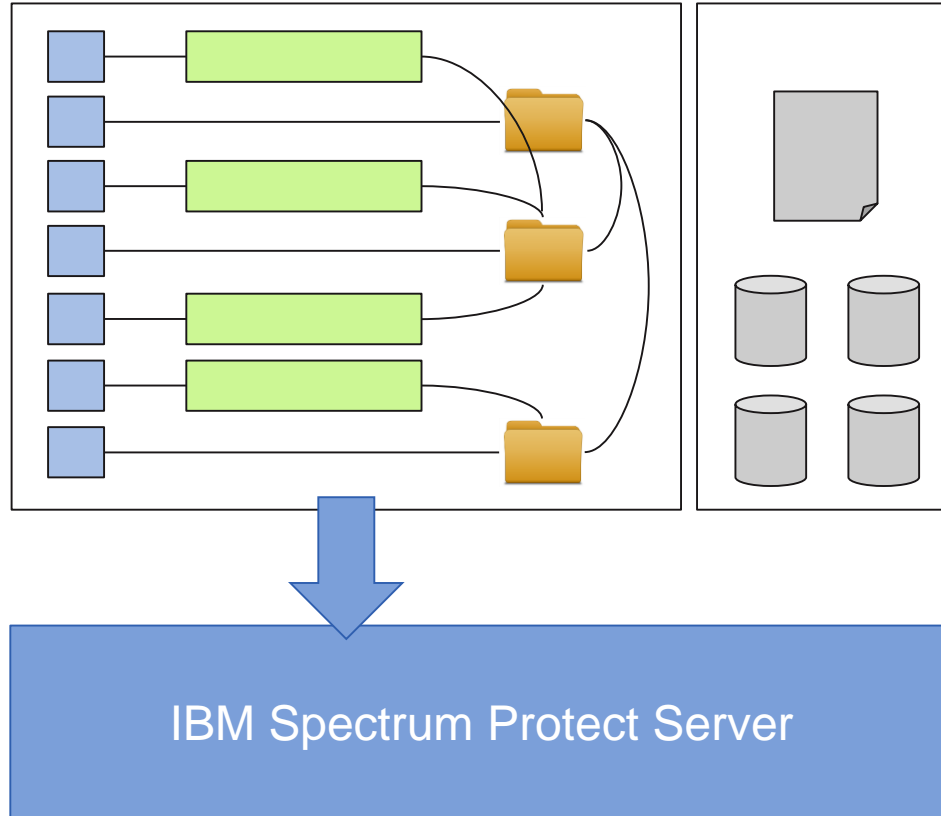
Scale Out Backup And Restore – Preparation

Preserve cluster
configuration data:

