



Snapshot – Point-in-time Copy in a Flash

1. Abstract

Snapshots are fast point-in-time copies of volumes. Taking snapshots is so fast that the “Backup Window” problem bothering most IT people virtually does not exist. Up to tera-bytes of data can be backed up in less than one second. Not impacting server performance and productive activities, it is the ultimate solution to satisfying short-term backup needs.

2. What is snapshot?

Snapshots are read-only copies of file-systems at a specific point in time. Snapshot distinguishes itself in its speed. Not copying any user data, creating a snapshot usually takes less than a second.

The concept of snapshot is very different from tape backups. Data are not copied to any media during backup. Instead, it just informs the NAS that all the data blocks in use should be preserved, not being overwritten. That is why it can be so fast. The “copy”, or “backup” occurs during everyday file access. When a file is modified after a snapshot is created, its original data blocks are protected from being overwritten. The new updates are written to a new location. The file-system maintains records and pointers to keep track of the snapshot data and file changes.

3. The Advantages of NASTorage™ Snapshot

i. Shrinking the “Backup Window”

Usually IT people spend a lot of time, resource and efforts in dealing with data backups. One of the biggest challenges is backup window. In the case of traditional tape backups, usually the server is under heavy load when doing backup jobs. What makes it worse is the long backup time, which is often known as the “backup window” issue. When data grow to several tera-bytes or more, it becomes a mission impossible to



make backups within limited time and budget.

The snapshot technology solves the backup window problem. It usually takes much less time than traditional tape backups or even disk-to-disk backups, which are usually involved with data copy or replication. Taking a snapshot does not copy any user data. It just puts a record in the NAS that all data blocks at this point of time should be preserved. Snap! That's it! All process takes less than a second.

ii. Solving the “Open File” Issue

The “open file” issue is also a big challenge to backup software. One of the reasons why open files cannot be backed up is that users keep updating them during backing up. The middle of a file might have been updated while the backup software read through the beginning of it. It often causes corruptness in data integrity.

To take a snapshot is to make a point-in-time image copy of a file-system. When creating a snapshot, the file-system is temporarily frozen and does not allow any writes to the file-system. The freezing will not be a problem because it only takes a fraction of a second. Since files are not modified during snapshot creation, data integrity is maintained.

iii. Retrieving Previous Versions of Files Without Intervention of IT People – Easy and Fast

Snapshot backups are kept on the hard disks of a NAS server. Different from tape backups, which are off-line backups, snapshot data are on-line.

The NASStorage Snapshot function puts additional special folders in the file-system – called snap folders – for clients to access their previous versions of files. In the snap folders are read-only snapshot backups. Users can browse the snap folders just as they do to their own data. To restore a file, simply copy it from the snap folders. It is easy and fast.

iv. Make Frequent Backups Possible

It is possible and recommended to take snapshots often. Hourly snapshots are easy jobs. Even several times an hour is fine. Tape backups cannot do so because of long backup time and the concerns of performance impacts.

Another issue of frequent backups is about the usage of disk space. This is not an issue on snapshot backups because the disk space used by snapshots is not determined by backup frequency, but the amounts of files changes. The more file changes are, the more disk space it takes. It has nothing to do with backup frequency. If 100 snapshots are taken, but no files are changed, it does not use any extra disk space.

v. Space-Saving

The snapshot keeps track of file changes in blocks. A block contains 4K-bytes of data. When a block is updated, it writes to a new block. For example, a 20M file is consisted of 5,000 blocks. If 1MB of that file is modified, it only uses additional 250 blocks of disk space to write the changes. In contrast, tape backups or Windows FRS (File Replication Service) backup data file by file. If 1MB of the 20MB file is updated, all 20MB will be copied to tape or replicated to remote sites.

