

WHITE PAPER

# FLASHBLADE DATA PROTECTION

WITH RUBRIK

# TABLE OF CONTENTS

**INTRODUCTION ..... 3**

**FLASHBLADE CONFIGURATION ..... 3**

**RUBRIK CONFIGURATION ..... 4**

**UNSTRUCTURED DATASETS USED IN TESTING ..... 4**

**DATASET CREATION ON FLASHBLADE ..... 5**

**FLASHBLADE AND RUBRIK DATA PROTECTION SETUP**

**FOR DATASET #1 AND DATASET #2 ..... 6**

**DATASET#1 BACKUP RESULTS ..... 13**

**DATASET#1 RESTORE RESULTS ..... 15**

**DATASET#2 BACKUP RESULTS ..... 18**

**DATASET#2 RESTORE RESULTS ..... 21**

**CONCLUSION ..... 25**

**ADDENDUM: USING A LINUX HOST ..... 26**

## INTRODUCTION

Rubrik and Pure Storage have partnered to provide an integrated, performant, and simple-to-manage solution that will enable IT professionals to seamlessly protect and restore content for the Pure Storage FlashBlade™ array.

These complimentary products will enable customers to truly have the best of both worlds: FlashBlade for performant, dense, scale-out storage for unstructured workloads, combined with robust data protection and recovery from Rubrik that enables granular point-in-time recovery for individual files or entire massive directories. No longer do system administrators need to waste time managing multiple siloes and a seemingly endless range of products that are expensive, difficult to integrate, and provide a point solution to a single use case.

FlashBlade with data protection from Rubrik will perform and protect whatever unstructured workloads you can throw at it – be it Deep Learning, HTPC and/or EDA workloads, as well as many others. This paper will showcase real world dataset examples which confirm that setting up and managing this integration is exceedingly easy – yet comprehensive as far as meeting your data retention requirements.

## FLASHBLADE CONFIGURATION

For this whitepaper, we used a half-populated FlashBlade chassis comprised of seven 8TB blades. Because this chassis was only half-populated, note that overall available network throughput would approximately double if some of the remaining seven blade slots were filled.

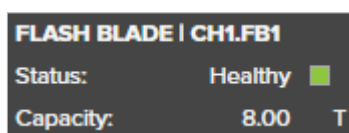
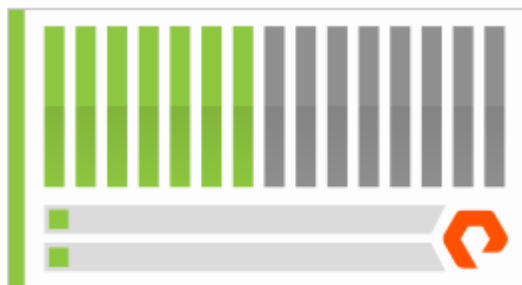


FIGURE 1. FlashBlade configuration used in this whitepaper

For more detailed information about FlashBlade and available configurations, please follow this [link](#).

## RUBRIK CONFIGURATION

The Rubrik appliance used in this whitepaper was a four-node Brik with hybrid storage comprised of a 1.6TB SSD and a 95.3TB HDD, for roughly 60TB of usable secondary storage space with overhead considered. We used firmware version 3.2, which includes native NAS share protection. For more information on Rubrik, please follow this [link](#).

## UNSTRUCTURED DATASETS USED IN TESTING

We created two different source datasets for our benchmarking activities and will provide explanation and results for both options. One dataset focuses on many different file sizes and types, and directory depths, spread across multiple FlashBlade file systems to emphasize the parallelism that both products provide for unstructured data of any type. The second dataset contains over 20 million smaller files hosted on a single FlashBlade share and is meant to simulate DL (Deep Learning) and EDA (Electronic Design Automation) data we commonly see across our joint customers.

For the first source dataset used in testing (Dataset #1), we opted to include a wide range of files of different types and sizes to provide a challenging and realistic workload for backup that would simulate a customer using a mixture of small HPC files as well as larger media and entertainment files, all residing on FlashBlade and Rubrik. The files used in our testing were comprised of multiple copies of the [Login VSI Pro Library](#) traditionally used for VDI testing and included a wide variety of PDFs, large MP4 movies, as well as Microsoft Office files. We also used [Splunk](#) index and raw data files, and larger data files of varying sizes and compressibility generated via the [Oracle VDBench tool](#). Lastly, these files were stored in subdirectories varying in depth from 1 to 10 levels.

To host the above data, we created four identical FlashBlade NFS file systems that were each about 500GB in written size and contained approximately 550,000 files to show that FlashBlade and Rubrik can execute as well as ingest a wide range of file types and sizes from multiple sources in parallel without sacrificing usability or performance. The combination of these four file systems resulted in over 2 million files and 2 TB of overall data for our testing.

For the DL/EDA simulated dataset (Dataset #2), we opted to include a huge number of files for backup in a single location that would simulate a customer using a mixture of small DL/EDA files, generally between 250KB to 2.5MB in size. The files used in our testing were comprised of multiple copies of the [ImageNet](#) 2011 and 2012 datasets, which is a large database designed for visual object recognition software challenges. The dataset is frequently cited and used as a standard for AI/DL workloads but is also similar in both file size and quantity to many EDA and HPC workloads.

In this dataset, we created a single FlashBlade NFS file system that was about 3TB in written size and contained approximately 23 million files, occupying about 3TB of space.

