IBM **Spectrum Protect**

Spectrum Protect / Spectrum Scale Integration

Chicago Spectrum Scale Users Group

Ron Henkhaus| IBM System Storage SoftWare Advanced Technology (SWAT)| rhenkha@us.ibm.com

IBM®

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.

IBM Spectrum Storage™ Family
Securely 'unboxing' storage to revolutionize data economics

Family of Storage Management and Optimization Software

Control Protect Archive
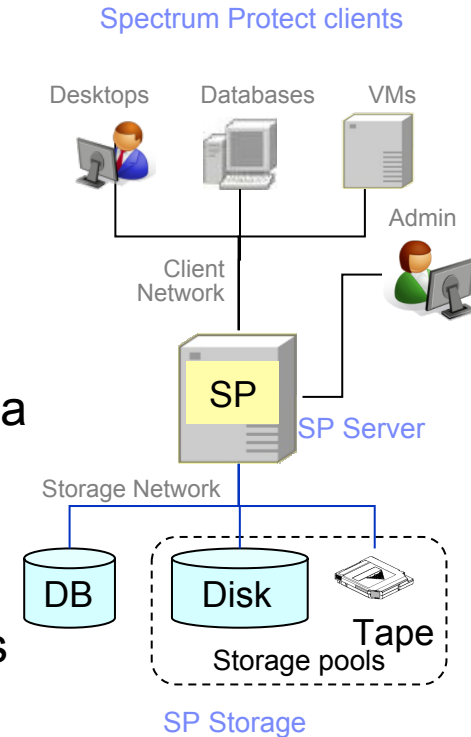
Virtualize Accelerate Scale

Any Storage   **Flash Systems**   **Private, Public or Hybrid Cloud**

IBM Spectrum Control™	Analytics-driven data management to reduce costs by up to 50 percent
IBM Spectrum Protect™	Optimized data protection to reduce backup costs by up to 38 percent
IBM Spectrum Archive™	Fast data retention that reduces TCO for active archive data by up to 90%
IBM Spectrum Virtualize™	Virtualization of mixed environments stores up to 5x more data
IBM Spectrum Accelerate™	Enterprise storage for cloud deployed in minutes instead of months
IBM Spectrum Scale™	High-performance, highly scalable storage for files, objects & analytics

3

Spectrum Protect Architecture

- **Spectrum Protect (SP) client – server architecture**
  - Backup Client selects and backs up or archives the data
  - Backup Server catalogs data and stores it in storage pools

- **Spectrum Protect server has internal database as catalog**

- **Storage pools can be on a variety of storage media**
  - Flash, disk, NAS, optical and tape
  - Storage pool tiering allows automated migration
  - Includes transparent migration between the pools

