TechFAQ

The Interesting Challenge of Backing Up NAS

Released October 2017 v3

The Interesting Challenge of Backing Up NAS

Rubrik is pleased to announce support for backing up Network Attached Storage (NAS) such as NetApp, Isilon, VNX, and more with our 3.0 Firefly release and continuing enhancements in the release since then. Given NAS support being one of the most common customer requests, we are excited to announce an innovative solution for managing NAS data that provides faster Recovery Point & Recovery Time Objectives (RPO & RTO), incremental forever backups, and cloud archive capabilities without sacrificing storage efficiency or tying the customer into a complex NDMP-based (Network Data Management Protocol) solution. To be clear, this is an approach that does not require NDMP or "proxy VM's".

To provide some variety from regular papers, we'll use an FAQ format for this discussion.

WHAT IS NAS?

NAS is one of the two main mechanisms for serving storage over the network. Typically sold as an appliance, NAS is known for its simplicity in serving files via protocols such as NFS or SMBv3 as opposed to a SAN serving blocks via FC, FCoE, or iSCSI. Another difference is that NAS is served over ethernet, whereas SAN is typically served over FCP (albeit with some adoption of iSCSI or FCoE). A final key differentiator is the location of the file system. For NAS, the file system is controlled by the NAS appliance with the file system features often being a key competitive differentiator. For SAN, the file system is controlled and put in place by the host connecting to the SAN.

WHAT MAKES NAS SO DIFFICULT TO BACKUP?

Two characteristics make NAS backups difficult:

1. It is not possible to install an agent since NAS is sold as an appliance.

2. NAS systems and datasets have grown in size with the industry. When NDMP was first introduced, a very large NAS dataset might have been 10 TB. Today, NAS datasets typically tend to be huge running into 100's of TBs to PBs.

For these reasons, traditional backup vendors using ancient approaches like NDMP have failed to keep pace in providing fast, incremental forever backups, efficient storage, and fast RTOs.

In response to these needs, we have developed an innovative solution for managing NAS data that provides faster RPO/RTO, incremental forever backups, without giving up on storage efficiency or locking the customer into a proprietary solution.

WHAT RUBRIK CAPABILITIES ARE SUPPORTED FOR NAS?

We will be supporting all of the capabilities that the Rubrik platform provides including:

- Backup & Recovery
- Disaster Recovery (DR) Replication+Restore
- Archive to Cloud and Local Object Storage
- Encryption - Replication, Cloud Archive

All of the core values of the Rubrik platform also underlie our NAS capabilities, including:

- A simple, intuitive policy-driven interface
- REST API-driven functionality
- Infinitely scalable architecture
- Storage efficiency
- Global file catalog & search

WHAT VENDORS AND PROTOCOLS ARE SUPPORTED?

The approach we are using is intentionally vendor agnostic and hence not bound to any particular vendor. Using this approach we are able to support all NAS vendors, including NetApp, Isilon, HNAS, VNX, SONAS, and more.

Most NAS filesystems are exposed through either NFS or SMB (sometimes referred to as CIFS) protocols. We are able to backup filesystems exposed through either protocol and preserve the corresponding metadata including the ACLs.

In the same vein, we are agnostic to the NAS protocol version. We support NFS 3.x, NFS 4.0 and NFS 4.1. Similarly, for SMB, we support SMB 1.x, 2.x, and 3.x.

HOW DOES IT WORK? WHAT IS OUR APPROACH?

Overall, the workflow includes creating filesets (the selected folders and files to be protected) and setting include/exclude information when applicable. See below for additional information on filesets.

WHAT ARE THE ADVANTAGES/DIFFERENTIATORS OF OUR APPROACH?

There are several advantages of our approach:

- **Vendor Agnostic** - It works with all NAS vendors and with all NAS protocols. There are no vendor specific plugins and no need to keep up with changing vendor proprietary formats and APIs.
- **Native Format** - The data is stored on Rubrik in the "native" format. This is in contrast to an opaque format like NDMP, which needs to be "unpacked" before any restores can take place. In our approach, the backup data is instantaneously recoverable. This results in significantly reduced RTO.
- **True Incremental Forever** - Our approach supports "incremental forever," obviating the need for periodic full backups (unlike NDMP). This reduces backup windows (shorter RPO) while making efficient use of the network bandwidth as well as the storage on the secondary system.
- **Full Rubrik Capabilities** - As mentioned before, our approach to backing up NAS gives you instant access to all of the data management capabilities of the Rubrik platform, including replication to DR, cloud archive, global search, erasure coding, and more.

WHY ARE WE NOT USING NDMP?

NDMP is a 20-year-old protocol that was developed in an age when the datasets were small and the destination for backup data was tape. NDMP is only a control protocol and does not dictate the format of the backup stream. As a result, each vendor sends the backup stream in a proprietary format that is usually not meant to be unpacked.

To keep your RTO at a reasonable level, NDMP requires periodic fulls, resulting in wasted bandwidth and space on the secondary system. Our approach sidesteps all these issues and provides superior RPO/RTO plus the full suite of data management capabilities (see above).

