HYCU Protects Nutanix Files

Changed File Tracking Optimizes Incrementals

A DeepStorage Technology Validation Brief

DeepStorage

About DeepStorage

DeepStorage, LLC. is dedicated to revealing the deeper truth about storage, networking and related data center technologies to help information technology professionals deliver superior services to their users and still get home at a reasonable hour.

DeepStorage Reports are based on our hands-on testing and over 30 years of experience making technology work in the real world.

Our philosophy of real world testing means we configure systems as we expect most customers will use them thereby avoiding "Lab Queen" configurations designed to maximize benchmark performance.

This report was sponsored by our client. However, DeepStorage always retains final editorial control over our publications.

Contents

Introduction

Whenever the IT market adopts a new platform, the operations teams that keeps it running has to figure out how to back it up.

They could add the new platform to their enterprise backup solution, but enterprise solutions are slow to use platform-specific APIs and optimize for the new platform. This lack of optimization means backups may be slow, and have a significant impact on system performance.

At the beginning of a new technology wave the market isn't big enough for the enterprise players of the day to address it, which opens the door for more nimble startups.

The RISC/UNIX/open systems of the 1990s brought us NetBackup, NetWorker, and TSM--today's enterprise big 3. ArcServ and Backup Exec started backing up NetWare servers but really took off in the Windows Era.

Most recently, the move to vSphere spawned vRanger, PHD Virtual, and of course Veeam.

This pattern is repeating itself as Nutanix promotes its home-grown AHV hypervisor. The first generation of AHV users had to return to the backup dark ages, backing up VMs by installing agents. The folks at Comtrade Software saw this as an opportunity; they created and then spun off HCYU, the first backup application designed to protect AHV systems and integrated with AHV's snapshots.

Over the past few months other backup vendors have added basic support for AHV, but HYCU's dedication to this platform means HYCU supports a broader set of Nutanix features faster than the others.

For this Technology Validation Brief we examined HYCU's support for Nutanix Files, the file services feature of the Nutanix platform.

As of February 2019, when this Technology Validation Brief was written, HYCU is the only backup application that supports Nutanix's Changed File Tracking API

The Bottom Line

HYCU Data Protection for Nutanix (HYCU) is the first platform-specific backup solution for Nutanix's fast-growing HCI infrastructure. While enterprise backup applications provide some support for Nutanix's AHV hypervisor and Prism management server, platform-specific solutions such as HYCU have historically been easier to use, and offer tighter integration with their selected platforms.

With HYCU this tighter integration results in:

- *HYCU-managed Nutanix snapshots as restore points*
- *Automatic protection of newly created VMs with default policy*
- *Support for Changed File Tracking*
- *Eliminates the overhead of file system scans*
- *Significantly speeding incremental backups*

Testing of HYCU to protect a 3.8 million file, 587GiB revealed:

- *Policy definitions aligned to business-like RPO/RTO*
- *Policy compliance includes estimated restore time*
- *Incremental backup file enumeration <10 sec*
- *No noticeable performance impact*

Organizations planning investment in Nutanix infrastructure should take a good look at HYCU to take full advantage of that investment.

File Services: HCI's Missing Link

HCI solutions use software to turn the storage media installed in hypervisor hosts into a resilient storage pool that the VMs running on those hosts can use. The part that's missing from most HCI solutions is file services.

Your VSAN cluster provides plenty of space for VMs, but if you need file shares you'll need another solution--and you always need file shares. Even VDI, the prototypical HCI application, needs reliable, high-performance file services to hold all the users' persona information and data files.

Some organizations run virtual Windows file servers, or file server clusters. Windows servers are more robust than they get credit for, and are the gold standard for SMB and Active Directory integration. On the other hand, providing storage services in one or more VMs means more VMs to manage and patch, and complicates cluster startup. Other organizations use enterprise NAS filers, missing out on HCI's promise of eliminating expensive, dedicated storage hardware.

Enter Nutanix Files

Nutanix Files, formerly known as Acropolis File Services, provides SMB (2.0,2.1 and 3.0) and NFS (v4) services using three or more virtual machines to create a scale-out file server. Nutanix Files does a better job supporting Windows clients than most Linux-based NAS systems. Nutanix Files uses proper NTFS ACLs and even supports the Windows file explorer's previous versions feature to empower users to restore from snapshots themselves.

File servers and shares are created and managed via the Prism UI, which makes Nutanix Files significantly easier to manage than a Windows file server cluster running as VMs. To create a new file server you just enter the VM IP addresses and size into a wizard and wait for Prism to do all the hard work.

By providing SMB/NFS access to the Nutanix cluster's distributed data store, Nutanix Files fills an obvious gap in the HCI's feature set.