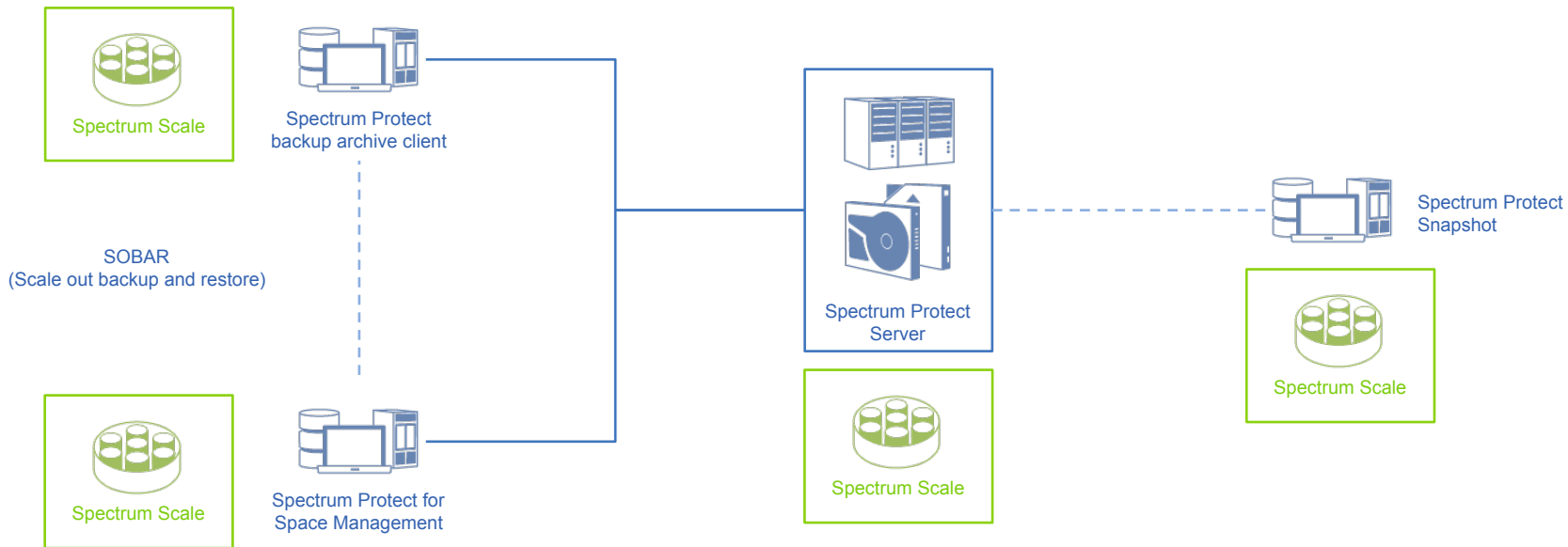- **Spectrum Protect server provides embedded tools for central management, monitoring and reporting**
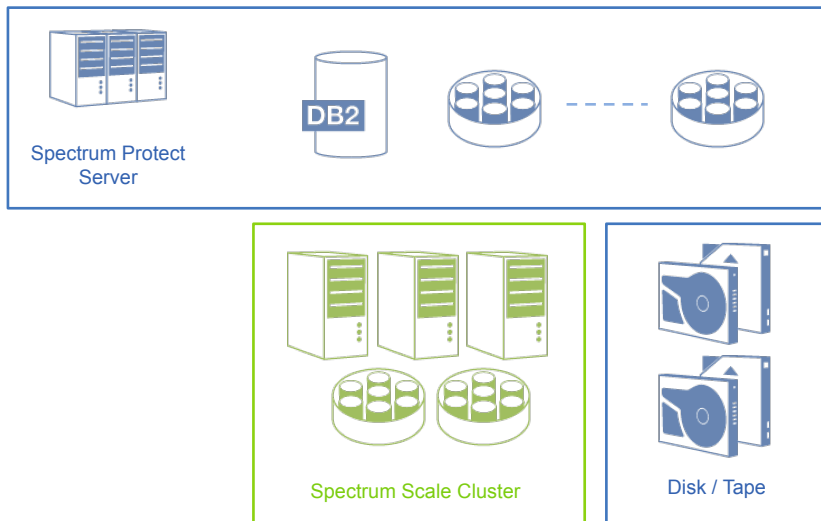
Spectrum Protect clients

Desktops   Databases   VMs

Admin

Client Network

SP

SP Server

Storage Network

DB   Disk   Tape

Storage pools

SP Storage

4

Spectrum Protect – Spectrum Scale Integration

High Level Functional Overview

Spectrum Protect / Spectrum Scale Integration Overview

Spectrum Scale

Spectrum Protect
backup archive client

SOBAR
(Scale out backup and restore)

