White Paper for Data Protection with Synology Snapshot Technology

Based on Btrfs File System

Synology®

Table of Contents

Introduction

Data loss is a real danger, no matter what storage device you use. Important files – such as your business files or critical work documents – deserve a great backup strategy to avoid unexpected hardware failure, natural disasters, or simple accidental deletion. Let DSM help with robust and flexible data protection situations.

Low RPO (recovery point objective), low performance impact data protection with instant recovery has been desired by IT admins for a long time. By 2016, 20% of organizations, up from 7% in 2014, will employ only snapshot and replication techniques, abandoning traditional backup/recovery.[1]

This white paper is intended for Synology customers, and partners looking to understand features that exist in the Synology NAS product that can learn the Synology's solution to data protection challenges and ensure maximized data availability.

[1] Gartner 2014 Magic Quadrant for Enterprise Software and Integrated Appliances.

Data Protection Technologies

Btrfs File System

Synology has already provided some advanced RAID technologies, including SHR (Synology Hybrid RAID), RAID 1, RAID 5, RAID 5 and RAID Groups to ensure maximum data availability in case any drives fail. On select models[1] targeted for businesses and enterprise, a new file system Btrfs is introduced to ensure the data correctness and strengthen data protection.

- **Two copies of metadata** - For a file system, metadata is critical as it includes all of the internal data structures of the file system, including folder structures, filenames, permissions, checksums, and the location of each file. On a Btrfs volume, two copies of the metadata will be stored on the data volume for recovery.
- **CRC32-C Checksums** - Checksums will be generated for user data and metadata to avoid silent corruption. If a checksum mismatch is detected during a reading process, it will first try to recover automatically, by obtaining a good copy of this block for metadata. If no good copy is available, an error will be reported to the user.
- **Silent data corruption detection and recovery (file self-healing)** - In DSM 6.1 and above, Btrfs file system has the ability to not only detect the silent corruption on the user data but also to recover it. If checksum mismatch is detected, the file system layer instructs MD layer to read redundant copy (parity or mirror) of the damaged block. It will then automatically recover the damaged block with the good redundant copy. This file self-healing feature requires the RAID volumes to run RAID 1, 5, 6, 10, F1, or SHR.
- **Instant server-side copy** - In DSM 6.1 and above, the Btrfs fast-clone feature leverages the copy-on-write technology to make instant file copy when the source and destination are both on the same Btrfs volume. This feature supports File Station, SMB, and AFP protocols.
- **Snapshot Technology** – Btrfs file system works on a copy-on-write mechanism, so system administrator can take snapshots of shared folders to preserve the history of the shared folder. Taking snapshot is done within seconds so IT administrators can protect the data quasi-continuously by taking a snapshot very frequently, up to every 5 minutes. To save storage, IT administrators can define the retention policy so that earlier snapshots will be automatically removed by the system. End users can find the previous versions of files and folders, including the deleted ones, via Windows File Explorer, File Station in Synology DSM or all file protocols, to recover the files on their own instead of consulting to IT helpdesk.

Snapshot Technology

The “snapshot” is a point-in-time copy stored in the same volume used to record the whole data status at the time upon being taken. Snapshots use only a small amount of additional storage space, and exert little impact on system performance.

With the snapshots, if a user accidentally modifies or deletes data on a shared folder with snapshots, you are able to quickly restore the data back to the previous time at which the snapshot was taken. In addition, it allows users to recover their own deleted or modified files in shared folders without assistance from the administrator.

[1] Please refer to the check if Btrfs is listed as supported file system in the Product Specification page.

In Snapshot Replication on Synology NAS, you can take manual (on-demand) snapshots, enable snapshot schedules, and perform snapshot restore and cloning operations. For scheduled snapshots, you can specify regular time intervals to take snapshot automatically.

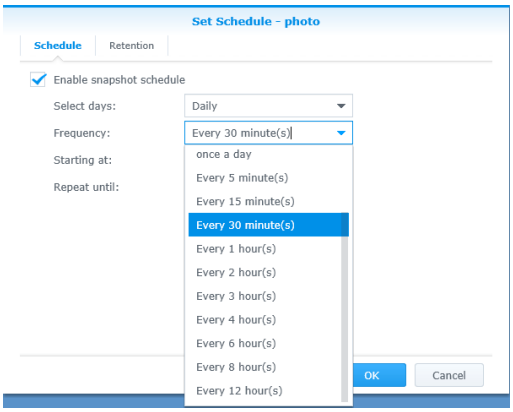

Figure 1: Configure the frequency of taking snapshots

Using snapshot to keep multiple versions of data is different from traditional backups which can take hours to complete. You can take a snapshot for every 5 minutes to reduce RPOs. You can retain up to 1024 snapshots for each shared folder and up to 65536 snapshot versions of all the shared folders. What's more, Synology NAS performs snapshot instantly and can be scheduled for many more point-in-time backups per day than backup to drives with little performance impact.