- IP addresses
- node names
- roles (quorum and
  server roles)
- cluster-wide
  configuration settings
  (mmchconfig)
- cluster manager node
  roles
- remote shell
  configuration
  authentication setup
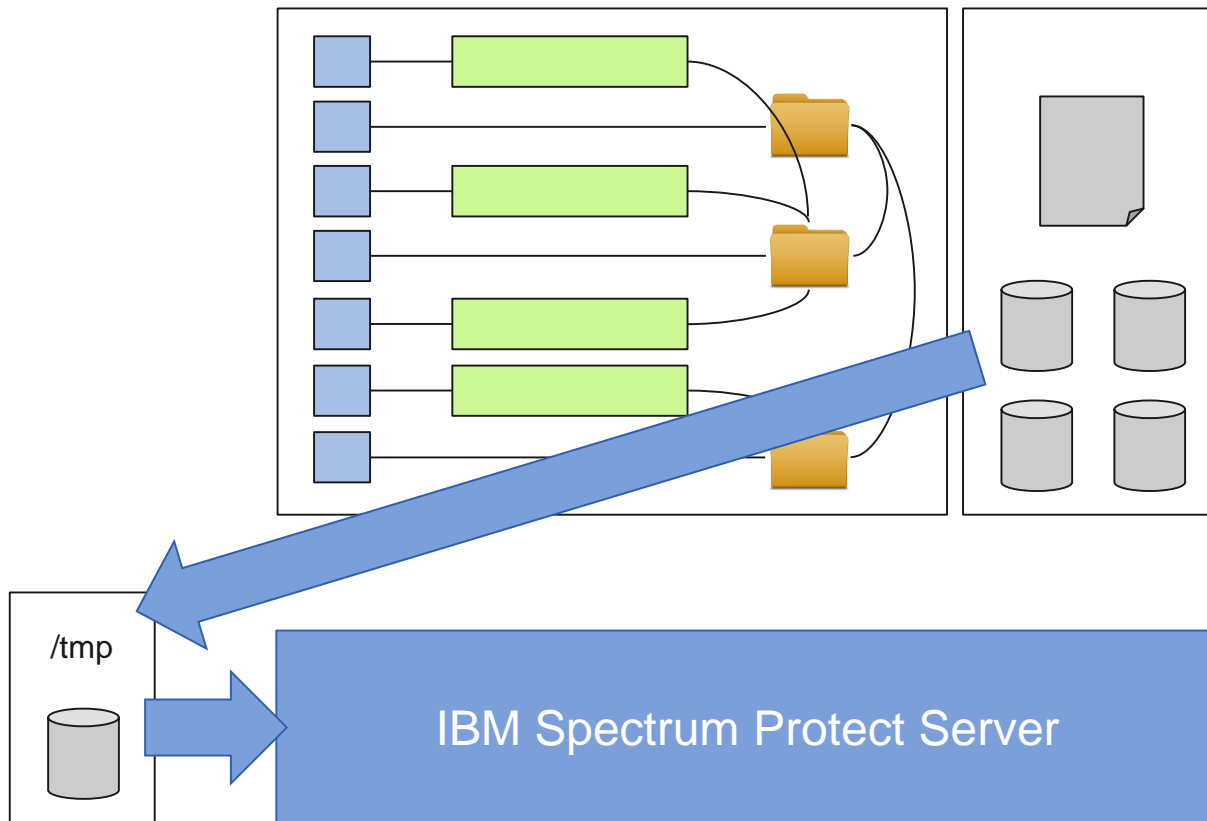- cluster UID

Potentially: mmsdrfs file
content.

Preserve disk
configuration data:

- Disk device names
- Disk device sizes
- The number of disk
  volumes
- NSD server
  configuration
- Disk RAID
  configurations
- Failure group
  designations

Potentially: mmsdrfs
file content

IBM Spectrum Protect Server

Scale Out Backup And Restore – Backup I

Backup file system volume configuration using

mmbackupconfig

Includes:
- block size
- replication factors
- number and size of disks
- storage pool layout filesets and junction points
- policy rules
- quota information
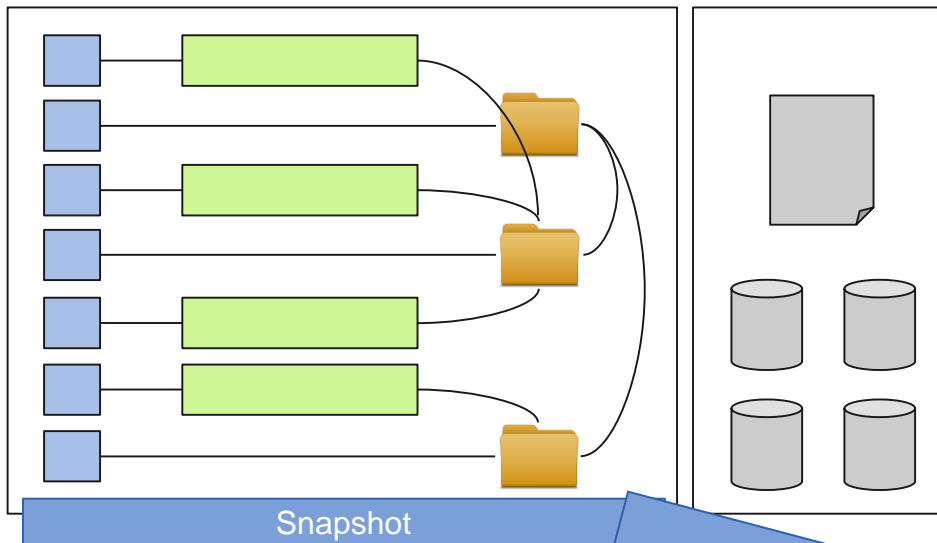- other file system attributes