Spectrum Scale

Spectrum Protect for
Space Management

Spectrum Protect
Server

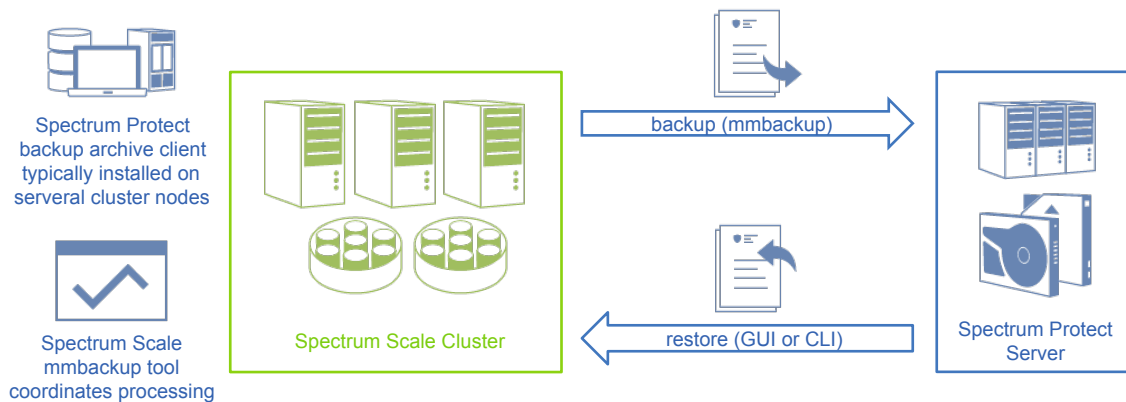Spectrum Scale

Spectrum Protect
Snapshot

Spectrum Scale

Spectrum Protect Server

- Spectrum Scale brings proven file system technology to IBM Spectrum Protect server storage (pools, DB, & logs) for a high level of performance and simplicity
- Simplified storage configuration
- Entire storage pools under a single file system
- Spectrum Scale storage pool concept to split Protect data and metadata (DB2)
- Global namespace sharable by multiple IBM Spectrum Protect servers
- Ideally suited for disk-based data protection solutions using IBM Spectrum Protect


Spectrum Protect Server

DB2

Spectrum Scale Cluster

Disk / Tape

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
- Spectrum Scale mmbackup uses local shadow of Spectrum Protect DB and uses policy engine to identify files for backup
- Spectrum Protect backup/archive client is used „under the hood" to backup files to Spectrum Protect Server
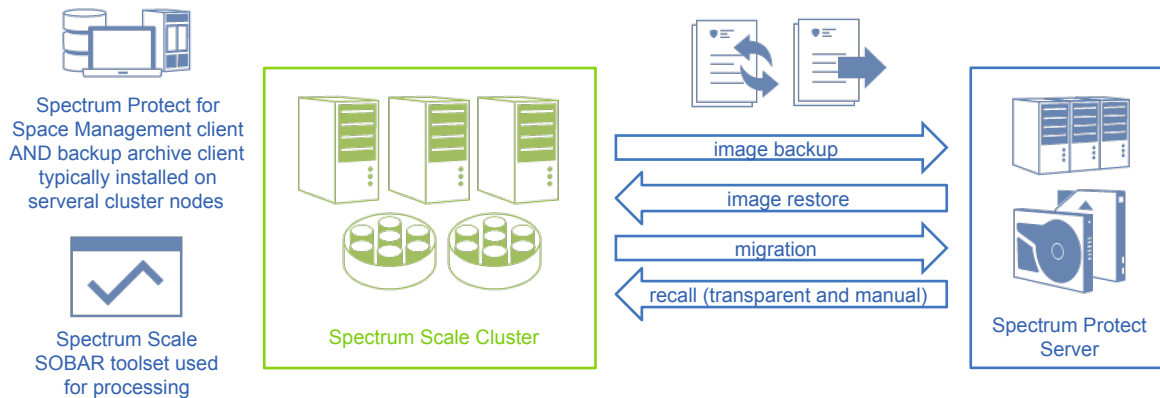- Spectrum Protect restore (CLI or GUI) can be used to restore files