WHY ARE WE NOT USING VENDOR NATIVE SNAPSHOTS?

Another approach to backing up NAS is to make use of the vendor's native snapshots. This approach is popular with other backup products.

At a high level, the disadvantage with this approach is that most NAS vendors require the backup data to be stored on their own arrays which results in a significantly higher TCO. Additionally, this results in a complex solution because the backup software and the secondary storage have to be managed through separate management interfaces.

At a deeper level, disadvantages center around performance inconsistencies between NAS

vendors having different snapshot engines, complex development SDK's to call these snapshot engines (when supported at all), and typically requiring a replicated copy of the data to the same NAS platform limiting and restricting ability to manage TCO. In addition, NDMP implementation is complex from a development and troubleshooting perspective and has to be developed individually to work with the NAS platform of choice. In stark contrast, using a standard Network Protocol like NFS or SMB to directly access the data to be protected allows ingest to be widely distributed across multiple nodes' interfaces to increase backup performance as well as deeper visibility into the data being protected.

As noted above, many large NAS vendors advocate for a simple Snap and Replicate solution. This leads to a potential 2x storage cost while becoming locked into a single NAS vendor. Using Rubrik paves the way for a vendor agnostic approach to NAS protection since you can backup from one NAS and restore into another while maintaining history in the cloud.

Some backup products focus on this solution because unlike Rubrik they do not have an integrated, hardware+software+storage solution and ignore the true storage cost in their TCO.

We believe the Rubrik approach is superior in that it provides all of the benefits of native NAS snapshots (faster RPO/RTO, great storage efficiency, etc.) without the prohibitive cost and complexity of the NAS array on the secondary storage.

ARE THERE ANY LIMITS OF THE CURRENT SOLUTION?

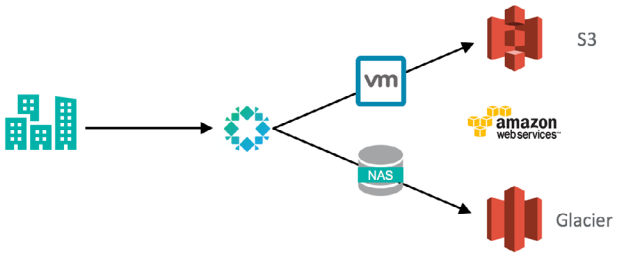
We are often asked if there are any limits such as size of data set or the number of files that can be protected. Generally speaking, as long as the sizing guidelines are followed, there really are not any scalability limits to the size of the data sets or the number of files that can be protected. The solution scales linearly as filesets are ingested in parallel across the entire cluster.

ARE THERE EXAMPLES OF EXISTING SUCCESSFUL IMPLEMENTATIONS?

We have several key customers already using this solution to backup their NAS. It is common for customers to have 50-100 TB of documents and ~40 million total files. One very large customer is using this solution to backup 1 PB of NAS with more than 1 billion files. This makes us confident that the solution is scalable and robust.

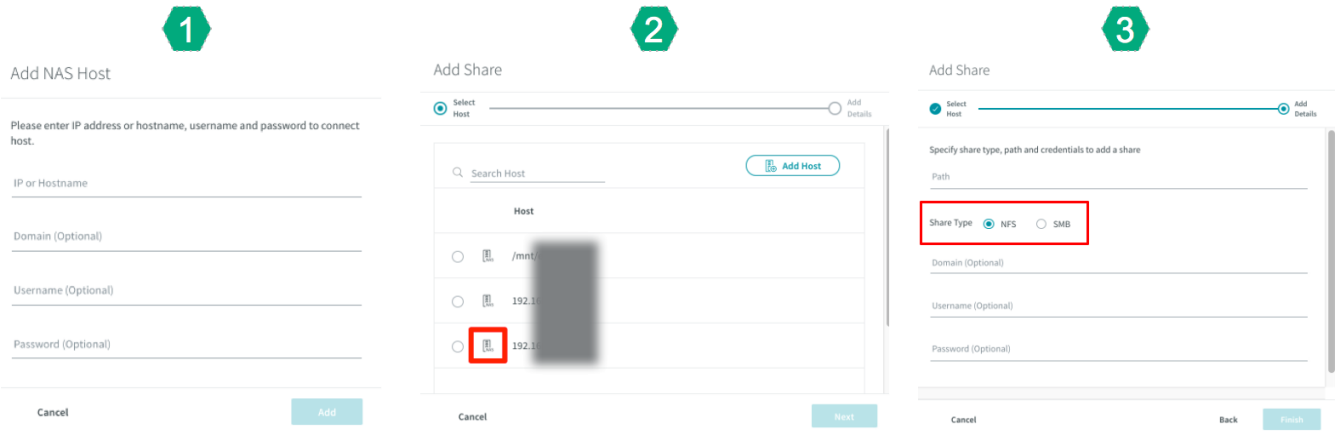
ARE WE ABLE TO CONFIGURE SLA DOMAINS TO ARCHIVE NAS DATA TO GLACIER AS A LOWER COST ALTERNATIVE?