/tmp

IBM Spectrum Protect Server

Scale Out Backup And Restore – Backup II

1 Create a global snapshot of the live file system, using command:

mmcrsnapshot

3 After the image backup is complete, delete the snapshot used for backup with command:
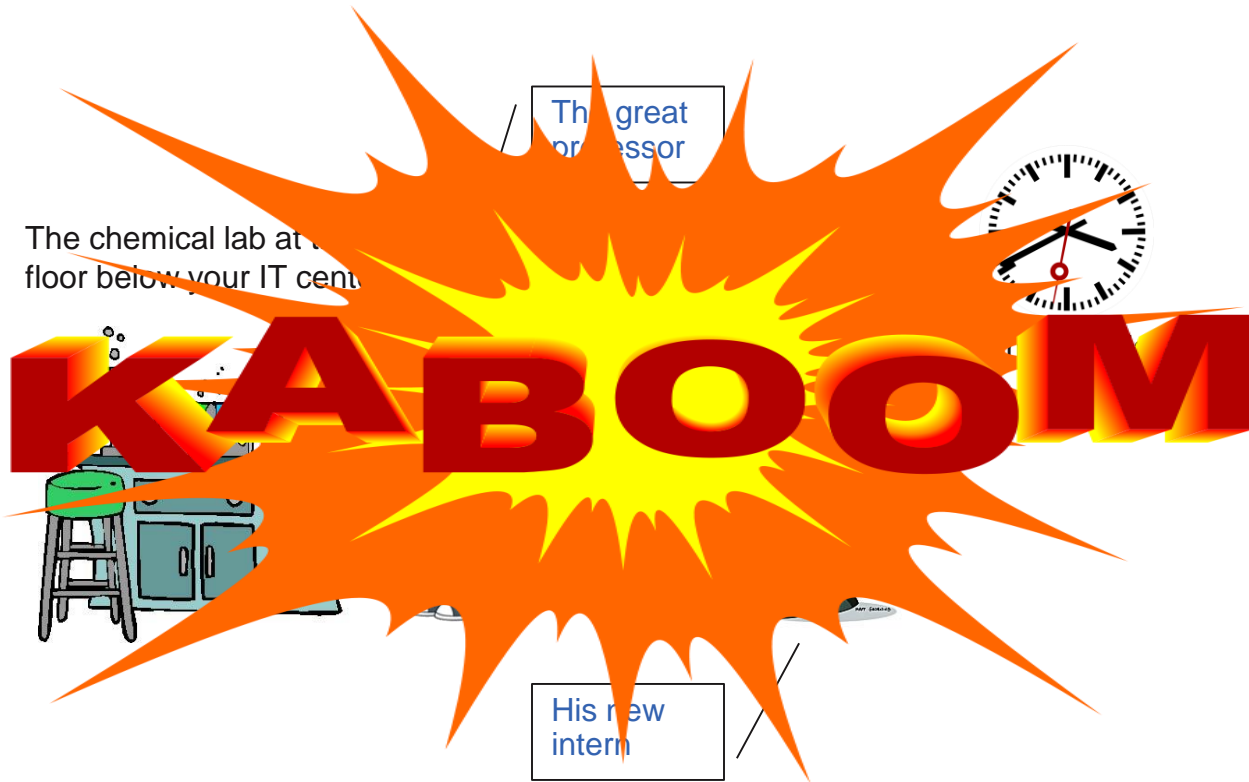
mmdelsnapshot

