BrightStor® ARCserve® Backup r11.5 Disk Staging Backup

Storage Management Team
July 2006

Table of Contents

Introduction

The purpose of this document is to introduce disk staging backup, also known as disk-to-disk to tape (D2D2T) backup, a new feature in BrightStor® ARCserve® Backup r11.5. It will discuss all the related BrightStor ARCserve Backup operations including backup, data migration policies, restore, job monitoring, and reports.

Why Disk Staging?
Traditionally, backups and restores are performed with tapes. While tape is still a popular choice for data archiving, it has several drawbacks for daily backup use. As storage capacity grows exponentially every year, backup window becomes smaller than ever. Backup simply takes longer and becomes harder to manage to fit in the backup window. By the same token, recovery time takes longer as well, which is crucial especially in disaster recovery scenarios.

Well, no problem! I can just keep upgrading my tape drives to the latest and fastest to stay within the backup window. True, tape technology may follow the Moore's Law of doubling performance and capacity every 18 months. However, there are other aspects in a backup environment besides tape drive speed that affects the overall throughput. In fact, the most likely bottleneck in any backup environment is not the tape drive, but the backup server and the network, which are not able to stream data fast enough to the tape drive. As a result, "shoeshining" occurs which causes the tape head wearing out quickly. Unless your data stream can keep up with the fast tape, the newer tape technology could be very costly and resource intensive without providing comparable performance. Media management can also be a problem with traditional tape backups. The number of tapes grows quickly and it becomes a nightmare for the backup operators to manage them.

Disk staging addresses all of these concerns. With disk staging backup, data is directly backed up to disk, then migrated to tape at a more convenient time. Disk in general is faster and less error-prone than most tape technology, thus shortening the backup and restore time. After a period of time based on some set policy, data is then copied from disk to tape. Because there is no network involved, data transfer from disk to tape is always local to the backup server. This ensures the fastest possible data stream from the server to the tape drives, thus greatly reducing "shoeshining" and operational costs. Lastly, because fewer tapes are used, the backup operator has greater control over the data protection environment and easier management of the media.

Wait a minute; I can do the same thing with previous versions of BrightStor ARCserve products. Yes and no. Customers can do D2D2T backup by manually setting up different tape copy scripts. However, the process could be cumbersome and error-prone because it is not automated. With r11.5, D2D2T backup is driven from the graphical user interface and tightly coupled to the base product. In addition, data migration process from disk to tape is automated through policies.

Disk staging is available in BrightStor ARCserve Backup servers running on Windows, UNIX and Linux platforms.

Backup

The underlying technology that disk staging uses is file system device (fsd), which has been available in BrightStor ARCserve Backup for quite some time. File system devices makes a disk directory look like a tape device to the BrightStor ARCserve Backup server. To prepare for disk staging backup, you have to first create one or more fsd(s). Then you designate the fsd(s) to be the staging device. That's all!

In Device Manager, click on the "Configure File System Device Groups" button. Choose the device group that you want for the staging. Then check "Enable Staging" in Figure 1 below.
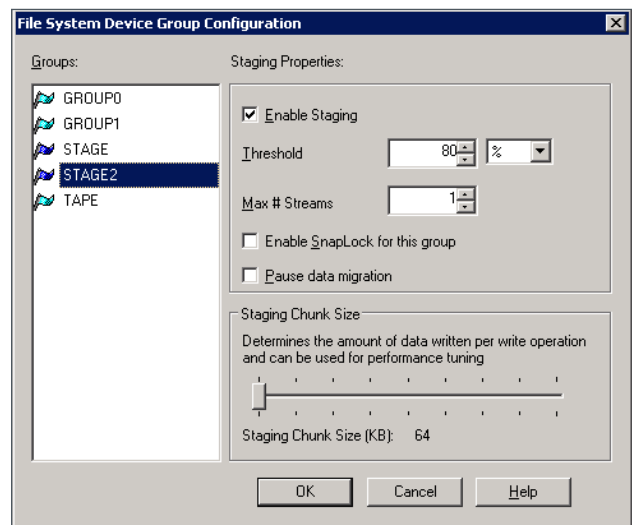

Figure 1. Enabling Staging for a Device Group.

When the disk capacity reaches the threshold, the backup job will fail. To ensure that backup continues, by default, a makeup job will be created automatically to back up to tape. This behavior will occur when the threshold is reached at both the start and the middle of a backup, maximizing the success rate of the backup.

