

---

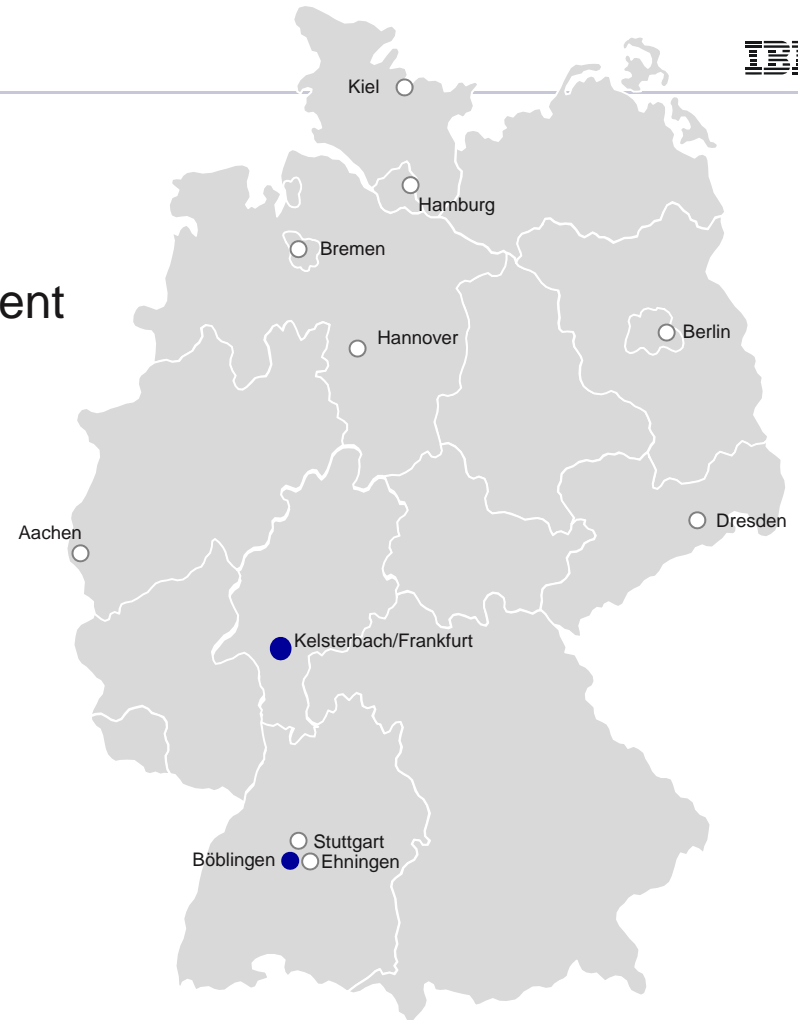
# Spectrum Protect Backup integration with Spectrum Scale

Spectrum Scale User Group Spring 2017  
Manchester UK – May 2017

Fabián Kuhl  
Teamlead IBM Spectrum Protect Research and Development Lab Frankfurt/Kelsterbach

● IBM Systems - Research and Development  
Lab in Kelsterbach/Frankfurt

- IBM Spectrum Protect clients
  - Backup/Archive
  - Space Management (HSM)
  - VE integration
  - Spectrum Scale and Archive integration



## Disclaimer

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.

Information about potential future products may not be incorporated into any contract.

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

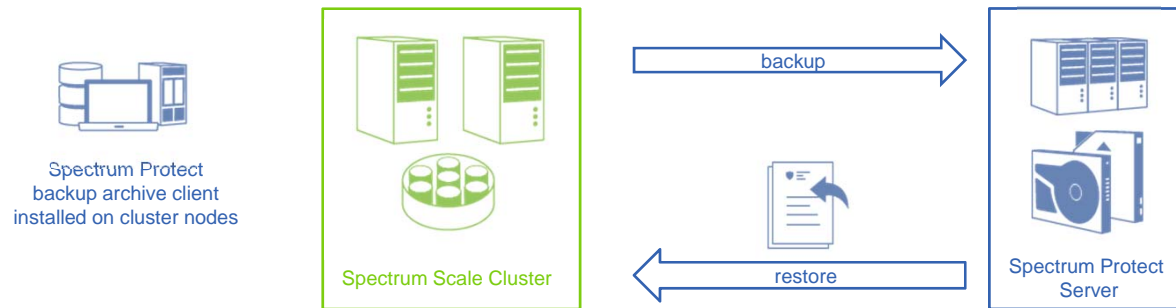
## Agenda

- Overview IBM Spectrum Protect clients
  - Backup/Archive
  - Space Management
  
- IBM Spectrum Protect Data Protection for Spectrum Scale
  - Tools like SOBAR, MMBACKUP, DSMMIGUNDELE
  - Backup Approaches