The table below summarizes our two test dataset characteristics:

NAME	NUMBER OF FLASHBLADE FILE SYSTEMS	NUMBER OF FILES PER FILE SYSTEM	TOTAL NUMBER OF FILES	FILESIZE RANGE	DATA SIZE PER FILE SYSTEM	TOTAL DATA SIZE	SUBDIRECTORY DEPTH RANGE
<b>DATASET #1</b>	<b>4</b>	<b>550,000</b>	<b>2,200,000</b>	<b>200KB – 1.5GB</b>	<b>500GB</b>	<b>2TB</b>	<b>1 – 10</b>
<b>DATASET #2</b>	<b>1</b>	<b>23,439,031</b>	<b>23,439,031</b>	<b>100KB – 2.5MB</b>	<b>3TB</b>	<b>3TB</b>	<b>2 – 3</b>

TABLE 1. Details of the two FlashBlade datasets used in testing

## DATASET CREATION ON FLASHBLADE

Creating a file system on FlashBlade is effortless. From the FlashBlade GUI, highlight the **Storage** option and then click on the **+** sign on the right.



FIGURE 2. FlashBlade GUI with file system creation highlighted

Next, give the file system a name, a provisioned size, and select the protocol(s) you wish to use. If using NFS, set the export rules. In our example, we used the NFS protocol exported to all VLANs and set the permissions to allow read and write privileges to all users. SMB and S3/Object Store are also fully supported protocols. Network export and permissions can be restricted as needed for your unique environment.

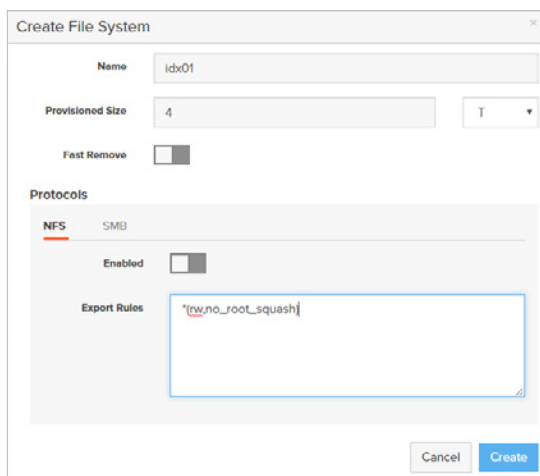


FIGURE 3. Simple NFS file system creation on FlashBlade

Once our file systems were created, we copied and extracted our source datasets to it.

## FLASHBLADE AND RUBRIK DATA PROTECTION SETUP FOR DATASET #1 AND DATASET #2

Despite the two datasets we introduced in the previous section being unlike one another, the steps needed within Rubrik to protect them are identical. How often a customer wants a backup job to run, the length of retention, along with optional archival/replication location, are variables that are dependent upon your unique environment requirements.

This section will illustrate how to set up a backup policy for both datasets described in the preceding section. The first step, creating the SLA policy for backing up the FlashBlade file systems, was easily accomplished from within the Rubrik GUI. Rubrik also has a full-featured REST API, so all the set up and functions shown here can also be done with the scripting language of your choice (e.g. Python, PowerShell).

Expand the **SLA Domains** menu and highlight the **Local Domains** option.

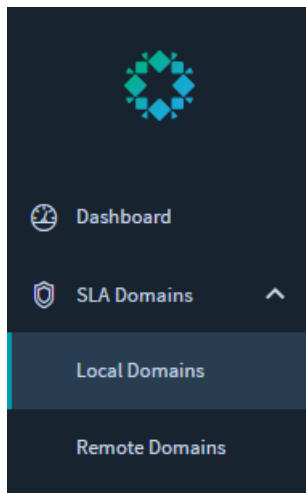


FIGURE 4. Rubrik Local Domains SLA creation

In the top-right of the GUI, click on the + sign that is spawned to get into the SLA domain creation wizard.



For both of our datasets, we will use identical but separate SLAs that will take snapshots every 12 hours with 5-day retention. Monthly backups will also be created and retained for two years. Time windows for when you want to run, or not run, the protection job can be assigned here, and you can also optionally set replication between multiple Rubrik appliances, or offload cold data to a separate archival FlashBlade for rapid onsite recovery or to a cloud-based tier such as AWS or Azure.

Create SLA Domain

SLA Domain Name  
FlashBlade-Dataset

Service Level Agreement  
Choose how often we take snapshots and the length of time we keep them.

Snapshot  
Take Snapshots: Every (Hours) 12 Keep Snapshots: For (Days) 5

Every (Days) For (Days)  
Every (Months) For (Years)  
1 2

Every (Years) For (Years)

Local retention set to 2 Years.

Snapshot Window  
Take snapshots from: to  
Take first full between: First Opportunity at

Cancel Configure Remote Settings Create

FIGURE 5. Options selected in creating our test SLA Domain

With our SLA domain built, the next step is to mount the four FlashBlade file systems from within the Rubrik appliance. In the main menu on the right, expand the **Protection** menu and highlight the **NAS Shares**\* option.

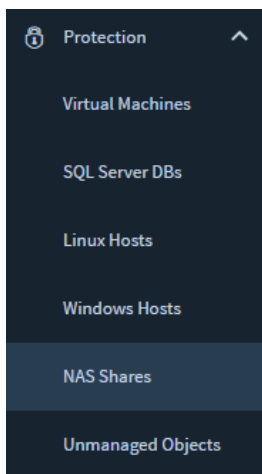


FIGURE 6. NAS Shares option under the Protection menu of the Rubrik GUI

Clicking on the **Add NAS Share** button in the top-right of the GUI will launch the wizard to mount FlashBlade to Rubrik.

\* Please note that in the addendum section we show how Linux hosts can be used as an intermediary mount point for FlashBlade to Rubrik, if desired.



FIGURE 7. Add Nas Share button in Rubrik GUI

Before we can access our test file sets, we first need to mount the NAS host. To do so, select the **Add Host** button.



FIGURE 8. Add an NAS host mount point in Rubrik GUI

Enter a data IP address for FlashBlade and the pureuser username and password combination.