Once a staging device group is defined, enabling disk staging for a job is as easy as a checkbox. In Backup Manager, there is now an additional tab called "Staging". Simply set up your backup job like you always do before with source, destination, and schedule. Then check the "Enable Staging" checkbox as shown in Figure 2. The job will now be submitted as a staging job.
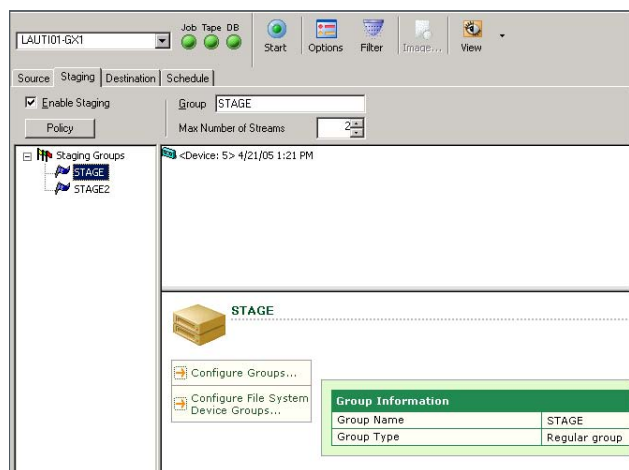

Figure 2. Enabling Staging for a Backup Job.

The number of streams going to a disk staging device is determined by the same logic as multiplexing job. For local backups, streams are created at the volume level. For example, if there are two volumes, c: and d:, two streams (child jobs) will be created for each volume. For remote backups, streams are created at the agent level. For example, if there are a client agent and a SQL agent, two streams (child jobs) will be created for each agent. The base product comes with two streams. Customers need to install and license the Disk Staging Option to support up to 32 streams.

Note that staging device cannot be formatted or erased. The media name for the staging device is whatever the name that the device is already formatted with. If the staging device started as blank, the disk staging backup job will automatically format it with the time-stamp in the regular BrightStor ARCserve Backup fashion. Conventional media name for GFS and simple rotation jobs is irrelevant for the staging device. The regular naming convention comes into play when data is copied from disk to tape.

The following functionalities are not supported by disk staging backup:

• NAS sources

• SAP Agent (Windows)

• Mutli-striping SQL Backup

• Hardware VSS

• Serverless Option

Job Queue/Activity Log

Similar to multiplexing/multi-streaming jobs, the job in the job queue is shown in a parent-child relationship fashion. As shown in Figure 3, the parent job spawns two child jobs, backup and migration. The backup job is for backing up from disk-to-disk. The migration job is for copying the data from disk to tape.
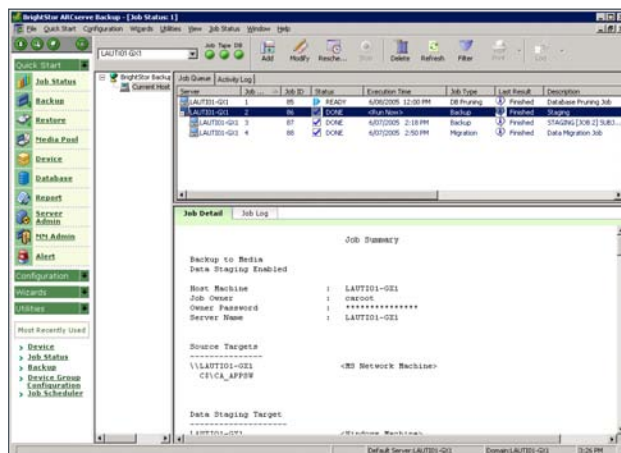

Figure 3. A Staging Job with Parent-Child Relationship.

When you first submit the staging job, you will only see the backup job that gets kicked off by the parent job. You will not immediately see the migration job. In fact, the status of the master job will be "DONE" for run-once jobs or "READY" for rotation jobs. The migration is kicked off based on the copy policy that you set when you submit the backup job. The job engine polls the database every minute. When a policy has been met, the migration job will be spawned to start the copy from disk to tape. At this point, the status of the master job changes to "COPYING". The copy policy will be discussed in later section.

At any point in time you want to know what migration jobs are scheduled, you can run migration reports which will be discussed in later section.

In r11.5, the Activity Log is modified. It now follows the same parent-child relationship design for better organization of the logs. As seen in Figure 4, the staging job has the two child jobs underneath and you can expand them to see their perspective log entries. To activate this view, right click on the log and choose "Organize..." Then pick the "Using Message Grouping" checkbox and make sure "Group by Jobs" is chosen.
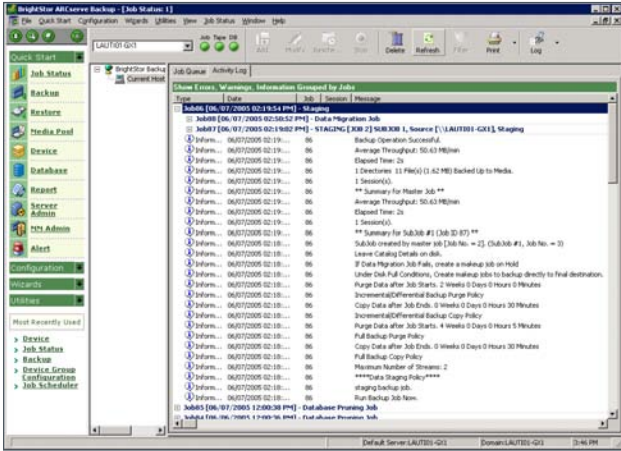

Figure 4. Activity Log Showing Parent-Child Relationship for a Staging Job.