Snapshot

2 Backup the file system image, using command

mmimgbackup

/tmp

IBM Spectrum Protect Server

The chemical lab at [the]
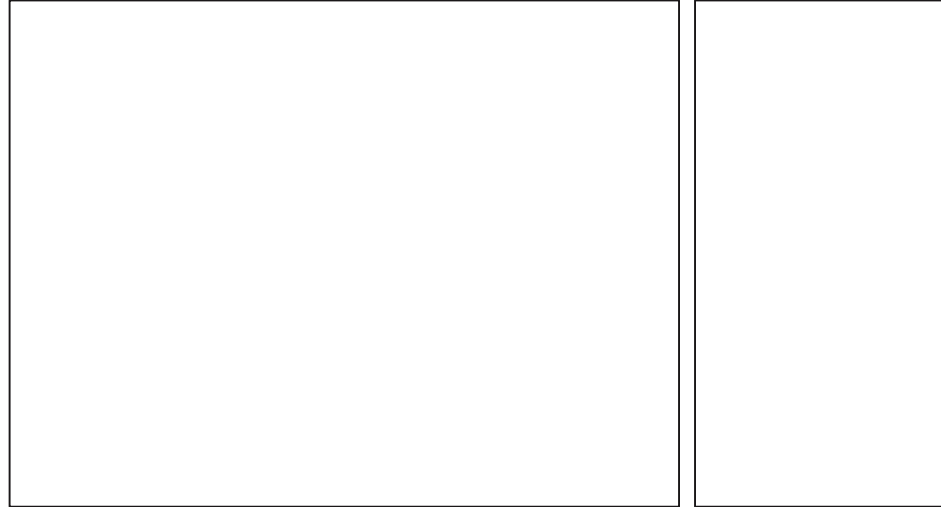floor below your IT cent[er]

The great
pr[of]essor

His new
intern

Scale Out Backup And Restore – Restore Overview

A Rebuild Spectrum Scale cluster and file system disk based on preserved information.
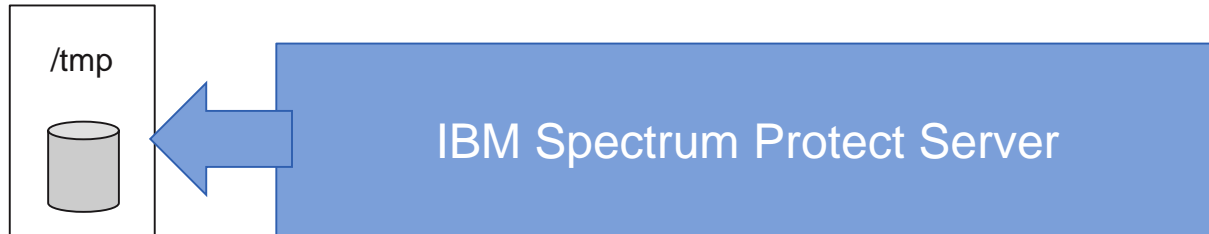
C Restore file system image files using Spectrum Protect Backup Archive Client

B

Retrieve the base file system configuration information to generate a configuration file, which contains the details of the former file system. Using command:

mmrestoreconfig [-F]

/tmp

IBM Spectrum Protect Server

Scale Out Backup And Restore – Restore I

1 Recreate NSDs if they are missing, using the output file generated in the previous step using command:

mmcrnsd

2 Recreate the base file system. The output query file specified in the previous commands can be used as a guide using command:

mmcrfs

/tmp

IBM Spectrum Protect Server
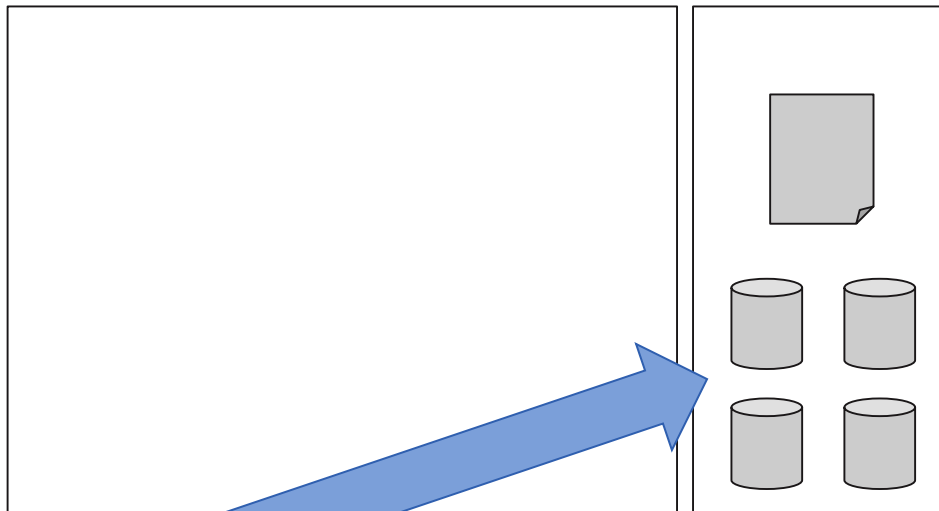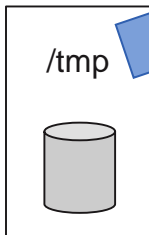
3 Restore essential file system configuration using command:

mmrestoreconfig

This recreates the essential file system configuration. Filesets will also be restored and linked.