A screenshot of the "Add NAS Host" form in the Rubrik GUI. The form has a title "Add NAS Host" and a subtitle "Please enter IP address or hostname, username and password to connect host." It contains four input fields: "IP or Hostname" with the value "10.21.108.15", "Domain (Optional)", "Username (Optional)" with the value "pureuser", and "Password (Optional)" with masked characters. There are "Cancel" and "Add" buttons at the bottom.

FIGURE 9. Include NAS hostname or IP address along with credentials

The host should now appear in the **Add Share** wizard. Click on the radio button for the FlashBlade IP or FQDN from the previous step, and click **Next**.



**Add Share**

Select Host Add Details

---

Search Host Add Host

Host	
<input type="radio"/>	10.21.107.189
<input type="radio"/>	10.21.108.12
<input checked="" type="radio"/>	10.21.108.15

Cancel Next

FIGURE 10. With the host added, select it and click the Next button

We will now enter in the first FlashBlade file system location as a share that Rubrik will be able to protect.

**Add Share**

Select Host Add Details

---

Specify share type, path and credentials to add a share

NFS Path (/mnt/exports)

/idx01

Share Type ☒ NFS ☐ SMB

Domain (Optional)

Username (Optional)

Password (Optional)

Cancel Back Finish

FIGURE 11. Specify the FlashBlade file system to add to the mount point as well as share type

While you can host multiple shares from the same NAS mount point, we elected to mount four separate NAS hosts for DataSet #1, each on a separate data IP address, to maximize throughput and resiliency. After creating the four NAS host mount points, we added one file system to each, as can be seen in the following NAS Share summary screen.









<input type="checkbox"/>	 /idx02	10.21.108.13	 Connected	NFS
<input type="checkbox"/>	 /idx04	10.21.108.14	 Connected	NFS
<input type="checkbox"/>	 /idx01	10.21.108.15	 Connected	NFS
<input type="checkbox"/>	 /idx03	10.21.108.15	 Connected	NFS

FIGURE 12. Four FlashBlade file systems mounted to Rubrik

With all four file systems mounted and available within Rubrik, we selected all four and clicked on the **Manage Protection** option in the top right.

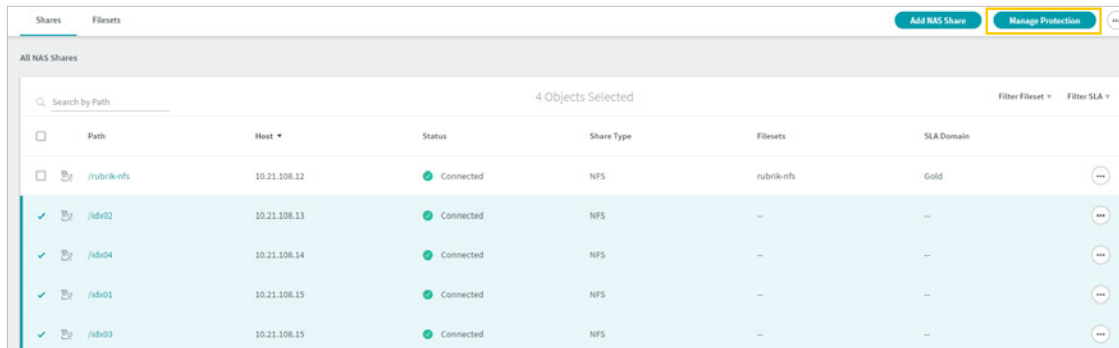
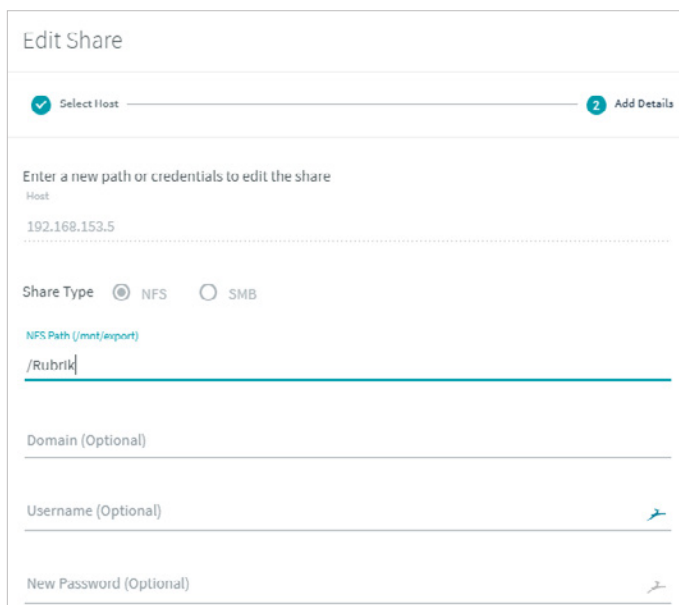


FIGURE 13. Select the four FlashBlade file systems and select Manage Protection

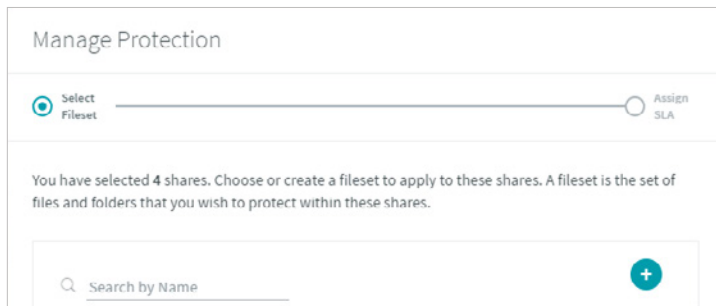
For DataSet#2, we used a single NAS host mount point since it was comprised of a single file system. In such a scenario, however, all four Rubrik nodes will ingest data from the single file system in a parallel manner.