Introducing HYCU

As is only appropriate for a backup application that protects a hypervisor's resources, HYCU is distributed as a virtual machine image. This saves administrators from the drudgery of spawning a VM, applying patches and pre-requisites like runtimes, and installing the backup application while ensuring the VM includes all the libraries and packages HYCU depends on.

Once the image file is deployed and assigned an IP address, HYCU's HTTP interface makes it simple to complete the configuration, including connecting the backup controller to Prism or vCenter, and operating HYCU.
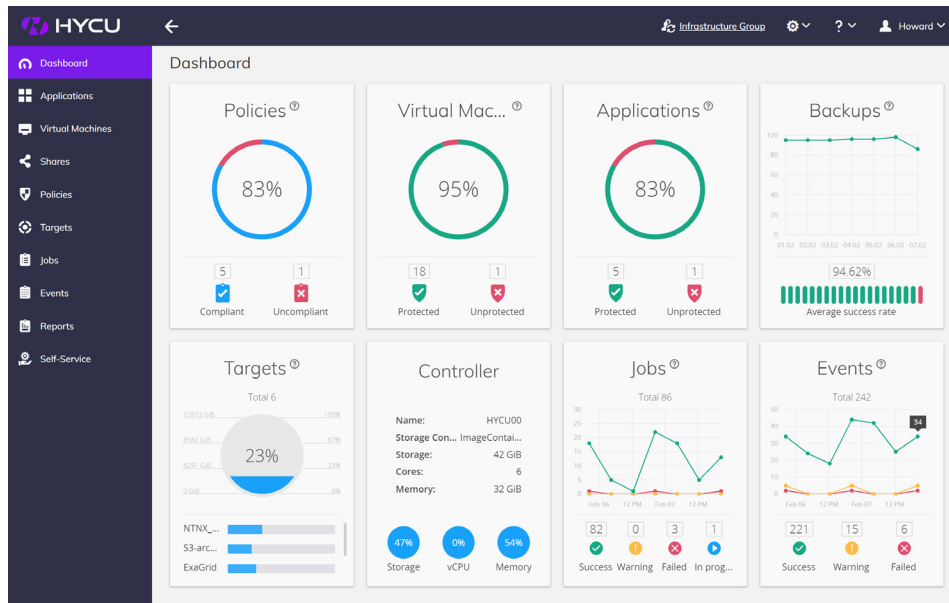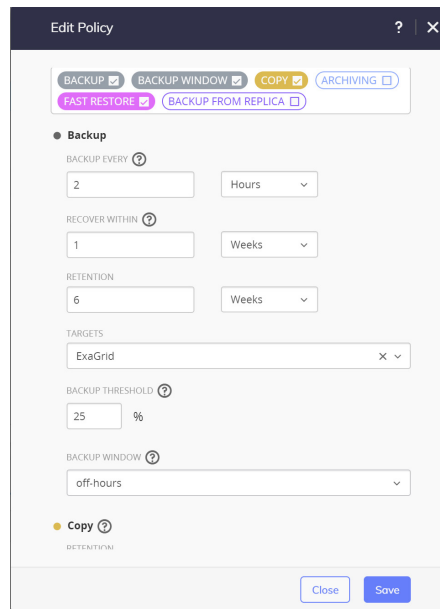

Figure 1: HYCU Dashboard

To start protecting data, just set up a backup target on an SMB or NFS share such as another Nutanix Files cluster, Azure, an S3-compatible object store, or an iSCSI LUN. (At present you can't use tape or virtual tape libraries, which is probably a good thing.) Then you just have to assign backup policies to the VMs, file shares, and applications that HYCU discovered in Prism or vCenter and direct them to a target.

Modern Backup: Policies Not Jobs

Traditional backup applications center around backup jobs. Administrators have to decide how often to perform full and incremental backups and then assign a schedule to each type. Pretty soon the backup application's console becomes a schedule of hundreds of jobs with any dependencies between them buried in notes and institutional memory.

More modern backup solutions like HYCU use policies. Administrators specify the frequency, retention, and placement of backup copies to the resources to be protected. Defining a policy only requires specifying the backup frequency, which correlates to RPO, a recovery time target, or RPO and a retention period.

Figure 2: Defining a Policy

HYCU tests the backup and restore performance as it performs each copy, and reports a resource as being out of compliance with its policy not only if a backup job fails but also if HYCU projects that a restore of this resource will exceed the policy RPO.

Rather than scheduling full and incremental backups on a static schedule, HYCU policies specify a change threshold. HYCU runs a full backup when changes exceed 25% in the default profile. That alone should reduce backup traffic significantly on the large amounts of cold data in most corporate data centers.