You can restore a shared folder to a previous version of a particular date and time when you need it.

How shared folders snapshot works

To take a shared folder snapshot, the file system of the data volume must be Btrfs. Snapshots for Shared Folder (and also VMware NFS Datastores, if used as a Datastore in VMware) is performed on the entire folder; Snapshots are based on a copy-on-write mechanism.

The following figure illustrates the shared folder view and snapshot view if the block is modified after the snapshot is created. These snapshots are read-only, so data must be restored or cloned as a new shared folder before it can be written or edited. A snapshot resides on the same volume of the shared folder, helping IT to keep multiple versions of data while only the different data blocks will occupy drive space.

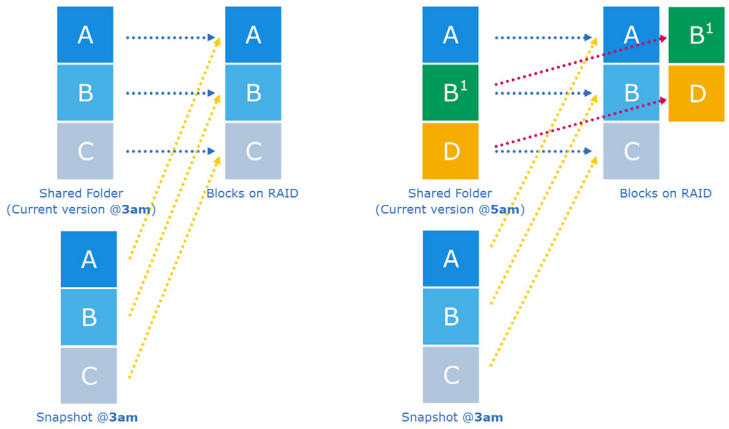

Figure 2: The relationships among shared folder, snapshot and actual data blocks in RAID

Custom Scripting for Snapshot

If your data stored in the shared folder is application data, for example, database and its logs or virtual disks for running VMs, you may want to do some custom scripting to integrate your applications with Synology shared folder snapshots, to make these snapshots application consistent. Synology provides a set of shared snapshot commands to allow you to automate the process of taking snapshots in your customized scripting. To view the supported commands, please enter synosharesnap – help in the SSH console.

Retention Policy

Retention policy allows you to specify the maximum amounts of snapshot versions to save your storage space, but you may need to retain your snapshots for longer periods of time. Synology employs the GFS, or Grandfather-Father-Son retention policy. You can configure the maximum amounts of snapshot versions to be retained for the following time ranges respectively: hourly, daily, weekly, monthly, and yearly.

Figure 3: Define the number of hourly, daily, weekly, monthly and yearly snapshot to keep in the system

Synology Snapshot Replication will retain as many as the configured maximum number of snapshots for each time range. If more than one snapshot version exists within a time range, only the earliest one will be kept. For example, if you

set a policy as 10 weekly snapshot(s), Snapshot Replication will retain the earliest snapshot (if more than one snapshot is taken in a week) for each of the latest 10 weeks.

Moreover, all snapshot versions within one hour since the taking of newest snapshot will be kept by default, for your convenience to find and restore from recent snapshot versions. You can also lock a snapshot to prevent it from automatic removal by your retention rules.

Self-Service Recovery

The end users, if they are granted with the access permissions when the snapshot is taken, can access and restore a file that was accidentally deleted by accessing the snapshots from all file protocols (SMB, FTP, AFP, File Station, WebDAV, NFS) to the shared folder by adding “\#snapshot” to the end of the shared folder path. If users are connecting to a shared folder via CIFS/SMB, they can browse the snapshot using Windows File Explorer under the Previous Versions tab in the share’s Properties window to perform the restore. This reduces the workload of IT administrator because users can perform self-service recovery on their own instead of consulting to IT helpdesk.