Hierarchical Storage Management (aka Space Management)
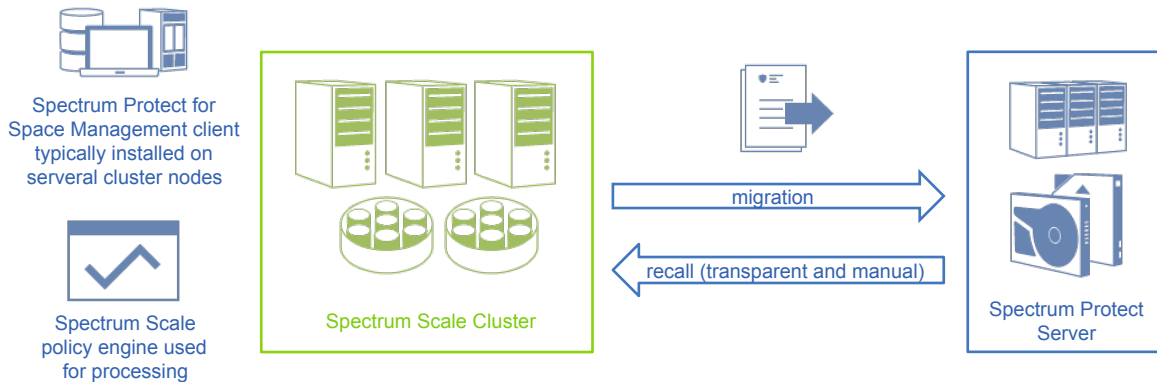
- Massive parallel migration processing to efficiently reduce filessystem to defined threshold
- Cluster wide distributed parallel recall of files in file specific mode
- Tape optimized recall for mass recalls in batch mode
- Close integration with Spectrum Protect Backup-Archive Client
- Recovery of migrated files in stub format

Fast Disaster 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

migration

recall (transparent and manual)

Spectrum Protect
Server

Function Backup

• Spectrum Protect HSM used to premigrate files
• SOBAR toolset used to generate filesystem configuration file and metadata image
• Spectrum Protect Backup-Archive client used to backup configuration and image files

Function Restore

• Spectrum Protect Backup-Archive client used to restore config and image files
• SOBAR toolset used to recreate file system structure
• Spectrum Protect HSM used to pre-fetch files and allow direct access by applying transparent recall
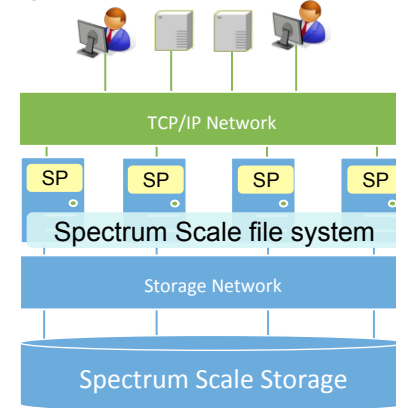
Spectrum Protect on Spectrum Scale - Overview

- **Multiple** Spectrum Protect (TSM) instances store DB and storage pools in a Spectrum Scale file system (GPFS)
  - Spectrum Scale provides global name space for all Spectrum Protect instances
  - Instances share all file system resources

- Spectrum Protect instances run on cluster nodes
  - Accessing the file system directly
  - Server instances can run on NSD servers or clients

- Spectrum Scale file systems balances the workload and capacity for all Spectrum Protect instances on disk

- Standardized, scalable and easy to use storage infrastructure for the multiple instances
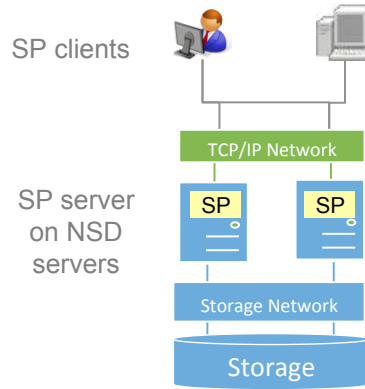
Spectrum Protect Clients