Scale Out Backup And Restore – Restore II

4

Mount the file system in read-only mode for image restore using command:

mmmount -o ro

Perform the image restore using command:

mmimgrestore

/tmp

IBM Spectrum Protect Server

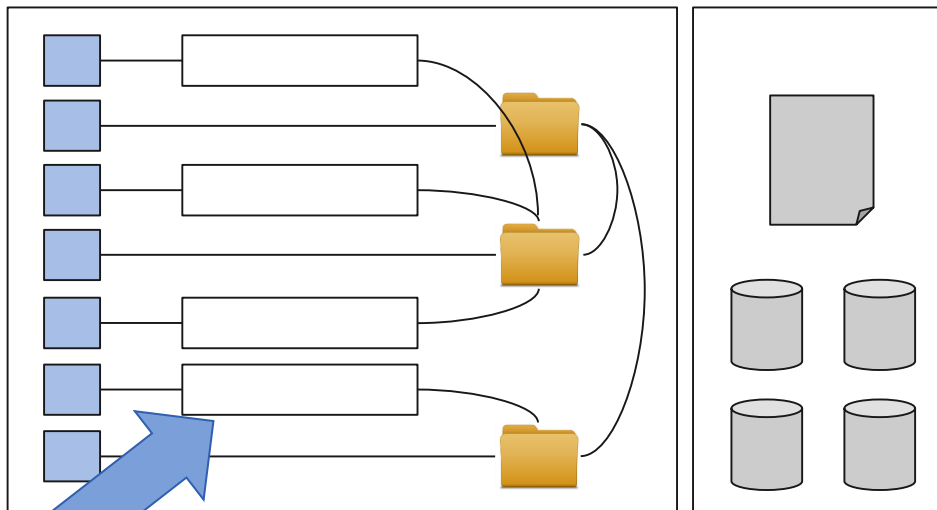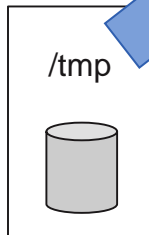**5**

Unmount the file system with the following command:

mmumount

Restore quota configuration using command:

mmrestoreconfig

Scale Out Backup And Restore – Restore III

6

Mount the file system in read-write mode with the following command:

mmmount

Delete the unusable HSM directory.

rm -rf /fs/.SpaceMan
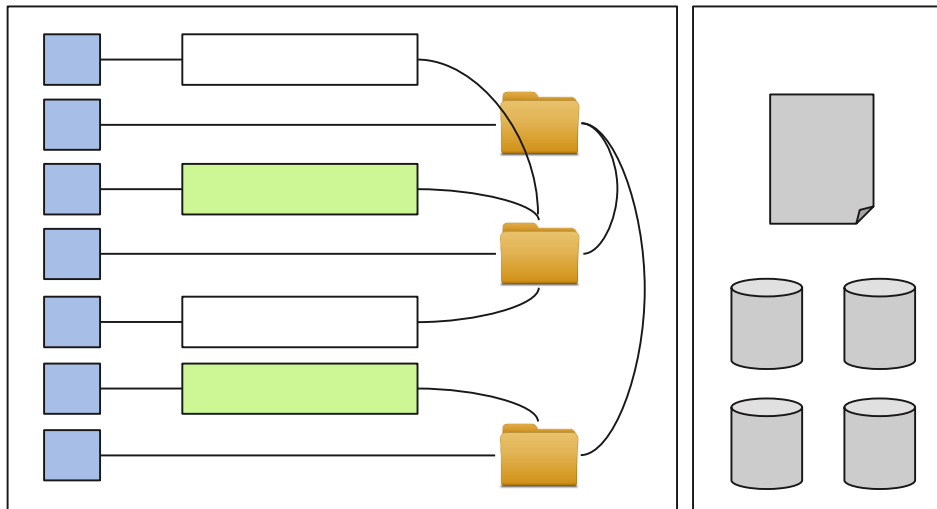
IBM Spectrum Protect Server

