IBM **Spectrum Protect**

Spectrum Protect / Spectrum Scale Integration

News and recommendations 1Q16

Christian Müller | Spectrum Protect Development | chmuelle@de.ibm.com
Dominic Müller-Wicke | Spectrum Protect Development | dominic.mueller@de.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.

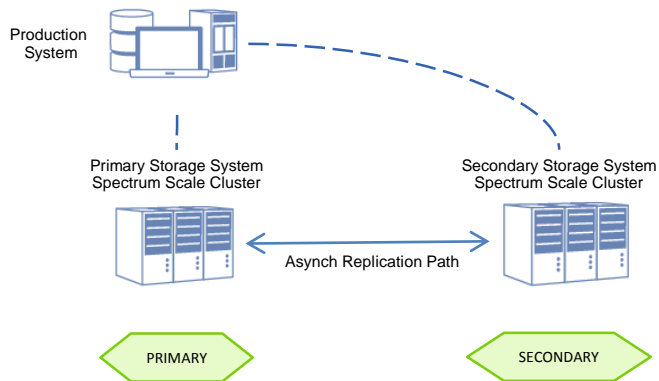Information about potential future products may not be incorporated into any contract.

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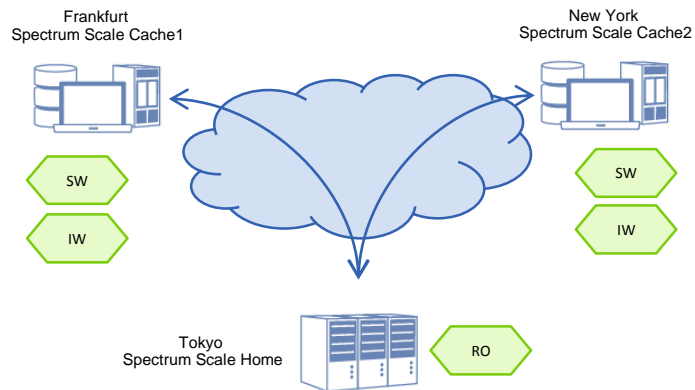
Agenda

- What's New?
- Best Practices

What's New?

Spectrum Protect Support For Spectrum Scale AFM (Spectrum Protect 7.1.3)

- Support for DR, System Migration and Branch Office scenarios described and defined in whitepaper (link in References page)
- General Recommendations if using IBM Spectrum Protect for Space Management and AFM
  - Prevent cache eviction
  - Prevent snapshots (DR: peer snapshots) if possible
  - Prevent mass recalls because of AFM replication processing

Backup On Fileset Granularity (Spectrum Protect 7.1.3)

- Use mmbackup to process filesets independent from others (Spectrum Scale 4.1.1)
- Use snapshots to backup filesets
- Spectrum Protect 7.1.3 improvement: -snapshotroot option was enhanced to allow more complex snapshotroot paths.

Example:

Before 7.1.3: snapshot root must be right after the FS name in a file path. e.g.:

```
Path:          /fsname/dir1/file1
Path in snap: /fsname/snapshotroot/snap1/dir1/file1
Snaproot:      /fsname/snapshotroot/snap1
```

Starting with 7.1.3 snapshotroot can be anywhere in the path. e.g.:

```
Path:          /fsname/dirN/fileset1/dir1/file1
Path in snap: /fsname/dirN/fileset1/snapshotroot/snap1/dir1/file1
Snaproot:      /fsname/dirN/fileset1/snapshotroot/snap1
```

Prevention Of Random Recalls For Backup

- 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
  - The file has ACL/EA data and was renamed or moved
  - (File data changes)
- Worst case: rename of file system root directory ☹
- Backup leads to random access of files in Spectrum Protect server. Potentially one tape mount per file to be recalled

- Improvement (stage 1 – Spectrum Scale 4.2.0):
  - mmbackup skips migrated files from backup processing and marks file in shadowDB
  - A file list is created that includes all migrated files eligible for backup
  - User can use tape optimized recall function to bring eligible files back to disk
  - User can start mmbackup again

- Future: mmbackup uses tape optimized recall internally

Further Improvements (Spectrum Protect 7.1.4)

- 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

Hints and Tips

Spectrum Protect Server

- Use IBM Spectrum Protect blueprint v2.2 for reference:
  *ibm.biz/IBM**SpectrumProtectBlueprints***

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

- New container pools configuration included

- Important performance tuning options: DIRECTIO & DIOENABLED
  - For the time of one transaction write data to memory
  - Actively flush memory before commiting data

Backup

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

- Consider work load on Spectrum Scale cluster nodes

  - mmbackup puts load on each node involved
  - Backup nodes need access to storage with enough bandwith
  - Backup nodes need access to Spectrum Protect server
  - Serialize backups of different file systems
  - Serialize backups with other activities on the Spectrum Protect server (other node backups, Server maintenance,...)

