

Cloud Scale with NAS Backup and NAS Direct Archive

ENABLING FAST INGEST AND AFFORDABLE BACK-END STORAGE

CHALLENGES WITH EXISTING SOLUTIONS

Enterprise NAS backup is challenging. Over the past two decades, NAS file systems have grown into petabyte (PB) scale in many vertical industries. File types and use cases have gone beyond the general-purpose office files and home directories. Increasingly NAS stores either very large multimedia files such as 4K video or billions of small files such as bank check images or IoT sensor data. Each major file type or workload often has a different performance or storage profile. The sheer size of NAS and the variety of content profiles make backup very difficult for backup admins.

As all enterprise NAS systems have proprietary file systems, there is no industry standard for a backup agent like an agent for Windows or Linux. The prevalent network data management protocol (NDMP)-based backup for NAS is not efficient enough to protect large NAS environments. Specifically, NDMP backup has the following drawbacks:



NDMP was created about 20 years ago for tape, often requiring periodic full backups whose backup windows are too long to be practical.



As NDMP is a volume-based backup, it has poor searchability and granular restore capability unless it's indexed in a staging area by backup software.



To index NDMP backups, a non-scale-out media server can be the performance bottleneck for a large file system, while multiple media servers could generate isolated catalogs and storage silos, making data search and access complex and slow.



Each NAS vendor has its own NDMP implementation with its own data format, which prevents data from being restored to a different NAS platform. This results in vendor lock in.

For large NAS file systems that have many hundreds of TBs or PBs of storage, organizations often ditch NDMP backup and use NAS-native snapshots and replication as a high-performance alternative. However, primary storage snapshots and replication are not considered a true backup, as they have the following weaknesses:



Many NAS snapshots are not isolated and immutable backup copies. Reverting to a point-in-time snapshot could render all snapshots taken after the recovery point useless. As space-efficient snapshots use virtual pointers, there is no physical copy and isolation to prevent them from being affected by malware.



Primary NAS snapshots consumes expensive storage, which is typically not suitable for retention longer than a week. The short retention could be a vulnerability explored or leveraged by ransomware criminals.



For NAS systems that do not have snapshot-based replication, there would be no remote recovery points in case of site failure. For those that do, they often do not allow different RPO and retention SLAs on the target side.



There is no central index/catalog for fast search and granular file restores.



The replication target has to be the same platform as the production NAS system, which is not compliant with the industry's best-practice recommendation of using a different type of storage as a backup target to mitigate risks.



NAS replication has its own management console (managed by storage admins), separate from the overall data center backup control plane (managed by backup admins), fragmenting the NAS data protection infrastructure.

Newer backup solutions, typically scale-out integrated backup appliances, tend to leverage NAS snapshots as the data capture mechanism and then parallel ingest incremental-forever snapshot data directly to the appliance storage repository, while performing file indexing and other operations. Such a backup method (incremental forever in combination with parallel throughput) can drastically speed up the backup performance. The drawback is that the scale-out architecture of scaling compute and storage together can quickly make the appliance too expensive to store a large amount of data.