7

Restart HSM, using command:

dsmmigfs restart

Resume HSM management on the newly reconstructed file system using command:

dsmmigfs add

All files are in the migrate state. Recall files using command:

dsmrecall

Spectrum Protect Scalable Backup and Spectrum Protect for Space Management

Hints and Tips

Backup

- Use latest versions: Spectrum Scale 4.x.x and Spectrum Protect 7.1.x have good improvements for mmbackup

- Consider additional workload on Spectrum Scale cluster nodes
    – mmbackup puts load on each node involved
    – Backup nodes need access to storage and Spectrum Protect server with enough bandwith

- Serialize backups of different file systems

    Example:
    Cluster has nodes N1 to N9 and files systems FS1 to FS6
    Run mmbackup on nodes N1,N2,N3 for FS1 and on nodes N4,N5,N6 for FS2, ... in parallel
    After FS1 has finished run mmbackup on nodes N1,N2,N3 for FS3, ...

- Do not mix platforms. Run mmbackup on either AIX, xLinux, pLinux, zLinux nodes
- Do not mix platforms for backup and restore.
- Rebuild of the shadowDB takes time. Use option –q or –rebuild as seldom as possible.

Backup

- Consider Spectrum Protect character limitations (especially for Windows environments)
    - Files with control-X, control-Y, carriage return and the new line character in their name cannot be backed
    - Use `QUOTESARELITERAL` (if mmbackup is used with --noquotes), if file names contain " or ' .
    - Use `WILDCARDSARELITERAL`, if file names contain * or ?.

- Do not use Spectrum Protect processing options:
    - `SUBDIR YES,`
    - `QUIET,`
    - `SCROLLPROMPT,`
    - `SCROLLLINES`

- Check your log files on a regular base
- Prevent move or rename of files or directories. These changes lead to a new backup of all affected files
- If ACL or EA metadata is used prevent changes of POSIX attribute.
    - Leads to a new backup of all affected files

Backup

Align the mmbackup, Spectrum Protect client and server options

Define number of bytes transferred in a single backup transaction

`--max-backup-size` ——————— (N *) ——————— `TXNBYTELIMIT`

Define number of files transferred in a single backup transaction

`--max-backup-count`
`--max-expire-count` ——————— (N *) ——————————————— `TXNGROUPMAX`

Define number of parallel backup processes

`--expire-threads`
`--backup-threads` ——— (* #Nodes *) `RESOURCEUTILIZATION` ——— `MAXSESSIONS`
`MAXNUMMP`

Spectrum Scale
mmbackup command
line options

Spectrum Protect **backup archive
client** processing options

Spectrum Protect
Server processing options

Backup - Include and Exclude

- Spectrum Protect offers a rich set of include and exclude options to control which files and directories are backed up.

- Spectrum Scale mmbackup translates these options into policy rules for backup
  – Translation is complex
  – Include and Exclude options may have significant impact on scan performance

- Some rules to consider:
  – Use as few EXCLUDE statements as possible
  – Aviod using INCLUDE. Use EXCLUDE instead
  – Do not use „EXCLUDE /dir/.../*" . Try EXCLUDE.DIR instead.
  – Avoid EXCLUDE and INCLUDE for the same subtree, like
    ```
    exclude /home/dominic*
    include /home/dominic/important*
    ```
  – If INCLUDE is only used to assign right management class in Spectrum Protect
    ```
    „INCLUDE <pattern> MGMT"
    ```
    use mmbackup service flag is used MMBACKUP_IGNORE_INCLUDE
    ```
    export MMBACKUP_IGNORE_INCLUDE=1
    ```

- Technote on this theme: http://www-01.ibm.com/support/docview.wss?uid=swg21699569

Space Management

- General recommendations:
  - Do not mix platforms with backup and Space Management. Run both application on either AIX, xLinux, zLinux nodes.
  - Use policy driven migration. See whitepaper: http://www-01.ibm.com/support/docview.wss?uid=swg27018848
  - Share workload between cluster nodes and serialize processing with backup.

- Migration performance:
  - Use processing option HSMGROUPEDMIGRATE YES to improve migration performance to tape
  - Align the settings for mmapplypolicy command arguments –m with Spectrum Protect option MAXSESSIONS and MAXMUNMP.
  - Align the settings for mmapplypolicy command arguments –B with Spectrum Protect option TXNBYTELIMIT and TXNGROUPMAX.
- Recall performance:
  - Use tape optimized recall if it can be combined with business processes.