TCP/IP Network

SP SP SP SP

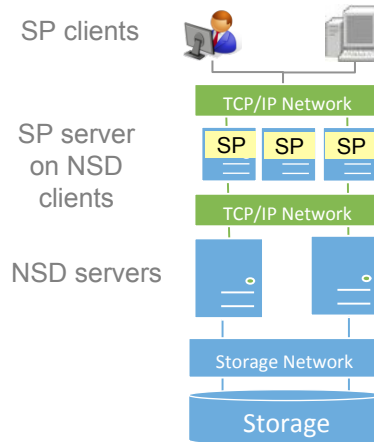Spectrum Scale file system

Storage Network

Spectrum Scale Storage

Spectrum Scale storage for Spectrum Protect

Spectrum Protect on Spectrum Scale – Deployment Options

On NSD server

SP clients

TCP/IP Network

SP server on NSD servers

SP   SP

Storage Network

Storage

On NSD client

SP clients

TCP/IP Network

SP server on NSD clients

SP   SP   SP

TCP/IP Network

NSD servers

Storage Network

Storage

With Elastic Storage Server

SP clients

TCP/IP Network

SP server on NSD client
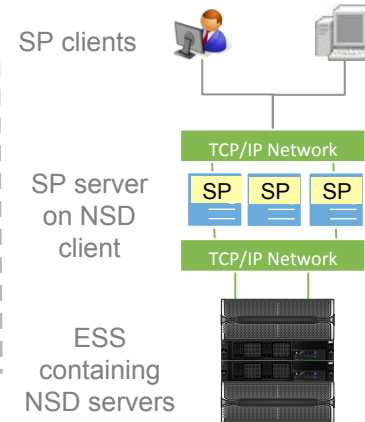
SP   SP   SP

TCP/IP Network

ESS containing NSD servers

- SP server on NSD servers

- Direct SAN storage access

- Requires less infrastructure

- Good if Scale is used just by SP

- SP runs on client connected to Scale / ESS NSD server

- SAN or LAN access from SP server to Scale/ ESS server

- Requires more infrastructure

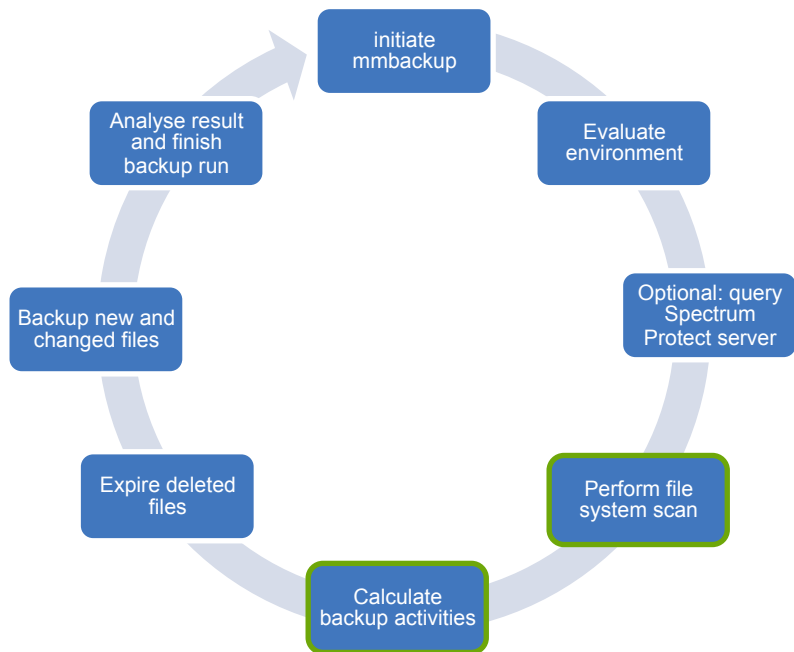- Good if Scale is used by multiple applications