The screenshot shows the 'Edit Share' form in the Rubrik GUI. The form has a progress bar at the top with 'Select Host' (completed) and 'Add Details' (current step). The 'Host' field contains '192.168.153.5'. The 'Share Type' is set to 'NFS'. The 'NFS Path (/mnt/export)' field contains '/Rubrik'. There are optional fields for 'Domain', 'Username', and 'New Password', each with a small icon to its right.

FIGURE 14. DataSet#2 mount within Rubrik GUI

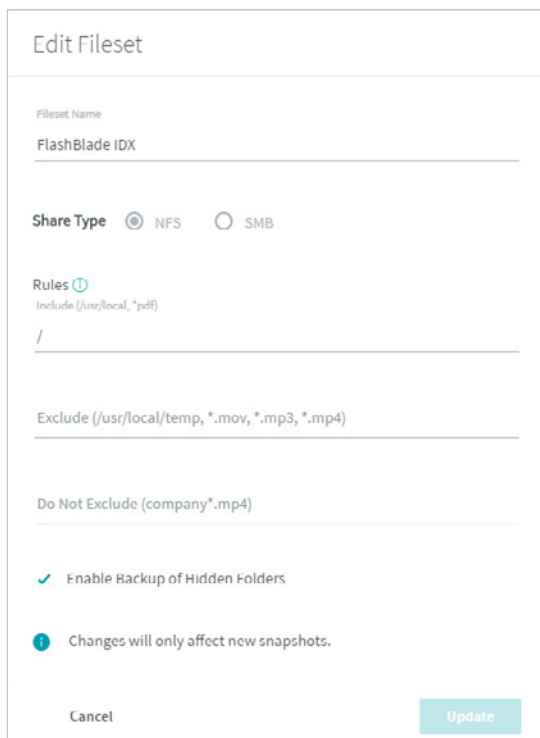
After that, we needed to create and select a fileset to apply to our shares. Filesets enable you to filter the types of files to include or exclude from your protection job. To create it, click on the + button in the window below.



The 'Manage Protection' window features a progress bar at the top with 'Select Fileset' (active) and 'Assign SLA'. Below, a message states: 'You have selected 4 shares. Choose or create a fileset to apply to these shares. A fileset is the set of files and folders that you wish to protect within these shares.' At the bottom, there is a search bar labeled 'Search by Name' and a green circular button with a white plus sign.

FIGURE 15. Click on the + sign to add a fileset

We are protecting all data on our FlashBlade file systems, so our fileset rules were very straightforward:



The 'Edit Fileset' window shows the following configuration: 'Fileset Name' is 'FlashBlade IDX'; 'Share Type' has 'NFS' selected; 'Rules' include 'Include (/usr/local, \*.pdf)' and 'Exclude (/usr/local/temp, \*.mov, \*.mp3, \*.mp4)'; 'Do Not Exclude (company\*.mp4)' is empty; 'Enable Backup of Hidden Folders' is checked; and a note states 'Changes will only affect new snapshots.' At the bottom are 'Cancel' and 'Update' buttons.

FIGURE 16. Provide a name for the fileset and assign any exclusion and/or inclusion rules

With our fileset rules established, we assigned them to our source FlashBlade file systems.

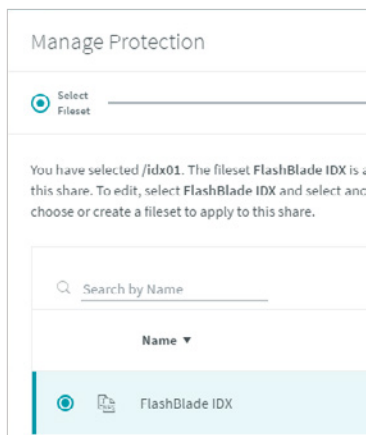


FIGURE 17. Select the newly created file set for use

The final step is to assign the SLA we built in the first step, which is illustrated in the screenshot below.

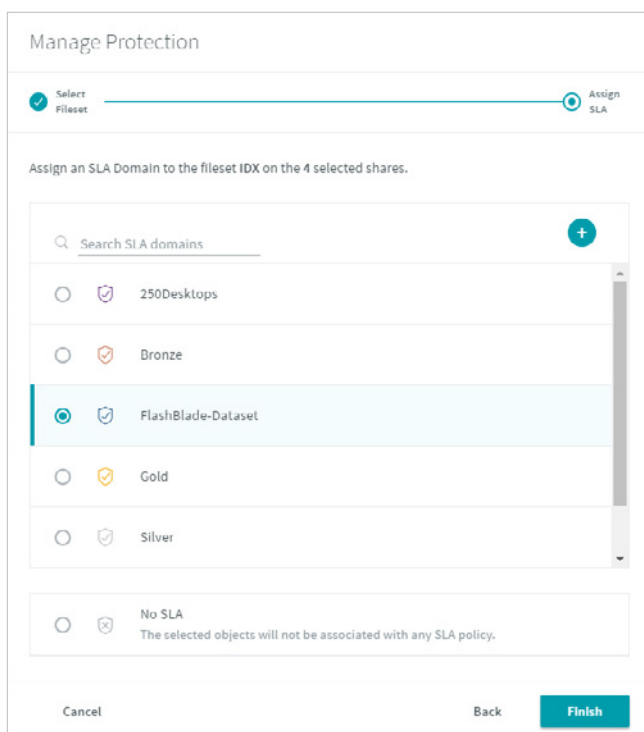


FIGURE 18. Assign the previously created Rubrik SLA to the four FlashBlade file systems

A summary screen shows that all four FlashBlade file systems have been successfully assigned the fileset and SLA Domain.