- Data protection and recovery:
  - Do not set processing option MIGFILEEXPIRATION to 0. Recovery of expired files will be impossible.
  - Space Management is not backup. Use option MIGREQUIRESBACKUP YES.

Space Management and backup integration consideration

- Spectrum Protect for Space Management migrated files can become eligible for backup if:
  - The file has ACL/EA data and this metadata changes
  - The file has ACL/EA data and a POSIX attribute changes
  - (File data changes)
- Backup leads to random access of files in Spectrum Protect server. Potentially one tape mount per file to be recalled
  - Worst case: rename of file system root directory ☹

- Spectrum Scale 4.2 mmbackup skips migrated files from backup processing
  - A file list is created that includes all migrated files eligible for backup
  - File list can be used perform tape optimized recall prior to the next backup

- Future: mmbackup uses tape optimized recall internally

Latest releases

Spectrum Protect 7.1.4 (4Q15)

- Spectrum Protect Backup Archive Client is available on IBM Power 8 Little Endian
  - Full functional support for Spectrum Scale mmbackup
- Spectrum Protect for Space Management is available on IBM System z Linux.
  - Full functional support for Spectrum Scale
  - Full integration with backup equal to xLinux
- Spectrum Protect for Space Management pre/migration of zero length files
  - Spectrum Scale stores file smaller than Inode size inside inode. Number of blocks is zero for these files
  - These files can be migrated now to have it in place for SOBAR restores

Spectrum Protect Server sizing

- Use IBM Spectrum Protect blueprint v2.3 for reference:
  *ibm.biz/IBM**SpectrumProtectBlueprints***

- Blueprints describe server and storage configuration based on daily and total data volumes to be managed

- Latest version describes configurations up to 4PB leveraging ESS GL4

- Based on new container pools

References

IBM Knowledge Center
IBM Spectrum Scale: http://www.ibm.com/support/knowledgecenter/STXKQY/ibmspectrumscale_welcome.html
IBM Spectrum Protect: http://www.ibm.com/support/knowledgecenter/SSGSG7/landing/welcome_ssgsg7.html

IBM Spectrum Protect resources landing page
http://www.ibm.com/support/docview.wss?uid=swg21684850

Overview on Spectrum Protect – Spectrum Scale Integration
https://www.ibm.com/developerworks/community/wikis/home/wiki/Tivoli%20Storage%20Manager/page/Integrating%20IBM%20Tivoli%20Storage%20Manager%20with%20IBM%20Elastic%20Storage

Configuration of Spectrum Protect for Spectrum Scale AFM
https://www.ibm.com/developerworks/community/wikis/home/wiki/Tivoli%20Storage%20Manager/page/Configuring%20IBM%20Spectrum%20Scale%20Active%20File%20Management

Spectrum Protect for Space Manangement whitepaper
Setup policy driven threshold migration: http://www.ibm.com/support/docview.wss?uid=swg27018848
Setup cross platform cluster: http://www.ibm.com/support/docview.wss?uid=swg27028178

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.

Important notes:

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 makes no warranties, express or implied, regarding non-IBM products and services, and any implied warranties of merchantability and fitness for a particular purpose. IBM makes no representations or warranties with respect to non-IBM products. Warranty, service and support for non-IBM products is provided directly to you by the third party, not IBM.

All part numbers referenced in this publication are product part numbers and not service part numbers. Other part numbers in addition to those listed in this document may be required to support a specific device or function.

When referring to storage capacity, GB stands for one billion bytes; accessible capacity may be less. Maximum internal hard disk drive capacities assume the replacement of any standard hard disk drives and the population of all hard disk drive bays with the largest currently supported drives available from IBM.

IBM Information and Trademarks

The following terms are trademarks or registered trademarks of the IBM Corporation in the United States or other countries or both: the e-business logo, IBM, system x, system p, System Storage, GPFS

Microsoft Windows is a trademark or registered trademark of Microsoft Corporation.

Linux is a registered trademark of Linus Torvalds.

Other company, product, and service names may be trademarks or service marks of others.