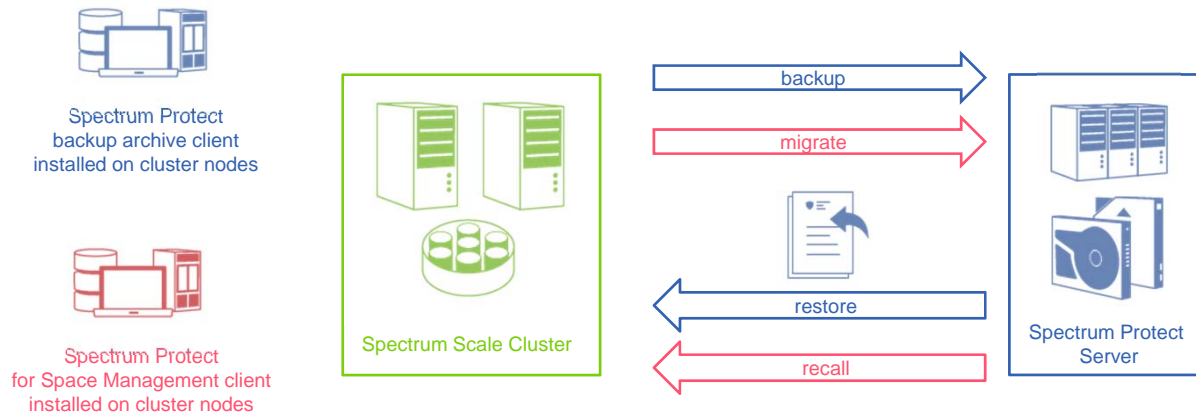
## Agenda

- Overview IBM Spectrum Protect clients
  - Backup/Archive
  - Space Management
- IBM Spectrum Protect Data Protection for Spectrum Scale
  - Tools, like SOBAR, MMBACKUP, DSMIGUNDELE
  - Backup Approaches

# Overview IBM Spectrum Protect clients



# Overview IBM Spectrum Protect clients



## Agenda

- Overview IBM Spectrum Protect clients
  - Backup/Archive
  - Space Management
- IBM Spectrum Protect Data Protection for Spectrum Scale
  - Tools like SOBAR, MMBACKUP, DSMMIGUNDELE
  - Backup Approaches



## Overview IBM Spectrum Protect client tools

- Restore options with tools dsmmigundelete, dsmc, SOBAR



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**):