4. The Magic of Snapshots – How It Works

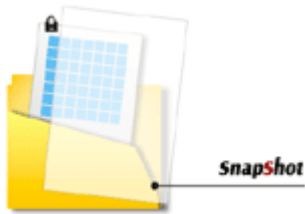
How can snapshot creation be so fast? Actually the magic comes from the fact that it does not really “back up” the data when creating a snapshot version. It simply writes a record to inform the file-system to preserve the data blocks which are currently in use. When users try to modify a file, it does not overwrite the original data blocks. Instead, it writes the new updates to a new location, keeping the original data blocks unchanged. The original data blocks belong to the snapshot backup, not being seen by the current active file-system. This is called the “Copy-On-Write” (COW) operation.

Please see the figure below about the operation.

1. Before snapshots are created, the file-system works as non-snapshot file-system does. Modifications to a file actually update the data blocks in use by the file.



2. After a snapshot is created, all the data blocks in use will be preserved, not being overwritten.



3. When users modify a file, it writes the updates to a new location.



4. Both the original and the new version of the file are available on disks.



5. Snapshot Version Control

Same as tape backups, media resource (disk space in snapshot's case) are valuable and should be recycled based on backup policy. Version controls are used to free



up media resource automatically.

All NAStorage snapshots are named by their types.

- hourly
- daily
- weekly
- monthly
- manual
- auto

The first four types are those snapshots created by snapshot schedules. ‘manual’ snapshots are created manually by admin. ‘auto’ snapshots are created by NAStorage internal backup software.

Version controls are implemented by limiting the maximum number of snapshots for each type. For example, administrators can choose to keep the latest 24 hourly snapshots. If a new snapshot is created when there are already 24 hourly snapshots, the oldest snapshot will be deleted automatically to release disk space.

6. Comparison Between NAStorage Snapshot and Windows VSS (Volume Shadow Copy)

	NAStorage Snapshot	Windows Volume Shadow Copy
Snapshot Creation	Less than one second	About one minute
Maximum Number of Snapshots	256	64
Restore files	Using special snap folders – provides uniform access to all kinds of clients	Using the Shadow Copies of Shared Folders service. Availability of the service: 1. Windows Server 2003 – default 2. Windows 2000/XP – must install separately on every client 3. Windows 95/98 – not available

Snapshot Version Control	Yes. Can limit the maximum number of hourly, daily, weekly and monthly snapshots respectively.	No version controls
Pre-allocated Disk Space for Snapshot	Not necessary.	Yes

7. Back Up the Snapshot Data!

Although the snapshot technology provides a powerful backup solution, it cannot replace existing backup solutions, like tape backups or remote data replication. Since snapshot backups are kept on the same disks where user data reside, snapshot backups will not be available if, for example, the RAID volume crashes. It is still necessary to back up data using tapes or remote replication to protect from disk crashes or natural disasters.

i. Using NASTorage SmartSync, NASTorage Tape/Tape Autoloader Backup

The NASTorage built-in backup solutions take advantages of the snapshot technology. When a backup task is activated, a snapshot is taken automatically. Then the backup software read data from the snapshot area instead of from the active file-system. Since snapshot data are read-only and static, there will be no open-file issue.

ii. Using 3rd-party backup software

3rd-party backup software, like CA ARCserve or Veritas NetBackup, can also benefit from the NASTorage Snapshot function, which provides read-only and static data for backups.

When snapshots are created and the snap folders are made visible, all the snapshot backup versions appear as sub-folders under the snap



folders. Among them are hourly-latest, daily-latest, weekly-latest and monthly-latest. Those folders are virtual links pointing to the latest versions of their schedule types. They provide a source for 3rd-party backup software to read static data from.

For example, when setting up a weekly backup job in CA ARCserve, choose the data in the weekly-latest snap folders for backup, instead of the data in the normal (currently active) file-system.

When restoring data, it is not possible to restore data to their original location since snap folders are all read-only. Alternatively, choose to restore data to a new location, and then move the restored data to the original location.