While snapshots are taken on a shared folder basis, users can restore a single file or folder, instead of the whole shared folder, allowing more granular recovery.

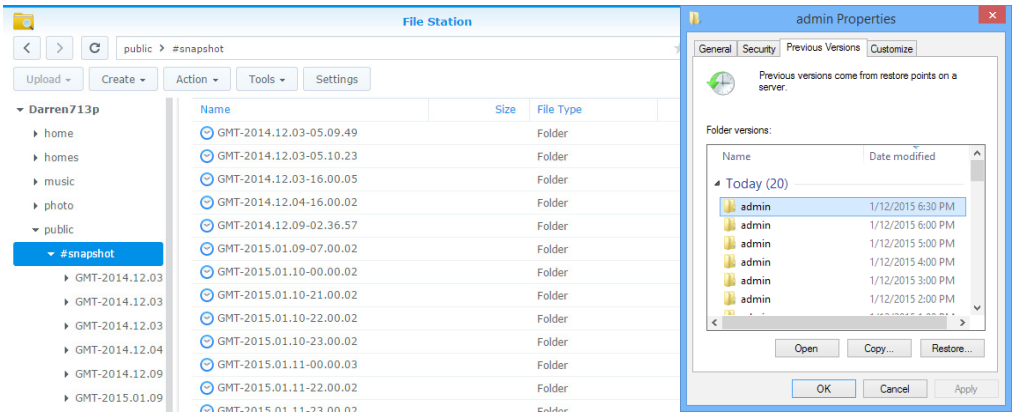

Figure 4: Recover files from File Satiation in DSM or Previous Versions in File Explorer

Integrate Snapshot Technology with Synology Applications

Multiple Tiers Data Protection

Synology provides a set of features and technologies that helps IT admins to achieve data protection, backup, disaster recovery purposes to recover data easily when it is needed. For different purposes, different technologies can manipulate:

- **Snapshot:** Features high performance and rapid snapshot & point-in-time recovery, while previous versions of data will take online, primary storage, which probably implies higher costs if high-performance disk or SSDs are used. Up to 1024 snapshots for a single shared folder. The retention policy can be defined by specifying the number of hourly, daily, weekly snapshots to keep in the volume.
- **Cloud Station:** Features one-way (by controlling the shared folder permission) or two-way, near-synchronous file shared folder synchronization between multiple Synology NAS servers; Data can be replicated for content distribution, co-authoring and site recovery purposes; Up to 32 versions of files can be retained for recovery.
- **Shared Folder Sync:** Features one-way, can be configured as semi-synchronous or scheduled file shared folder synchronization between multiple Synology NAS servers; Data can be replicated for content distribution and site recovery purposes. Only the latest version will be kept.
- **Backup to Volume:** Streaming backups for longer term retention or lower tier SLA requirements. It supports data deduplication, unlimited number of restore points to keep, and can also be rotated by simple rendition policy (keeping the latest N recovery points) or smart recycle (keep hourly backups for 24 hours, daily backups for 30 days and all weekly backups as long as free space is available)
- **Backup to Cloud:** The longest retention for business requirements, archives, and off-site. Takes longer time for backup and recovery to reduce the total cost of data backup.
- **Cloud Sync:** Supports one-way and two-way synchronization and use Public Cloud File Storage (Dropbox, Google Drive, OneDrive, and more) for disaster recovery and backup purposes.

Synology has integrated snapshot technology with Backup to Volume and Cloud Station to improve the data consistency and storage efficiency.

Integration with Cloud Station

Snapshot technology is leveraged when Cloud Station and file versioning are enabled to achieve better storage efficiency. Before Btrfs file system is introduced, file versioning in Cloud Station will occupy the double size of the shared folder to store previous file versions, one in file system and one in Cloud Station database. After Btrfs is introduced, Cloud Station will take advantage of the snapshot technology to get the previous versions of files, so the storage required will be significantly less than ext4 file system.

Figure 5: Cloud Station takes advantage of snapshot technology to enhance the storage efficiency

Integration with Hyper Backup

For some applications, such as database logs or indexes, the consistency between files is critical to ensure the backed up files which can be used for successful recovery. Before snapshot technology is integrated with Hyper Backup application, backups may be inconsistent due to continual changes of data. For example, if a file is removed or changed during the backup in progress, the result of backup cannot be predicted because the time each file being backed up will be different.