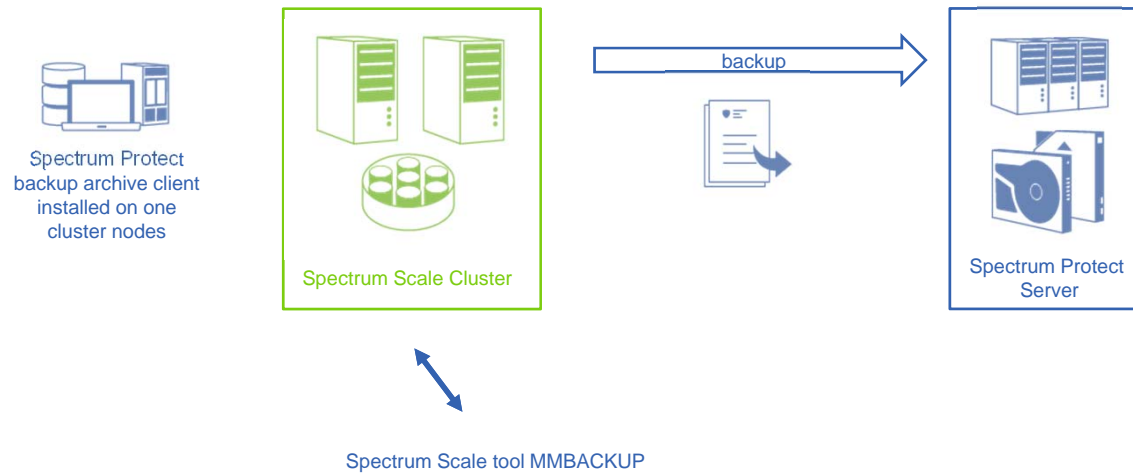
- SOBAR

## Overview IBM Spectrum Protect clients

### - Restore options

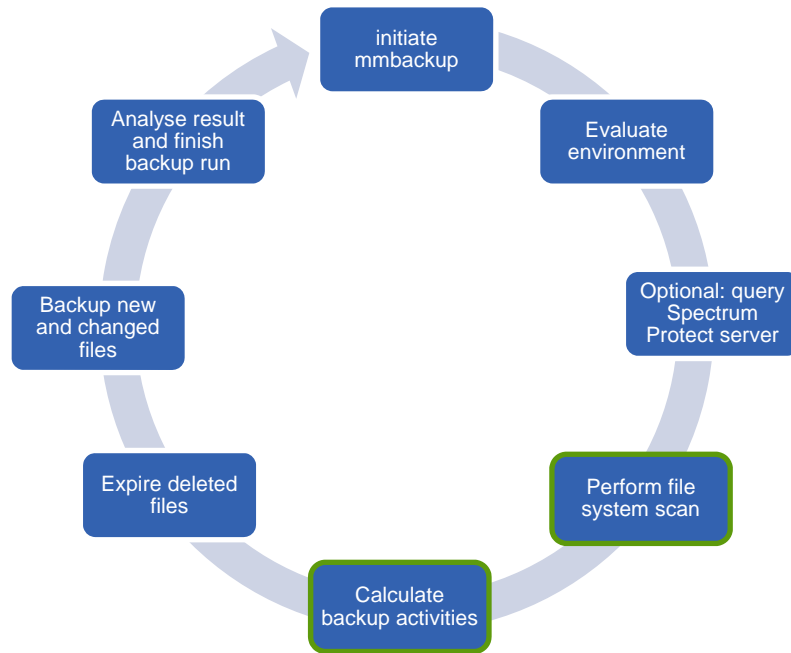
| Restore Type  | Restored Content   | Comments  |
|---------------|--|---|
| Full Restore  | <ul style="list-style-type: none"> <li>directories</li> <li>file data</li> <li>file metadata (POSIX/ACL/EA)</li> </ul> | <ul style="list-style-type: none"> <li>full data transfer over network</li> </ul>   |
| Stub Restore  | <ul style="list-style-type: none"> <li>directories</li> <li>file metadata (POSIX)</li> </ul>                           | <ul style="list-style-type: none"> <li>quick</li> <li>next backup requires recall of files</li> <li>reset of ACL/EA required</li> </ul> |
| SOBAR Restore | <ul style="list-style-type: none"> <li>directories</li> <li>file metadata (POSIX/ACL/EA)</li> </ul>                    | <ul style="list-style-type: none"> <li>fast</li> <li>scalable</li> </ul>  |