All NAS Shares						
Search by Path						Filter
<input type="checkbox"/>	Path	Host	Status	Share Type	Filesets	SLA Domain
<input type="checkbox"/>	/idx02	10.21.108.13	Connected	NFS	IDX	FlashBlade-Dataset
<input type="checkbox"/>	/idx04	10.21.108.14	Connected	NFS	IDX	FlashBlade-Dataset
<input type="checkbox"/>	/idx01	10.21.108.15	Connected	NFS	IDX	FlashBlade-Dataset
<input type="checkbox"/>	/idx03	10.21.108.15	Connected	NFS	IDX	FlashBlade-Dataset

FIGURE 18. SLA successfully assigned to the four FlashBlade file systems

## DATASET#1 BACKUP RESULTS

From the FlashBlade GUI, we observed a consistent high throughput of data ingest during the initial backup operation. Worth noting is that subsequent backup jobs to Rubrik will only include files that are changed or added, which will significantly decrease the amount of time and network bandwidth needed. As Rubrik features data compression technologies, the data footprint required will remain small over time. Snapshots on Rubrik are also immutable, allowing for recovery to any retained snapshot instantly.

A few moments after we assigned an SLA to DataSet#1, Rubrik begins to read and ingest data from our four FlashBlade file systems in parallel for content retention and protection.

Status	Name	Message	Date	
36%	FlashBlade IDX	Starting metadata scan for fileset "FlashBlade IDX" on "10.21.108.15/idx03"	5/18/17 10:31 AM	×
36%	FlashBlade IDX	Starting metadata scan for fileset "FlashBlade IDX" on "10.21.108.15/idx01"	5/18/17 10:31 AM	×
36%	FlashBlade IDX	Starting metadata scan for fileset "FlashBlade IDX" on "10.21.108.14/idx04"	5/18/17 10:31 AM	×
36%	FlashBlade IDX	Starting metadata scan for fileset "FlashBlade IDX" on "10.21.108.13/idx02"	5/18/17 10:30 AM	×

FIGURE 19. Data ingestion from FlashBlade to Rubrik starting

The following screenshot of the FlashBlade GUI shows that it took just under an hour to ingest the four file systems with over 2TB of data as an initial snapshot. Data is ingested and parity is built within the Rubrik cluster for further resiliency.

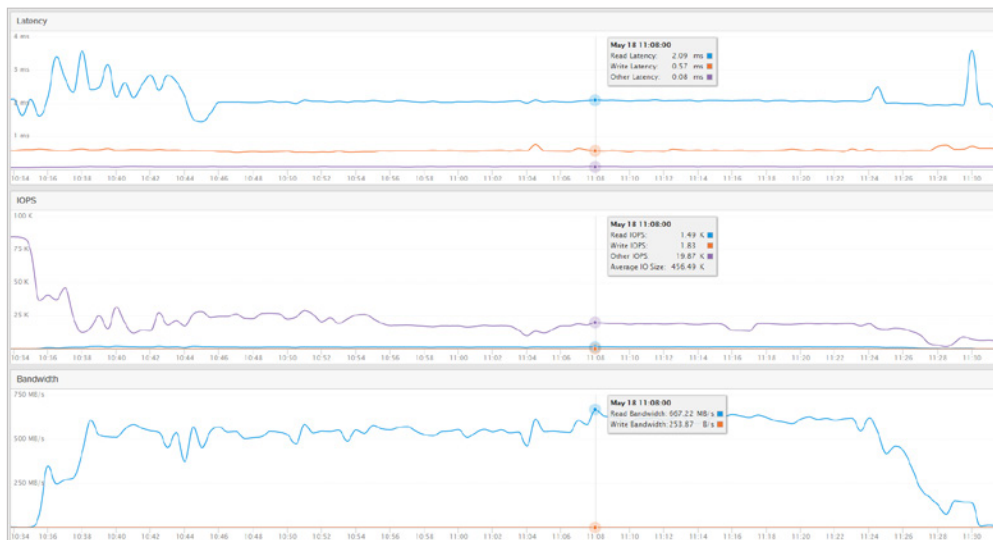


FIGURE 20. FlashBlade GUI showing data being read into Rubrik for protection

The example below of an incremental backup shows that subsequent file system snapshots are created in relatively little time with respect to the initial backup run:

Status	Activity	Date
✓	Queued backup of 'FlashBlade'	6/23 4:37 am
✓	Mounted NFS network share '10.21.108.12:/idx01' for fileset 'FlashBlade'	6/23 4:37 am
✓	Completed metadata scan for fileset 'FlashBlade' on '10.21.108.12:/idx01'. Fo...	6/23 4:45 am
✓	Completed data fetch for fileset 'FlashBlade' from '10.21.108.12:/idx01'. 0 byte...	6/23 4:45 am
✓	Unmounted NFS network share '10.21.108.12:/idx01' for fileset 'FlashBlade'	6/23 4:45 am
✓	Completed data copy for fileset 'FlashBlade' from '10.21.108.12:/idx01'	6/23 4:57 am
✓	Completed backup of Fileset 'FlashBlade' from '10.21.108.12:/idx01'	6/23 4:57 am

Figure 21. Incremental FlashBlade file system backup

DATASET#1 RESTORE RESULTS

The final test for DataSet#1 was to restore two different directories in parallel from Rubrik to FlashBlade from two separate file systems. One test directory included subdirectories of varying depth as well as many small files, while the second was a collection of a few very large files. To start, we selected one of our mounted NAS shares which spawned the window in Figure 23.