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 environments including Windows or machine generated file names)
  - Files with control-X, control-Y, carriage return and the new line character in their name can't be backed up to Spectrum Protect.
  - 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. These changes lead to a new backup of all affected files

Backup

Align the option setting between all components

Define number of bytes to be used in a single backup transaction

`--max-backup-size` ——————————————————— `TXNBYTELIMIT`

Define number of files to be used in a single backup transaction

`--max-backup-count`
`--max-expire-count` ————————————————————— `TXNGROUPMAX`

Define number of parallel backup processes

`--expire-threads` —————— `RESOURCEUTILIZATION` —————— `MAXSESSIONS`
`--backup-threads`   `MAXNUMMP`

Spectrum Scale
mmbackup command
line options

Spectrum Protect backup archive
client processing options

Spectrum Protect
Server processing options

Backup

Recommendations on mmbackup options

- `-S:` Use snapshots to reduce transaction failures due to re/moved files
- `-m:` DO NOT USE ! Use fine grain options below instead
- `--expire-threads:`
  - If you don't expect a peak in your deletions use default (4). Increase for "peak" days.
- `--backup-threads:`
  - Each thread will start one dsmc process
  - Align setting with free sessions available. Include potential restores in caclulation (MAXSESSIONS)
  - Align setting with free mount points if backup goes to tape (MAXNUMMP)

Backup

Recommendations on mmbackup options

- RESOURCEUTILIZATION option defines number of consumer and producer threads in dsmc command for backup
- Table shows values and session numbers for backup
- Expiration processing uses only one sessions

Number of backup sessions can be calculated as follows:
#sessions = (#backup-threads * #nodes * (RESOURCEUTILIZATION[VALUE] - 1))

Number of backup sessions can be calculated as follows:
#sessions = (#expire-threads * #nodes)

Number of backup threads for a given number of mount points can be calculated as follows:
#backup-threads = #mount-points / (#nodes * (RESOURCEUTILIZATION[VALUE] - 1))

Value	#Sessions (send+query)
1	1
0 (default), 2	2 (1+1)
3, 4	3 (2+1)
5, 6	4 (3+1)
7	5 (4+1)
8	6 (5+1)
9	7 (6+1)
10	8 (7+1)

Backup

Recommendations on mmbackup options

- `-B:` DO NOT USE ! Use fine grain options below instead
- `--max-backup-size & max-backup-count:`
    - Use multiple of TXNBYTELIMIT and TXNGROUPMAX.
    - Transactions are cheaper than sessions!

Example:
Assumption: Only small files in backup; TXNBYTELIMIT, --max-backup-size set to high values
Setting –max-backup-count to N * TXNGROUPMAX will lead to multiple sequential transactions within one dsmc CLI call per node.
Result is that overall number of sessions and mount points needed for backup can be reduced with factor N per node.

- `--max-expire-count:`
    - Expiration is metadata operation. Use parallelism to reduce overall expiration time
    - Increase count for „peak" days

Backup

- Spectrum Protect offers a rich set of include and exclude options to control which files and directories are backed up.
- mmbackup is building these options into its policy for backup

- 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

- 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.

- Use processing option HSMGROUPEDMIGRATE YES to improve migration performance to tape.

- Do not set processing option MIGFILEEXPIRATION to 0. Recovery of expired files will be impossible.

- Space Management is not backup. Use option MIGREQUIRESBACKUP YES.

- Use tape optimized recall if it can be combined with business processes.

Questions?

References

IBM Knowledge Center
http://www-01.ibm.com/support/knowledgecenter/

IBM Spectrum Protect resources landing page
http://www-01.ibm.com/support/docview.wss?uid=swg21684850

Overview on Spectrum Protect – Spectrum Scale Integration
https://www.ibm.com/developerworks/community/wikis/home?lang=en#!/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?lang=en#!/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-01.ibm.com/support/docview.wss?uid=swg27018848
Setup cross platform cluster: http://www-01.ibm.com/support/docview.wss?uid=swg27028178