13

Spectrum Scale File System
Scalable Backup Functional Overview

Backup Of Large Spectrum Scale File Systems

- initiate mmbackup
- Evaluate environment
- Optional: query Spectrum Protect server
- Perform file system scan
- Calculate backup activities
- Expire deleted files
- Backup new and changed files
- Analyse result and finish backup run

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 of 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 of all files with changed metadata in the file system since last backup run
- Selective backup of all files with changed data in the file system since last backup run
- While backup activities are ongoing, update shadow DB inline
- Analyse backup results from all used cluster nodes and finish backup cycle by selective backup the current shadow DB

Scalable Backup – Backup Methods Compared

Backup Related File & Environment Changes	IBM Spectrum Protect (progressive incremental)	IBM Spectrum Scale (mmbackup)
Detects changes in files and sends a new copy of the file to the server	Yes	Yes
Detects changes in metadata and updates the file metadata at the server or sends a new copy of the file to the server in terms of ACL/EA changes	Yes	Yes
Detects directory move, copy or rename and send a new copy of the file to the server	Yes	Yes
Detects local file deletion and expires the file at the server	Yes	Yes
Detects Spectrum Protect file space deletion or node/policy change and sends a new copy of the file to the server	Yes	No*
Detects file deletion from Spectrum Protect server and sends a new copy of the file to the server	Yes	No*
Detects additions of new exclude rules and expires the file at the server	Yes	Yes
Detects policy change due to new include rule and rebinds the file to the new storage pool	Yes	No**
Recognizes copy mode and copy frequency options	Yes	No*
Detects migration state changes (Spectrum Protect for Space Management) and updates server object	Yes	Yes
Detects that a file wasn't processed successfully during backup operation and attempts again at the next backup	Yes	Yes

* mmbackup queries the Spectrum Protect Server only once at the time of the first backup. Changes which are performed at the Spectrum Protect Server directly by using the Spectrum Protect administrative client can't be detected by mmbackup. It is recommended to rebuild the mmbackup shadow data base in the case of Spectrum Protect Server file space changes.
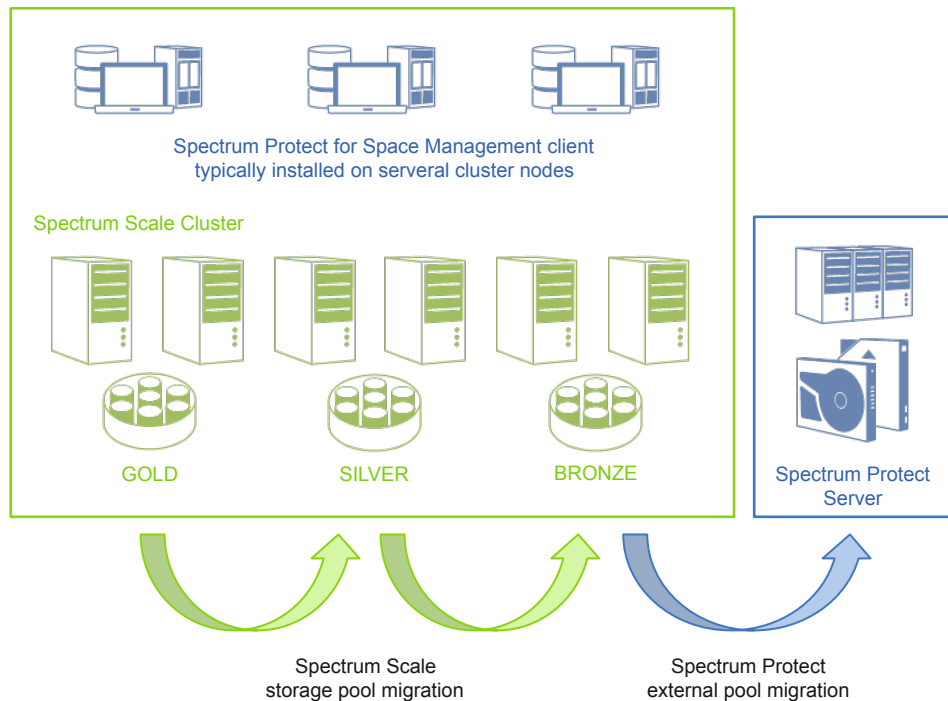** Spectrum Protect include rules with associated management class bindings can't be detected from mmbackup. Therefore mmbackup doesn't rebind a file in the case of management class changes in include rules.