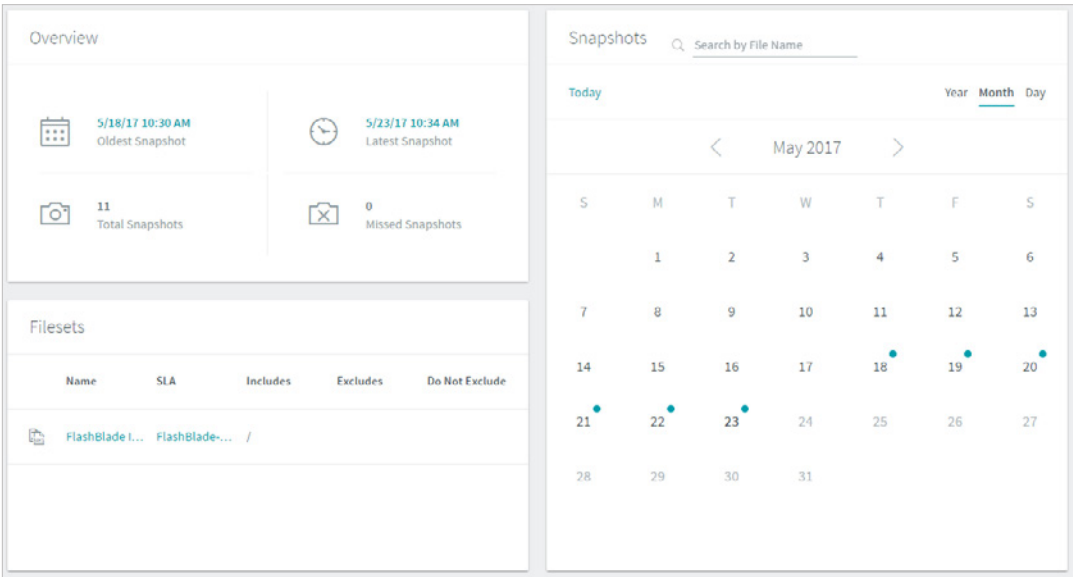


FIGURE 22. Single FlashBlade file system recovery pane in Rubrik GUI

Next, on the calendar, click on the date that you wish to restore back to.

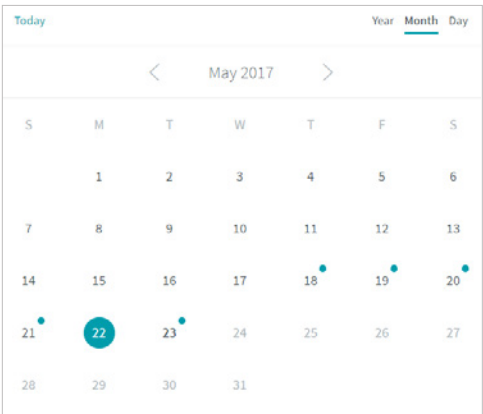


FIGURE 23. Select a snapshot to recover to from the calendar view within Rubrik GUI

Snapshots taken for the date selected will be displayed. Click on the icon to the right and select **Browse Files** to drill down to the directory or file that needs to be recovered. Individual files or folders can also be searched for and recovered.

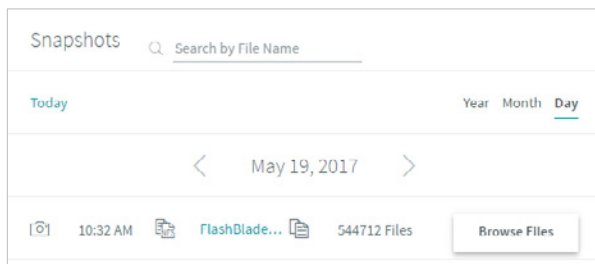


FIGURE 24. Select date, time, and files/folders to be restored

In this case, we are going to restore the entire LoginVSI-2 directory, which is comprised of over 25,000 unique files, and is about 20GB in size. Navigate to the directory and click on the circle icon alongside of it to show the available recovery options. We are choosing the **Restore** option in this test.

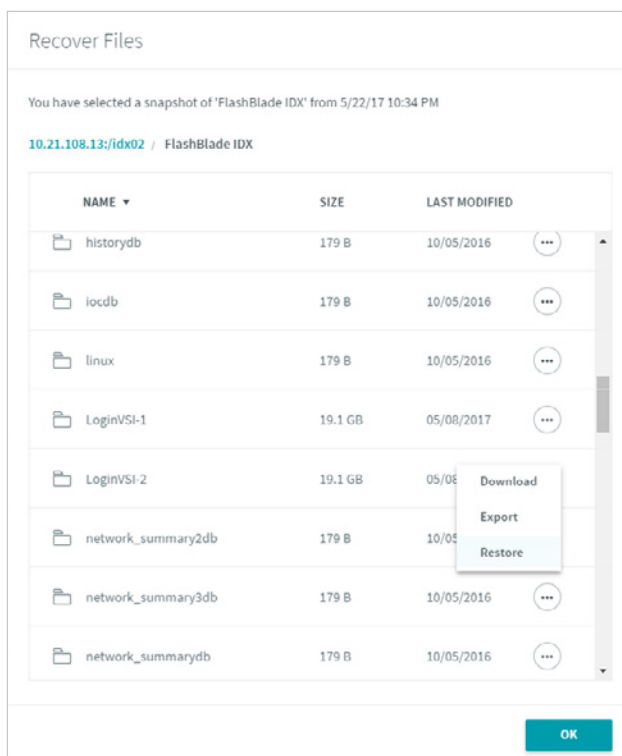


FIGURE 25. Once the directory is located, click on the circular icon and select Restore

The recovered directory will be entirely overwritten on our FlashBlade file system for this test, but customers can also choose to create a separate directory to restore to.

Clicking the **Restore** button will kick off the operation.



Restore

You have selected to restore **/LoginVSI-2** from a snapshot from 5/22/17 10:34 PM.

☒ Overwrite Original
   
☐ Restore to Separate Folder

Folder Name

☐ Ignore restore errors

Cancel
Restore

FIGURE 26. Choose to overwrite the original files or restore to a separate directory, then click Restore

Immediately after this restore operation was created, we repeated the process for a larger directory that was approximately 96GB in size and comprised of fewer and bigger files.