Some of the more advanced policy options reveal the close integration between HYCU and the AHV infrastructure. The **Fast Restore** option, for example, manages how long Nutanix snapshots are retained, while the **Backup from replica** option creates backups from replicated copies of remote VMs.

Snapshots and Changed File Tracking
The Modern Way to Backup Files

Backup applications have traditionally kicked off their incremental backups by walking the file system's metadata. They examine the last modified date or archive bit to build a list of files that have changed and therefore need to be backed up. But walking a file system takes a lot of work, which can slow real user access to files and take a long time.

It's not uncommon with large file systems for the file system scan at the beginning of each backup job to actually take longer than copying the relatively small number of changed files to the backup repository. We've seen several cases where users were not able to maintain their daily incremental backup SLA because the initial file system scan took more than 24 hours.

To avoid this problem, Nutanix has built changed file tracking APIs into Nutanix Files. Rather than walking the file system and copying changed files via SMB or NDMP, HYCU takes advantage of these APIs to perform a much more efficient backup.

When HYCU calls the Nutanix backup APIs, Nutanix Files creates a snapshot of the file system and presents HYCU not only access to the snapshot data but also a list of the files that have changed since the last time an application called the backup API.

> Conventional file scans can have a significant impact on NAS performance over many hours for each incremental backup

The snapshot provides a consistent, point-in-time version of the file system, which prevents two other problems with old-fashioned backups--open and inconsistent files. Standard file access protocols like SMB don't allow an application to access files a user has opened for writes. But some files, like Outlook PSTs are kept open, and therefore locked, as long as their applications are open. If a Sr. Executive VP never closes Outlook, his very important PSTs will never get backed up.

More insidious are inconsistent backups. A conventional backup may take minutes or hours from the time the first file is backed up to the completion of a job. If an application makes changes to two or more files while the backup is running, the backup system may backup one file before the group is modified and other files after the change. This makes the group inconsistent. Since Nutanix Files creates a snapshot of the whole file system at an instant in time, all the files are both available and consistent.

Once HCYU has access to the changed file list and the snapshot, it uses those Nutanix backup APIs to access multiple files in parallel so the data can be copied to safety as fast as the backup target can accept it.

Today, HYCU is the only backup application to fully leverage these Nutanix APIs to eliminate the overhead of a file system walk, produce complete and consistent backups, and move data in parallel. Compared to more conventional solutions, HYCU backups should run faster and reduce the impact on other application.

How we tested

HYCU's integration into the Nutanix ecosystem minimizes the overhead of protecting AHV VMs and makes it easy to maintain approximately two days of snapshot history on the Nutanix cluster to provide quick restores of full VMs, or individual files. But VM backup is just the first step in supporting the full Nutanix ecosystem.

We wanted to explore how HYCU protects Nutanix Files. Our experience tells us that while it's easy to provide basic protection of an SMB share, getting file system backup right requires greater integration between the file system and backup application.

First, we created a dataset to protect. The folks at HYCU gave us remote access to a 3-node Nutanix AHV cluster in their Boston headquarters, complete with a Nutanix Files server and HYCU instance. They also gave us access to a backup target within the HYCU environment.

Because actual backup performance depends on so many factors, and we were working on shared infarstructure, we believe the 15MB/s backup rate we saw in testing could be improved with a little tuning to HYCU and/or the generic backup target.

We created a share on the AFS (Acropolis File Server), mounted it from a Windows Server 2012 virtual machine, and ran a script to create a test file system.

The resulting file system contained:

- Files   3,831,596
- Folders   91,438
- Total size   237.3GiB or is it 586.23GiB

The initial set of data is big enough, in both number of files and total size, to make both walking the file system for changed files, and frequent full backups, impractical.

We created a backup policy to backup our file system every two hours. Then we started a script that created new files, appended to existing files, deleted files, and generally pretended to be a group of users working for a day on the same two-hour schedule.

After a couple of days we had a pretty good emulation of users working for a month, so we examined the results.

Our script was programed to create about 5% new data each run. When we looked, we saw that every 5th or 6th backup was a full copy, keeping in line with the 25% threshold specified in our policy.

We were also curious about how quickly HYCU would catalog the file share, the Achillies heel of traditional backup solutions. HCYU regularly took less than a minute from job kickoff to actually moving data to the generic backup target, with most of that time consumed by the Nutanix system creating a snapshot.

Regardless of whether we changed 100 files or 100,000, HYCU never took more than 10 seconds to catalog the share and start moving data once the snapshot was ready.

We Backup to Restore

More specifically, we backup file shares to be able to restore the files our users, and their applications, have overwritten, renamed, and otherwise deleted. Restoring files with HYCU is simple.