—

Spectrum Protect for Space Management for Scale File Systems
Functional Overview

Data Lifecycle Management Concept

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

Spectrum Scale Cluster

GOLD   SILVER   BRONZE

Spectrum Protect
Server

Spectrum Scale
storage pool migration

Spectrum Protect
external pool migration

- Spectrum Protect for Space Management installed
  on several cluster nodes to increase scalability and
  performance by applying parallelism.

- Connection to Protect server via LAN, Data via LAN
  or SAN (LAN-free using storage agent)

- Spectrum Scale storage pool concept and policy
  engine used for data migration to lower tiers.

Typical data lifecycle:

1. Data created in file system and frequently used is
   automatically stored in high performance storage pool
   GOLD (e.g. flash disk)
2. Less frequently used data automatically moved to
   cheaper and slower disk storage SILVER and BRONZE
   (e.g. SATA)
3. Unused data automatically migrated to external storage
   pool (Spectrum Protect server pools) from Spectrum
   Protect for Space Management to free space in file
   system. A stub file resides in file system.
4. Data deleted in the file system will be removed from
   Spectrum Protect server with reconcile process.

Access to stub files will automatically recall data into the file
system.

Links

- **Integrating IBM Tivoli Storage Manager with IBM Elastic Storage:**
  https://www.ibm.com/developerworks/community/wikis/home?lang=en#!/wiki/Tivoli%20Storage%20Manager/page/Integrating%20IBM%20Tivoli%20Storage%20Manager%20with%20IBM%20Elastic%20Storage

- **Spectrum Protect HSM – Scale Integration:**
  http://www-01.ibm.com/support/docview.wss?uid=swg27018848

- **IBM Spectrum Storage home:**
  http://www-03.ibm.com/systems/uk/storage/spectrum/

- **IBM Spectrum Protect Home Page:**
  http://www-01.ibm.com/software/uk/tivoli/csi/backup-recovery/

- **Spectrum Protect Knowledge Center:**
  http://www.ibm.com/support/knowledgecenter/en/SSGSG7_7.1.5/tsm/welcome.html

- **Spectrum Protect Wiki:**
  https://www.ibm.com/developerworks/community/wikis/home?lang=en#!/wiki/Tivoli%20Storage%20Manager/page/Home

- **IBM Spectrum Scale Home Page:**
  http://www-03.ibm.com/systems/storage/spectrum/scale/overview.html

- **Spectrum Scale Knowledge Center:**
  http://www-01.ibm.com/support/knowledgecenter/STXKQY_4.2.0/ibmspectrumscale42_welcome.html?lang=en

- **Spectrum Scale Wiki:**
  https://www.ibm.com/developerworks/community/wikis/home?lang=en#!/wiki/General%20Parallel%20File%20System%20%28GPFS%29

- **Spectrum Scale Redbook:**
  http://www.redbooks.ibm.com/abstracts/sg248254.html?Open

- **ESS home page:**
  http://www-03.ibm.com/systems/uk/storage/spectrum/ess/

- **ESS Knowledge Center and Scale Native RAID:**
  http://www-01.ibm.com/support/knowledgecenter/SSYSP8_3.5.0/sts35_welcome.html?cp=SSYSP8_3.5.0%2F0&lang=en