Policies

So how does BrightStor ARCserve Backup know when to copy data from disk to tape and purge the data to free up the disk? These two operations are based on policies that are covered in the following sections. The policies are accessible from the "Policy" button in the Staging tab in Backup Manager as shown in Figure 5: Data migration and purge are driven by policy.
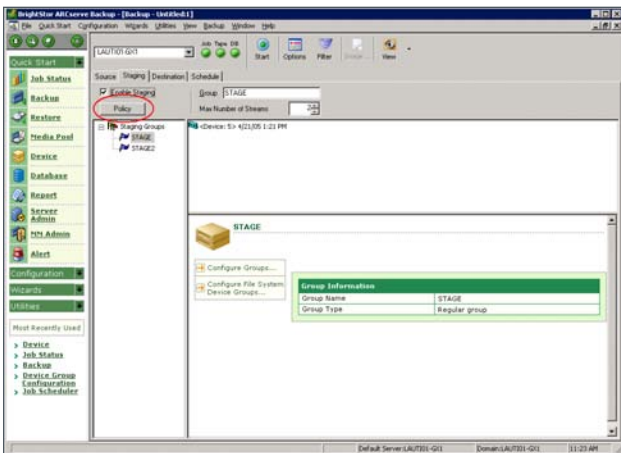

Figure 5. Data Migration and Purge are Driven by Policy.

Copy Policy

By default, data is copied from disk to tape 30 minutes after a backup job ends, either full or differential/incremental. You can change this in the "Copy Policies" section as shown in Figure 6. The options are quite self-explanatory. One interesting note about the "Do not copy data" checkbox, as the name says, data will never be copied from disk to tape when this box is checked. If you were a file system device user and using GFS backup schema, you would need to set up at least 21 file system devices for the job. This cumbersome effort can be circumvented by using this checkbox for both full and differential/ incremental backups.

Purge Policy

On the same screen shown in Figure 6, the "Purge Policies" section defines when data will be deleted from disk in order to free up more disk space. By default, data is purged 4 weeks and 2 weeks after job ends for full and differential/incremental backup, respectively.

The purge process happens in the background without showing it in the job queue. It will, however, show in the Activity Log. Even when the master backup job is deleted from the job queue, the purge policy associated with it will still kick off as scheduled.
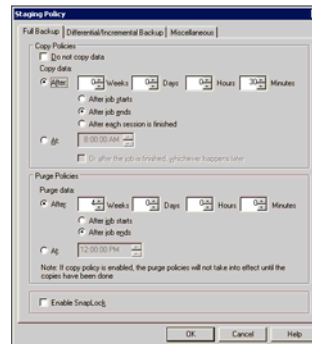

Figure 6. Staging Policy.

SnapLock Support

SnapLock is a feature for some Network Appliance NAS filers. It provides WORM capability for the filers. Disk staging supports SnapLock with a simple checkbox.

Restore

One of the benefits of disk staging is fast recovery time. Like backup, data is retrieved much faster from disk than tape. There is no change in terms of user experience when restoring from disk. Users would go through the same procedure when performing restores. It is transparent to them whether the data is retrieved from disk or tape.

In r11.5, a new feature called Smart Restore is designed to increase the overall success rate of restore operations. With disk staging, it is possible to have the same information or backup session in different backup locations such as disk and tape. If during the restore process, a media or drive error occurs, Smart Restore internally searches and finds the alternate media and starts restoring the data from the alternate media.

For example, if during a restore job, the restore source media jams and disables the library. BrightStor ARCserve Backup would then automatically search for duplicates of the backup session. If a duplicate of the session exists, regardless of whether it exists on a file system device or another media, the restore operation continues without user intervention. Note that if a second media error occurs during the restore job, the job will fail.

For a session that are duplicated in different media, customers have the flexibility of choosing where to restore the data from. In Restore Manager, choose Restore By Tree. In the Version History dialog box, there is a button called "Duplicates". It shows you all the duplicate media for that particular session. As shown in Figure 7, this session resides on two media, "DISK STAGING" and "TAPE".
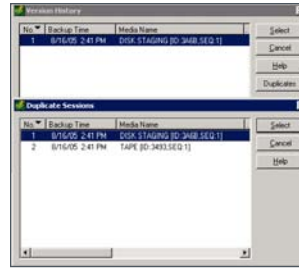

Figure 7. Duplicate Sessions.

Reports

There are four pre-canned disk staging reports.

- Disk Staging Migration Report shows migration status of data backed up using data staging jobs.
- Disk Staging Purge Failed Report shows sessions whose data failed to be purged from staging area.
- Disk Staging SnapLock Report shows SnapLock status of data backed up using data staging jobs.
- Disk Staging Summary Report shows summary status of disk staging jobs by job number and job execution duration.