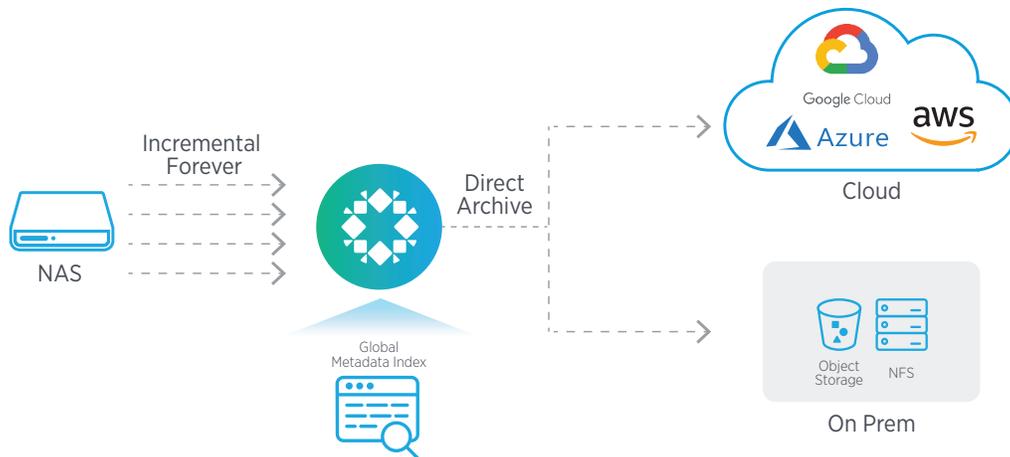
THE INNOVATION RUBRIK BRINGS WITH NAS BACKUP AND NAS DIRECT ARCHIVE

Rubrik designed its NAS data protection to achieve high ingest performance and low cost. On the performance side, Rubrik integrates with leading NAS solutions' native snapshot APIs to use their snapshots as the fastest backup data capture mechanism. It then leverages NAS-native change tracking APIs (when available for integration) to drastically reduce the time to scan for file changes. Once changed filesets are identified, they are ingested in many parallel streams into Rubrik's scale-out appliances. Unlike individual file-based incremental forever backup method, which requires a super-large index engine and database, Rubrik's fileset-based incremental forever method reduces the indexing load and increases ingest performance.

To address the cost issue associated with locking a large amount of backup data in a scale-out platform, Rubrik offers an option called NAS Direct Archive, where Rubrik's scale-out cluster only retains file metadata and passes through backup data and metadata to any third-party NFS or object storage including both on-premises and cloud object storage. This "open door" option drastically reduces NAS backup storage cost to customers with the freedom of choice. For on-premises only deployments, scale-up NFS storage or geo-distributed object storage are among the lowest-cost, but still scalable enterprise storage. For cloud deployments, organizations can enjoy the benefits of zero infrastructure maintenance for long-term retention and continuing decrease in cloud storage cost. NAS Direct Archive is especially useful for large image or multimedia files that, once created, rarely change.

Figure 1 illustrates Rubrik's innovative approach to protect NAS data.

Figure 1. Rubrik's NAS Backup and NAS Direct Archive



Rubrik's NAS data protection solutions stands out in the market with the following capabilities, which address the key limitations of traditional or alternative solutions:



High scalability with a lightweight, scale-out index engine: The solution overcomes NDMP backup's and traditional media server's limitation on performance scalability. It addresses the issue of lacking index-based fast search and granular restores with NAS-native snapshots and replication.



Separate scale between compute and storage: While the scale-out compute cluster can offer high ingest and index speed, the support of third-party back-end storage drives down storage cost at scale dramatically and eliminates vendor lock-in (as in the case with alternative backup appliance solutions or NAS replication).



Flexible options to protect different workloads: For small workloads or file applications with high change rates, Rubrik's local cluster fits perfectly to address granular RPOs with short retention. For image/media files that do not change, NAS Direct Archive creates affordable remote copies that can be easily searched with instant granular access.



Data immutability and resiliency: Data landed in Rubrik cluster is immutable to cyber-attacks and resilient for local failure, addressing the limitations of NAS-native snapshots and replication. Data landed in third-party storage is also isolated from the primary NAS and protected by their native protection schemes such as RAID, erasure coding and replication. As both backup data and metadata are stored in the archival, data remains available to another Rubrik cluster when the primary Rubrik cluster is down. Cloud storage, as well as third-party object storage with geo-distributed erasure coding, offers additional protection of local site failure.



Flexible restore: Data can be restored to the same or a different NAS platform than the primary NAS platform, overcoming the limitation of NDMP vendor lock in.

Table 1 compares Rubrik's approach with the alternative NAS backup methods.

Table 1. Rubrik's Approach of Protecting NAS Data – Best of All Worlds

	NAS Snapshots and Replication	NDMP-Based Traditional Backup Software with Indexing	Scale-Out Backup Appliances	Rubrik
Immutable Data Resiliency with Fast Search		●	●	●
Performance at Scale	●		●	●
Ease of Scaling Index and Backup Process			●	●
Flexible Back-End Storage		●		●

CONCLUSION

Rubrik's architectural design offers an industry unique solution that overcomes the main constraints with alternative NAS backup and protection methods. It drastically reduces NAS backup windows by combining fileset-based incremental forever method with fast scan and parallel ingest. Its source-side data reduction and lightweight index engine also offer a good balance between compute and storage costs. The ease of scaling compute to support large workloads and separate storage scaling via third-party storage allow customers to protect petabytes of NAS data at competitively low cost.



Global HQ

1001 Page Mill Rd., Building 2
Palo Alto, CA 94304
United States

1-844-4RUBRIK
inquiries@rubrik.com
www.rubrik.com

Rubrik delivers a single platform to manage and protect data in the cloud, at the edge, and on-premises. Enterprises choose Rubrik's Cloud Data Management software to simplify backup and recovery, accelerate cloud adoption, and enable automation at scale. As organizations of all sizes adopt cloud-first policies, they rely on Rubrik's Polaris SaaS platform to unify data for security, governance, and compliance. For more information, visit www.rubrik.com and follow @rubrikInc on Twitter.