# IBM Spectrum Protect Data Protection for Spectrum Scale - MMBACKUP



# IBM Spectrum Protect Data Protection for Spectrum Scale

## - MMBACKUP

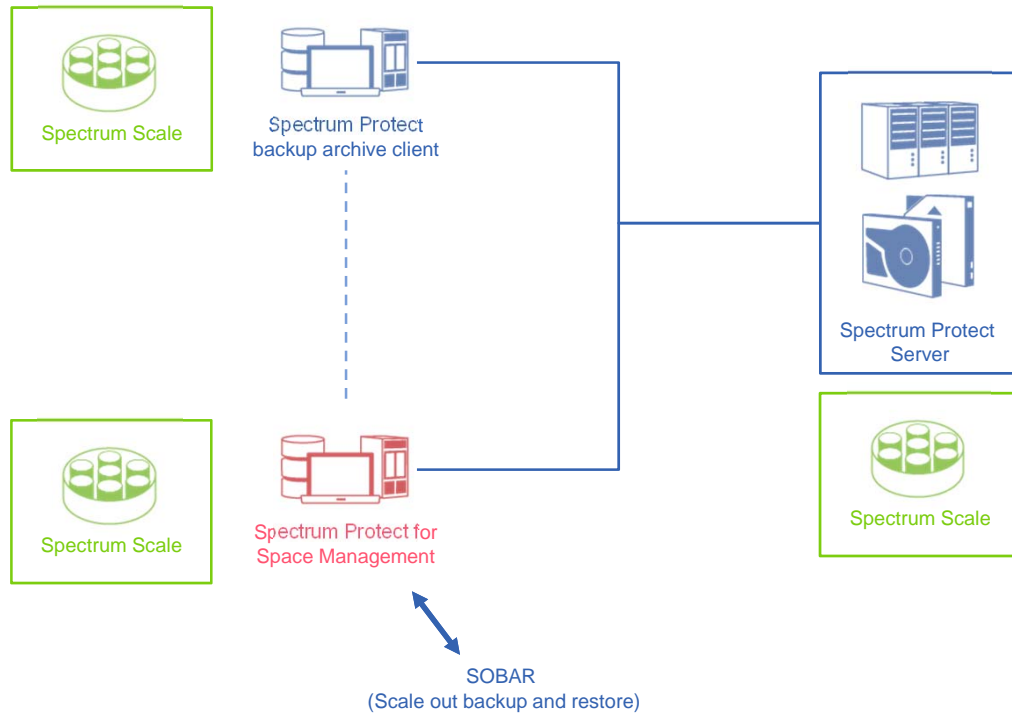


### Backup cycle:

- After start mmbackup evaluates the cluster environment and verifies product versions and settings
- Optional the Spectrum Protect server is queried for existing backup information. In other cases existing shadow DB is used for processing
- The policy engine is used to generate a list files currently eligible for backup activities
- Compare existing shadow DB and scan result to calculate file lists for required backup activities
- Expire all files deleted in the file system since last backup run
- Incremental backup all files with changed metadata in the file system since last backup run
- Selective backup all files with changed data in the file system since last backup run
- While backup activities ongoing update shadow DB inline
- Analyse backup results from all used cluster nodes and finish backup cycle by selective backup the current shadow DB

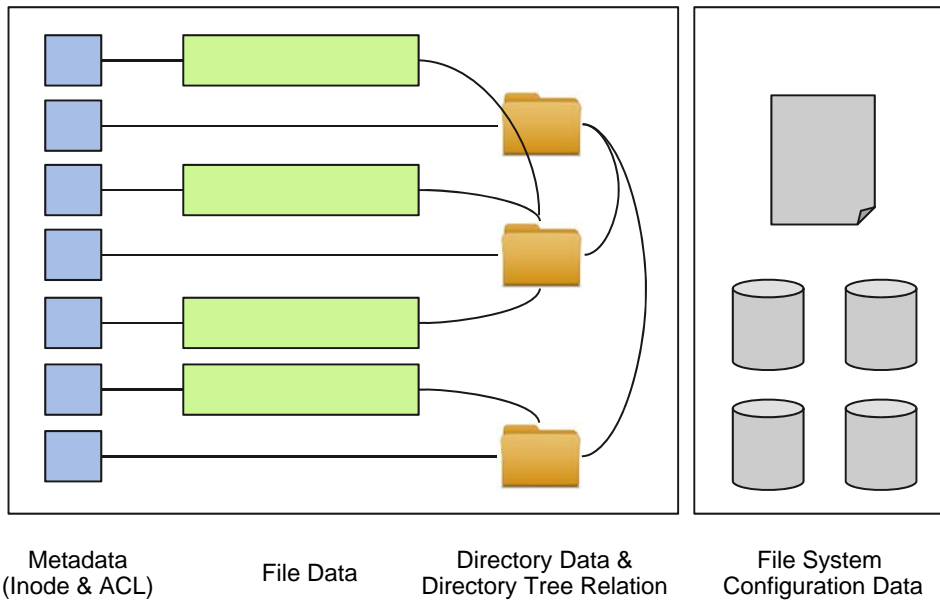
# IBM Spectrum Protect Data Protection for Spectrum Scale