Recover Files

You have selected a snapshot of 'FlashBlade IDX' from 5/22/17 10:34 AM

10.21.108.15/idx03 / FlashBlade IDX

NAME ▼	SIZE	LAST MODIFIED	
vdb.1_2.dir	95.8 GB	05/17/2017	⋮
vdb.1_3.dir	95.5 GB	05/17/2017	⋮
vdb.1_4.dir	95.6 GB	05/17/2017	⋮
vdb.1_5.dir	95.5 GB	05/17/2017	⋮
vdb.1_6.dir	95.9 GB	05/17/2017	⋮
vdb.1_7.dir	96 GB	05/17/2017	⋮
vdb.1_8.dir	95.8 GB	05/17/2017	⋮

Download
  
Export
  
Restore

OK

FIGURE 27. Repeating directory recovery on a larger directory with fewer files

The following two screenshots show our two simultaneous recovery tasks being executed. The smaller-sized recovery with the larger number of files completed in just about 14 minutes, while the second job with larger files finished in under an hour, proving that rapid recovery of very different sets of data is smoothly achieved, even in parallel.

Activity Detail		
Status	Activity	Date
✓	Mounted destination NFS network share '10.21.108.13:/idx02'	5/24 6:55 am
✓	Unmounted destination NFS network share '10.21.108.13:/idx02'	5/24 7:06 am
✓	Restore of '/LoginVSI-2' to '10.21.108.13:/idx02' based on snapshot taken at 'T... 5/24 7:06 am	
<div>CancelDownload LogsOK</div>		

FIGURE 28. Activity details for our first directory recovery.

Activity Detail		
Status	Activity	Date
✓	Mounted destination NFS network share '10.21.108.15:/idx03'	5/24 6:56 am
✓	Unmounted destination NFS network share '10.21.108.15:/idx03'	5/24 7:44 am
✓	Restore of '/vdb.1_7.dir' to '10.21.108.15:/idx03' based on snapshot taken at '... 5/24 7:44 am	
<div>CancelDownload LogsOK</div>		

FIGURE 29. Activity details for our second directory recovery.

**DATASET#2 BACKUP RESULTS**

This section covers results for the FlashBlade dataset that consisted of over 23 million small files hosted on a single FlashBlade file system being backed up and restored via Rubrik. The SLA was identical to that used in the previous dataset and this testing was performed separately to demonstrate expected results for this use case. Additionally, a single NAS mount point within Rubrik FlashBlade was used in this experiment.

After the SLA was assigned to our source data, we can see that Rubrik begins with an initial metadata scan of the data as pictured below from the FlashBlade GUI (note the purple IOPs line in the middle graph):

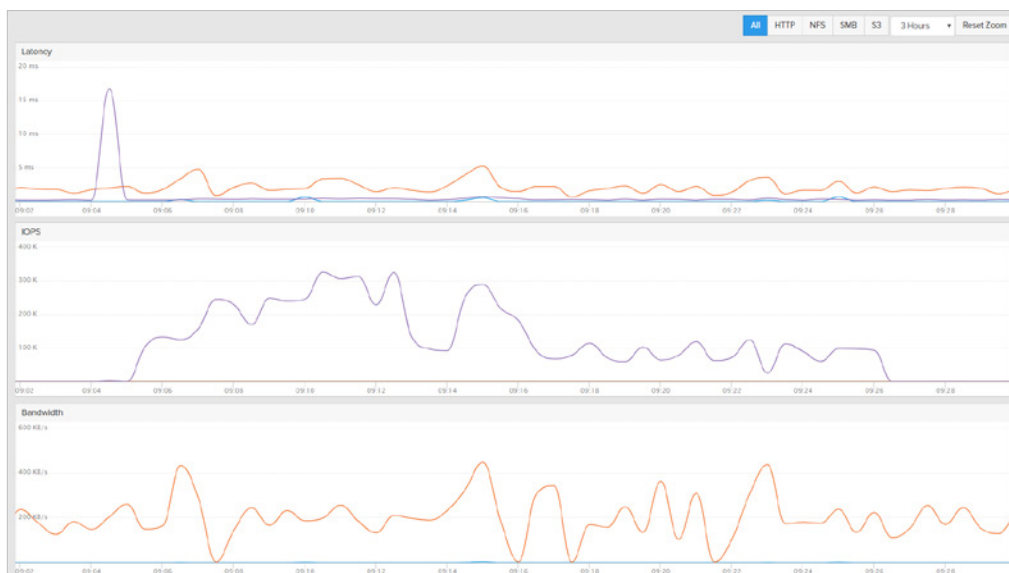


FIGURE 30. Initial metadata ingest shown in FlashBlade GUI

Once the initial metadata scan has completed, data was copied from FlashBlade to Rubrik across all four Rubrik nodes. Impressively, the entire dataset was copied to Rubrik in about 90 minutes, at a rate of over **11,000 files/second**. After this initial ingest is complete, data parity of the source data is built amongst the Rubrik nodes for further resiliency.