First, choose the share to display the various restore points available.
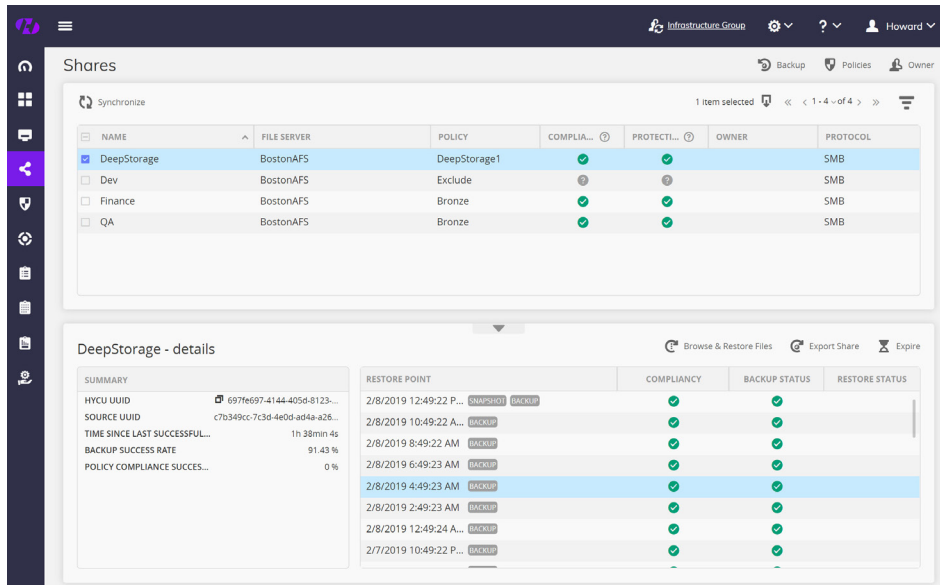

Figure 3: Selecting a restore point

Select a restore point and click the 'Browse and restore' button to open a tree view of the share at the time of the selected backup.
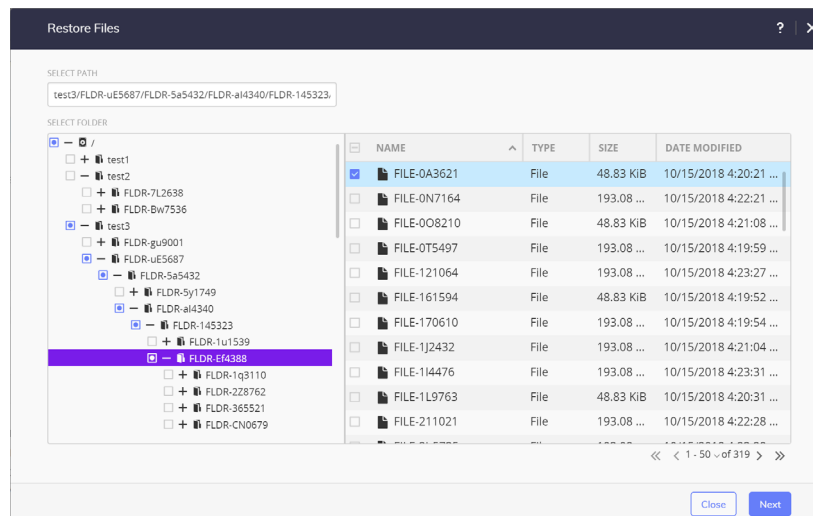

Figure 4: Selecting Files to Restore

Select the file(s) you need, then choose if you want the restored files redirected or renamed. Voila! That important contract your CEO "misplaced" is right where it should be.

The 'Export share' button, which we didn't test, will restore the whole share to another file server/NAS.

Conclusions

HYCU clearly demonstrates the advantages of a modern, policy-centric backup application that's tightly integrated with the platform it protects. Administrators can think in the RPO/RTO terms relevant to business needs to define policies, and view if the system complies with those policies.

The tight Nutanix integration lets HYCU manage Nutanix snapshots as restore points, automatically assign newly created VMs to a default policy, and leverage Nutanix' Changed File Tracking API to backup Nutanix Files shares with minimum impact.

Our testing of HYCU to protect Nutanix Files showed that it could identify the changed files for an incremental backup in ten seconds, even when we changed only 100 files in a file system of over 3.8 million files. A challenge that would have taken a conventional backup solution hours while placing a significant load on the file system scanning metadata.

Organizations investing in Nutanix infrastructure should take a good look at HYCU. Platform-specific solutions like HYCU have historically supported backup-enabling APIs like CFT six to eighteen months faster than enterprise backup providers.