## - SOBAR (Scale Out Backup and Restore)



# IBM Spectrum Protect Data Protection for Spectrum Scale

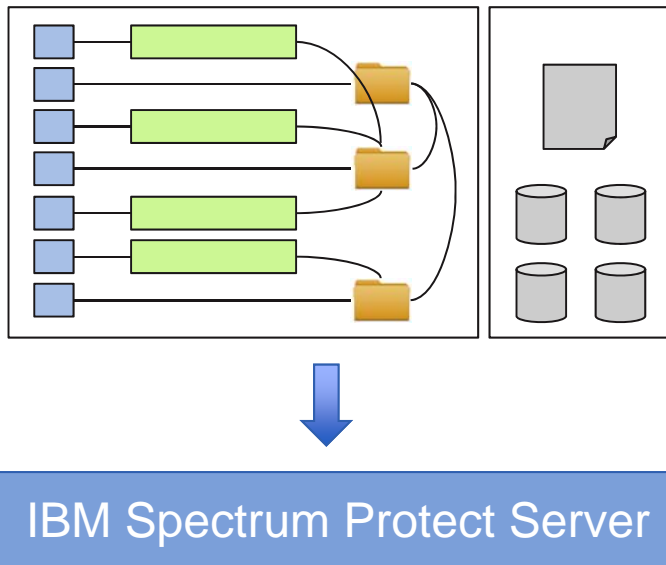
## - SOBAR (Scale Out Backup and Restore)



- SOBAR – protected items
- Metadata
  - File date
  - Directories
  - Spectrum Scale configuration

# IBM Spectrum Protect Data Protection for Spectrum Scale

## - SOBAR (Scale Out Backup and Restore)



### SOBAR – short procedure steps

- backup all file and directory data using Spectrum Scale `mmbackup` and `Spectrum Protect Backup Archive Client`
- premigrate all file data using `Spectrum Protect for Space Management` and Spectrum Scale policy engine
- preserve cluster and disk configuration `mmsdrfs`
- Backup file system volume configuration using `mmbackupconfig`



You are now ready for a disk crash!!

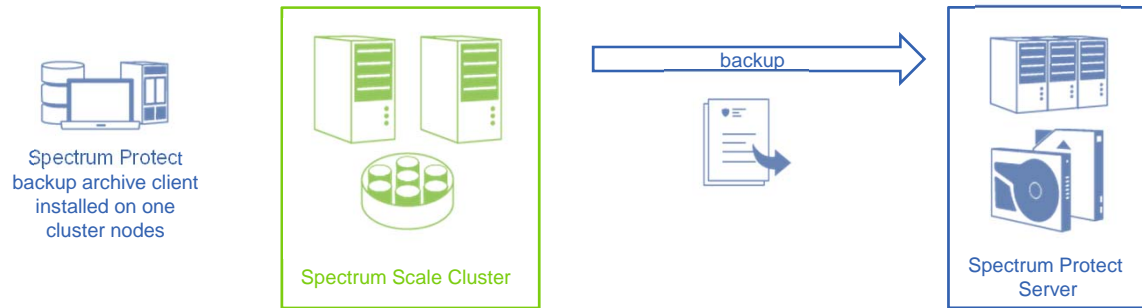
## Agenda

- Overview IBM Spectrum Protect clients
  - Backup/Archive
  - Space Management
- IBM Spectrum Protect Data Protection for Spectrum Scale
  - Tools like SOBAR, MMBACKUP, DSMIGUNDELE
  - Backup Approaches