Rubrik 4.1 introduces the ability to send long term archives to Glacier simply by configuring an SLA policy. AWS Glacier is extremely low cost and primarily used for long term archival workloads. As shown below, some customers may choose to send Virtual Machine archives to S3 while sending NAS archives to Glacier. For extremely infrequently accessed archives, this use case provides a 5x cost differential or greater over sending data to S3. Please see the 4.1 User Guide for more detailed information.

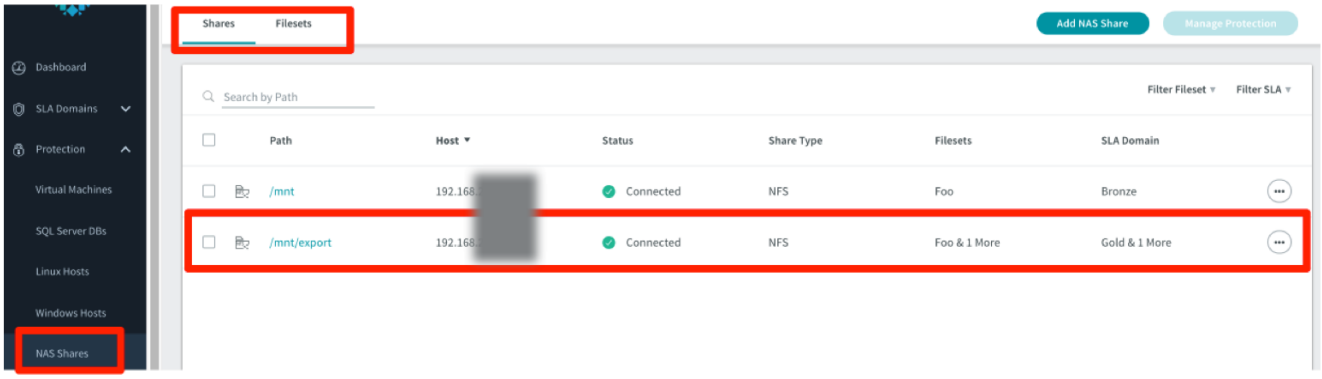


HOW DO I CONFIGURE NAS BACKUPS IN THE RUBRIK UI?

We offer a streamlined process for setting up NAS backups and having the Rubrik cluster connect directly to NAS servers. Setup is simple - add the IP address or FQDN and supply appropriate credentials.



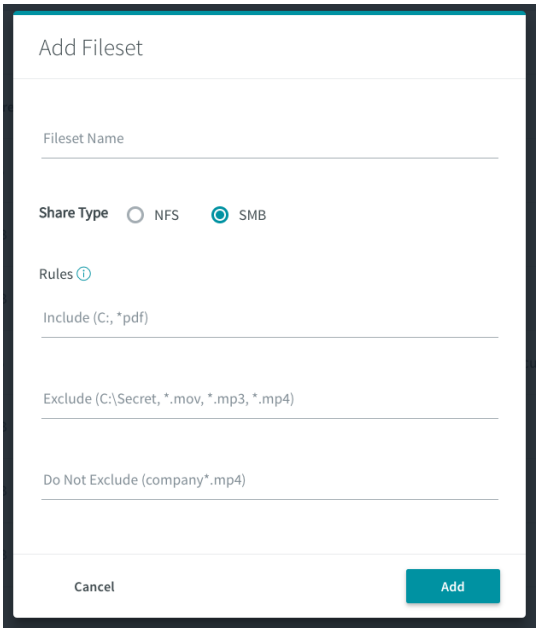
In addition, each share can have separate authentication (username/password) if needed. For simple access, there is a dedicated section for NAS in the sidebar.



From there it's regular fileset definition similar to Windows and Linux file backups.

WHAT IS A FILESET?

A fileset is a construct applicable to all file backups - a set of files and folders whether on a Linux host, a Windows host, or a NAS share. Rubrik uses filesets assigned to a host or share to determine the data to manage and protect. Rubrik further interprets a fileset based on the values provided in the Include, Exclude, and Do Not Exclude fields with each field accepting multiple values as illustrated here.



Filesets are persistent - no new full backups are required after the initial backup. Restore capabilities are what you would expect - download, restore to the original location, or export.

SLA assignment continues to match the core simplicity of the Rubrik architecture - the same SLA policies that can be assigned to VM's, Windows, Linux, or SQL can also be assigned to NAS datasets.

Please see the user documentation for more detailed information.

I'M IMPRESSED - WHAT DO I DO NEXT?

As always, just reach out to your Rubrik sales team if you have any questions.

https://www.rubrik.com/contact-us

info@rubrik.com

About the Authors

Demetrius Malbrough is a Technical Marketing Engineer at Rubrik. He is the host of the Data Protection Gumbo podcast, creator of the LinkedIn Backup & Recovery Professionals group and has been in the Data Protection industry for 18+ years. You can find him on Twitter @dmalbrough.

Andrew Miller is Technical Marketing Manager at Rubrik. He has random certs, blogs sporadically, and has lived on the customer and partner side for 15+ years. You can find him on Twitter @andriven.