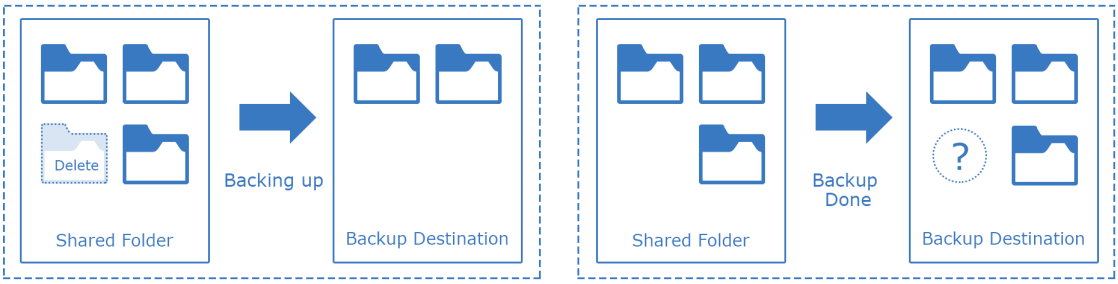

Figure 6: Without snapshot, backups may be inconsistent since data is always changing

Snapshot technology solves the issue by taking a snapshot when a backup process started, and backup application will transfer the data in the snapshot to the backup destination, and finally delete the snapshot after the backup process is completed. As a result, files will be backed up at the point in time when backup started, even though the files are changed or deleted afterwards. This is useful for some applications for which data consistency is critical.

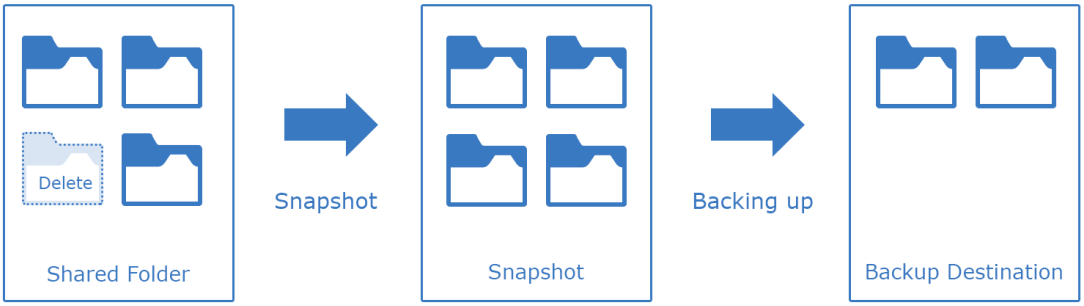

Figure 7: When a backup task started, a snapshot will be created and will be used for data backup to ensure data consistency

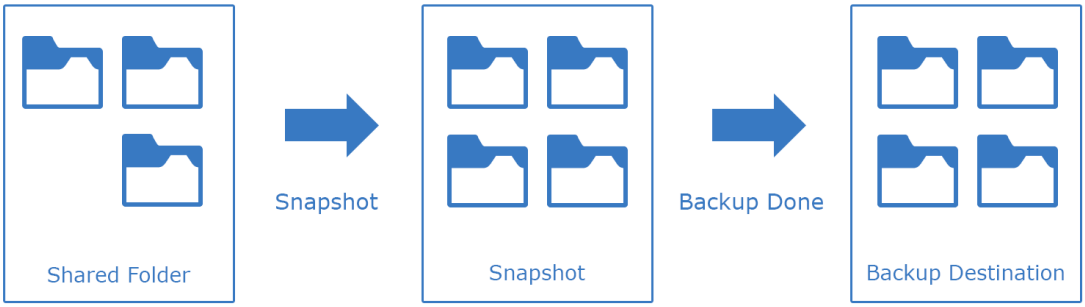

Figure 8: Regardless of the changes happened to backup source, the data backed up will be consistent to the point of time when backup task started

Summary

Snapshot completes Synology's Multiple Tiers Data Protection for SMBs and Enterprises. With the introduction of Btrfs file system and snapshot technology for shared folders on Synology NAS, IT administrators can protect their data more often by configuring a frequent scheduled snapshot, protect more by specifying the retention policy of snapshot and keep more recovery points. IT helpdesk workload is also lessened because users can access the previous versions of files and perform self-service recovery. For more information, please contact Synology at **www.synology.com**.