# IBM Spectrum Protect Data Protection for Spectrum Scale

## - Backup approaches

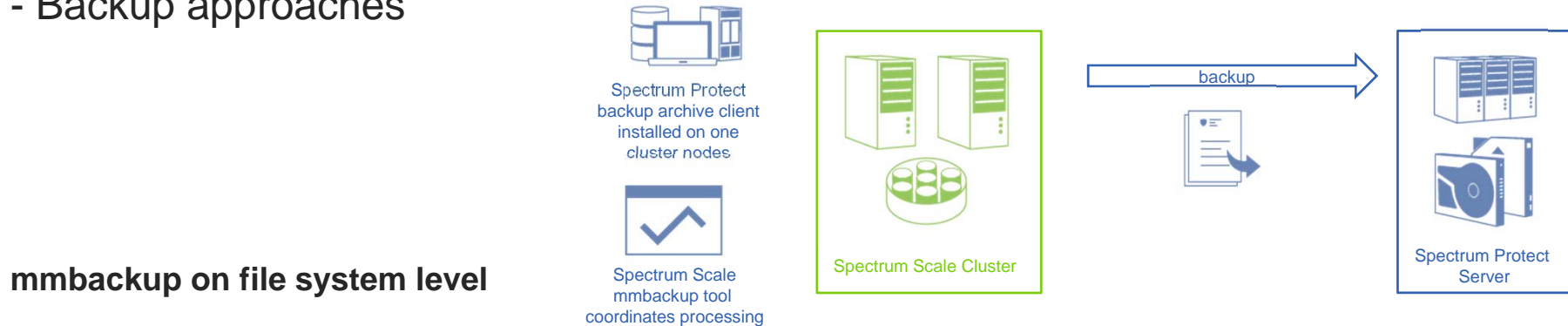


### Progressive incremental backup

- **Environment:** Small IBM Spectrum Scale installations with a small number of nodes and file systems  
IBM Spectrum Protect backup archive client installed on one or more cluster nodes
- **Scalability:** Millions of files, Terrabytes of data, up to 25.000.000 Objects (empirical value)
- **Processing:** Standard IBM Spectrum Protect backup archive client progressive incremental is used to perform file system backup
- **Pros:** Simple setup and usage
- **Cons:** Limited performance and scalability

# IBM Spectrum Protect Data Protection for Spectrum Scale

## - Backup approaches

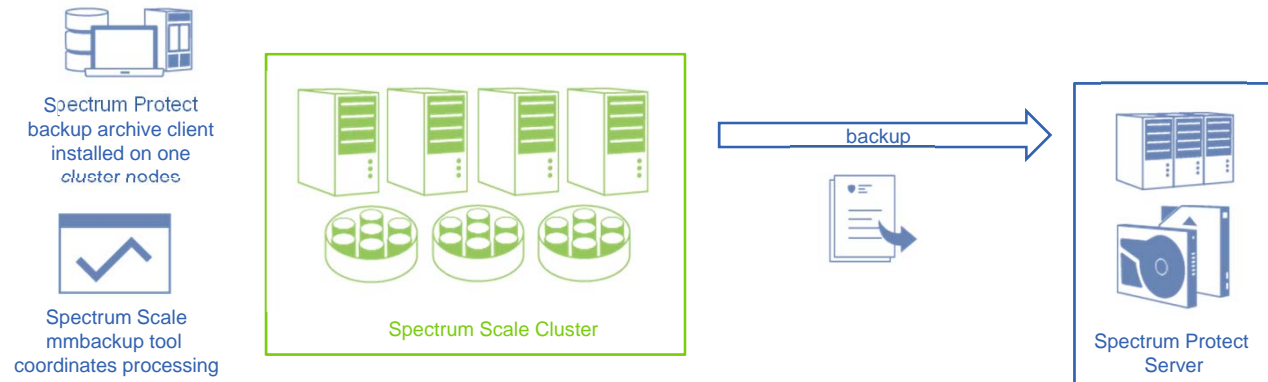


### mmbackup on file system level

- **Environment:** Medium IBM Spectrum Scale installations with a single digit number of nodes and file systems. IBM Spectrum Protect backup archive client installed on several cluster nodes
- **Scalability:** Tens of millions of files, Tens of terrabytes of data, up to 1.000.000.000 Objects (empirical value)
- **Processing:** IBM Spectrum Scale mmbackup scans file system and IBM Spectrum Protect data base and generates list of backup candidates. IBM Spectrum Protect backup archive client used by the mmbackup to perform file system backup.
- **Pros:** Simple setup and usage, good performance and scalability

# IBM Spectrum Protect Data Protection for Spectrum Scale

## - Backup approaches

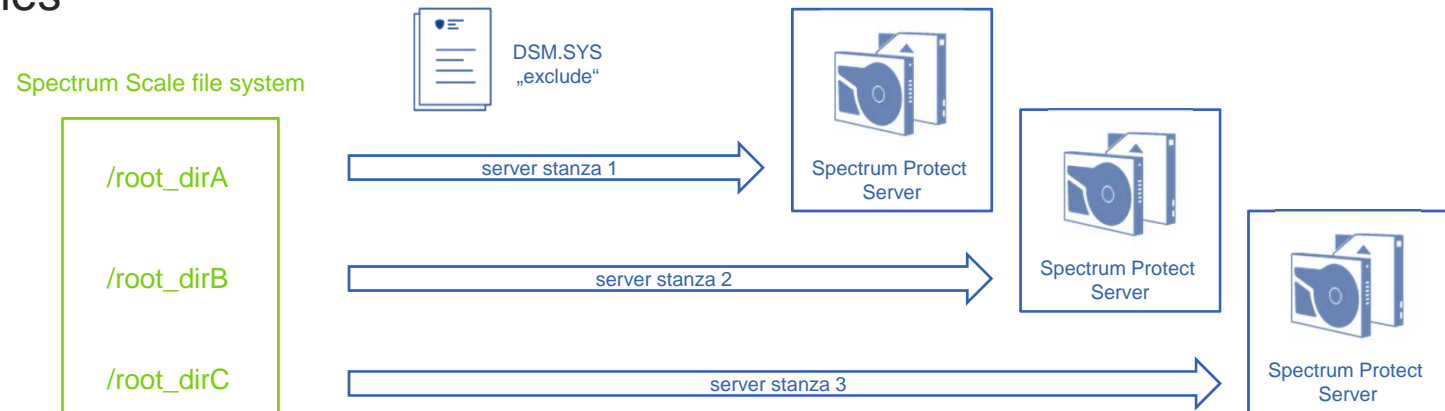


### mmbackup on root directories level

- **Environment:** Large IBM Spectrum Scale installations with a double digit number of nodes and file systems. IBM Spectrum Protect backup archive client installed on several cluster nodes
- **Scalability:** Hundreds of millions of files, Hundreds of terrabytes of data

# IBM Spectrum Protect Data Protection for Spectrum Scale

## - Backup approaches

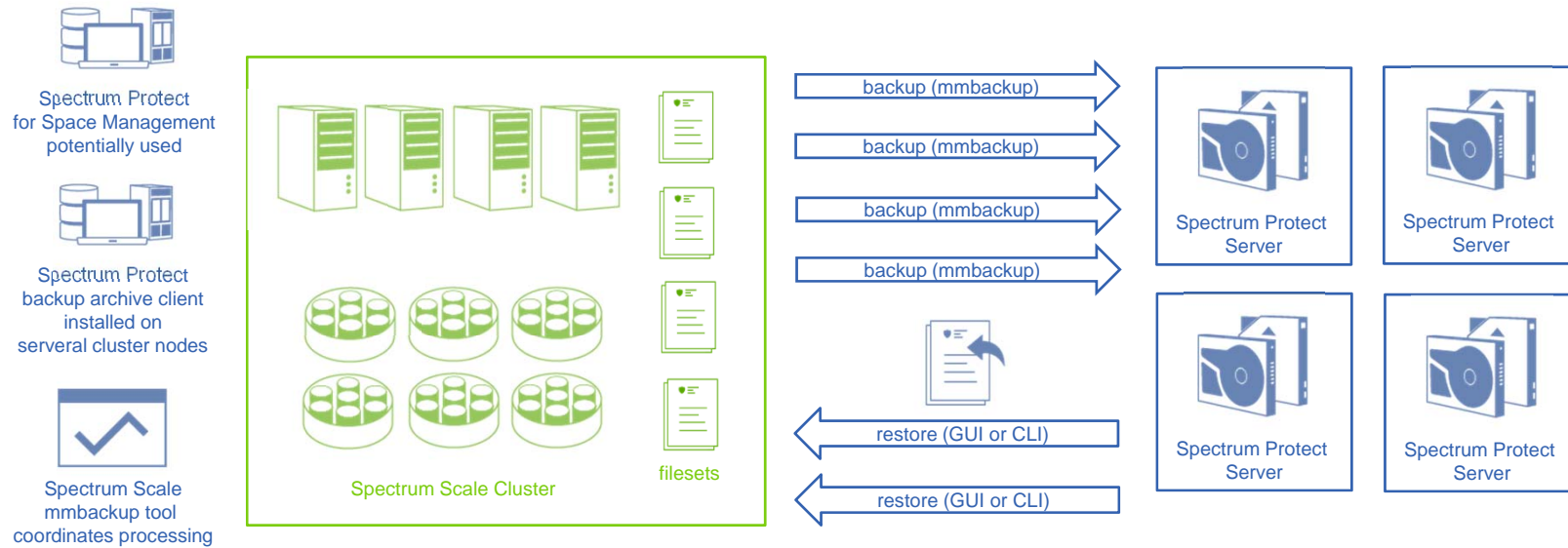


### mmbackup on root directories level

- **Processing:** IBM Spectrum Protect processing used to divide file system into backup parts on root directory level. One part goes to one server. IBM Spectrum Scale mmbackup is used to backup parts to different servers using `-tsm-servers` option.
- **Pros:** usable on existing data w/o full backup, scalable, IBM Spectrum Protect server housekeeping can be parallelized
- **Cons:** complex planning and setup, IBM Spectrum Scale mmbackup sequential processing

# IBM Spectrum Protect Data Protection for Spectrum Scale

## - Backup approaches

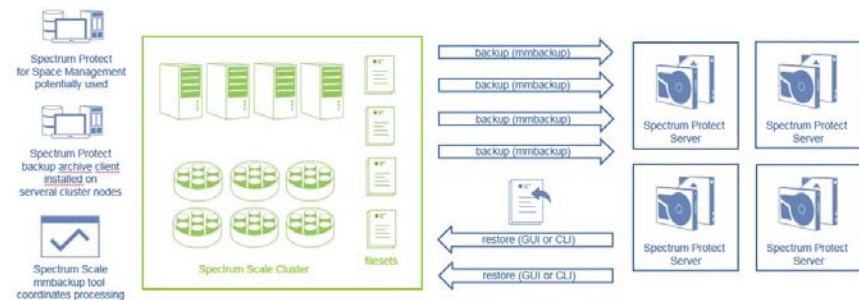


### mmbackup on fileset level including active server binding

- **Environment:** Extremely large IBM Spectrum Scale environments or fast growing environments
- **Scalability:** Billions of files, petabyte of data

# IBM Spectrum Protect Data Protection for Spectrum Scale

## - Backup approaches

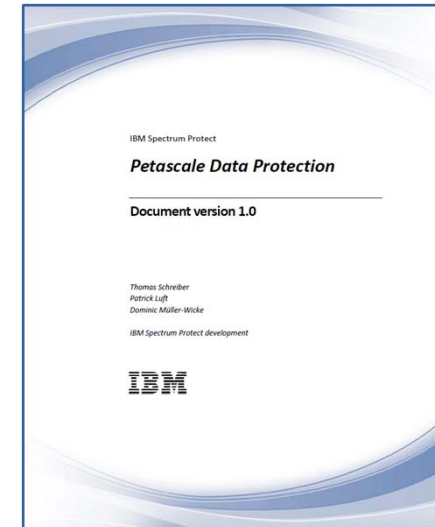


### mmbackup on fileset level including active server binding

- **Processing:** IBM Spectrum Scale file system divided into several independent filesets. IBM Spectrum Scale mmbackup running on fileset level. Each fileset belongs to a single IBM Spectrum Protect server. Potentially usage of IBM Spectrum Protect for Space Management to enable active server binding and SOBAR features.
- **Scalability:** extremely scalable, parallel IBM Spectrum Scale mmbackup processing for different filesets
- **Cons:** Fileset usage must be planned wisely, processing requires detailed planning to prevent resource over allocation.

## IBM Spectrum Protect Data Protection for Spectrum Scale - Backup approaches

### mmbackup on fileset level including active server binding



The paper describes a data protection approach scaling up to hundreds of petabytes for an IBM Spectrum Scale file systems using IBM Spectrum Protect backup-archive client and IBM Spectrum Protect for Space Management. The focus of this paper is to provide configuration guidance for the setup and operation of the data protection processes in such an environment.

<https://www.ibm.com/developerworks/community/wikis/home?lang=en#!/wiki/Tivoli%20Storage%20Manager/page/Petascale%20Data%20Protection>

Questions?

Thank you!