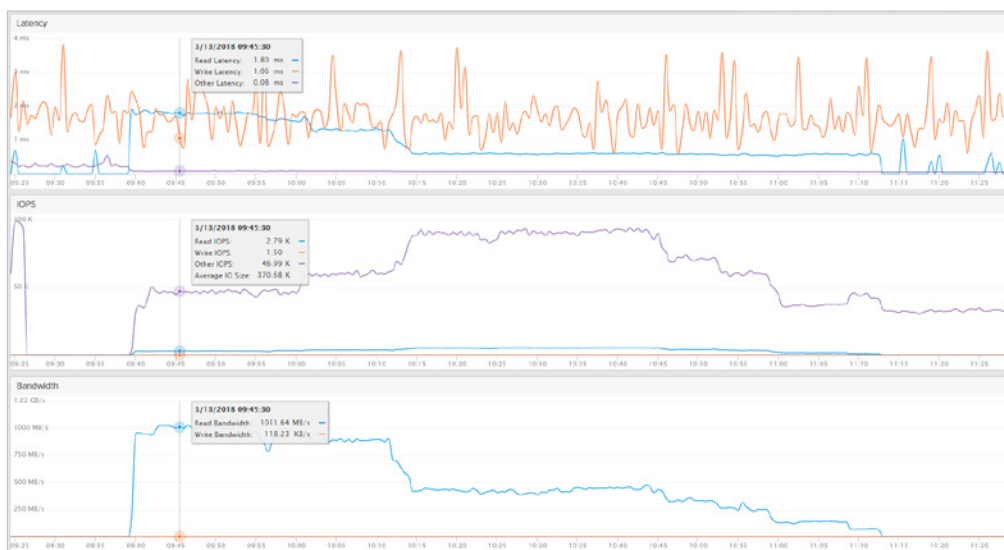


FIGURE 31. 23 million file copy from FlashBlade to Rubrik

The below summary screen from the Rubrik GUI shows us that initial metadata scan, data ingest, and backup verification on Rubrik was completed in under 24 hours.

Activity Detail		
Status	Activity	Date
✓	Queued backup of 'ImageNet'	3/13 9:04 am
✓	Completed metadata scan for fileset 'ImageNet' on '192.168.153.5/ [redacted]' using 35 partitions. Found 23439030 files, 900550 directories and 0 links. Scanned at a rate of 11698 files per second. Failed to scan 0 inodes	3/13 9:39 am
✓	Backed up data for fileset 'ImageNet' from '192.168.153.5/ [redacted]'. Successfully backed up 23439030 files consisting of 0 bytes in 10082 seconds at an average transfer rate of 0 Dps MBps. Failed to back up 0 files	3/13 3:04 pm
✓	Patch file conversion completed for Fileset 'ImageNet'. Converted 3267029895205 bytes with throughput 51001405 bytes per second	3/14 8:52 am
✓	Finished verification for fileset 'ImageNet'	3/14 9:11 am
✓	Completed data copy for fileset 'ImageNet' from '192.168.153.5/ [redacted]' in 65190 seconds	3/14 9:11 am
✓	Completed backup of Fileset 'ImageNet' from '192.168.153.5/ [redacted]'	3/14 9:11 am

FIGURE 32. Rubrik initial ingest job summary

We ran an additional incremental backup after executing a permissions change (**chmod 777**) against all 23 million files. This time, the entire incremental run completed in about 75 minutes.

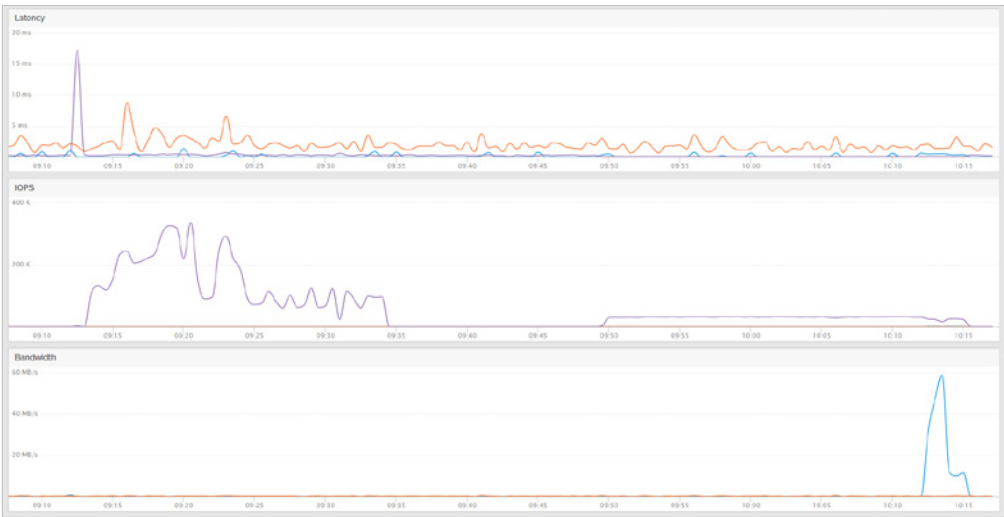


Figure 33. FlashBlade GUI during incremental backup job

Activity Detail		
Status	Activity	Date
✓	Queued backup of 'ImageNet'	3/14 9:12 pm
✓	Completed metadata scan for fileset 'ImageNet' on '192.168.153.5/Mentor' using 34 partitions. Found 23439029 files, 900550 directories and 0 links. Scanned at a rate of 3/14 9:48 pm 11752 files per second. Failed to scan 0 inodes	
✓	Backed up data for fileset 'ImageNet' from '192.168.153.5/Mentor'. Successfully backed up 23439029 files consisting of 0 bytes in 166 seconds at an average transfer rate of 0 Bps MBps. Failed to back up 0 files	3/14 9:51 pm
✓	Patch file conversion completed for Fileset 'ImageNet'. Converted 156706 bytes with throughput 148 bytes per second	3/14 10:09 pm
✓	Finished verification for fileset 'ImageNet'	3/14 10:27 pm
✓	Completed data copy for fileset 'ImageNet' from '192.168.153.5/Mentor' in 2170 seconds	3/14 10:27 pm
✓	Completed backup of Fileset 'ImageNet' from '192.168.153.5/Mentor'	3/14 10:27 pm

FIGURE 34. Rubrik summary of incremental backup job

## DATASET#2 RESTORE RESULTS

With our data successfully backed up, the next phase in testing was to delete a large amount from FlashBlade and recover it back to the original location through Rubrik.

From a Linux host we mounted the FlashBlade file system, went into a directory that contained over 14 million files (1.3TB) and deleted it:

```
[root@ Rubrik]# cd train/
[root@ train]# find . -type f | wc -l
14197087
[root@ Rubrik]# rm -rf ./train/*
[root@ train]# find . -type f | wc -l
0
```

To start the recovery process, we went into the Rubrik GUI and found the most recent backup of our entire FlashBlade file system:

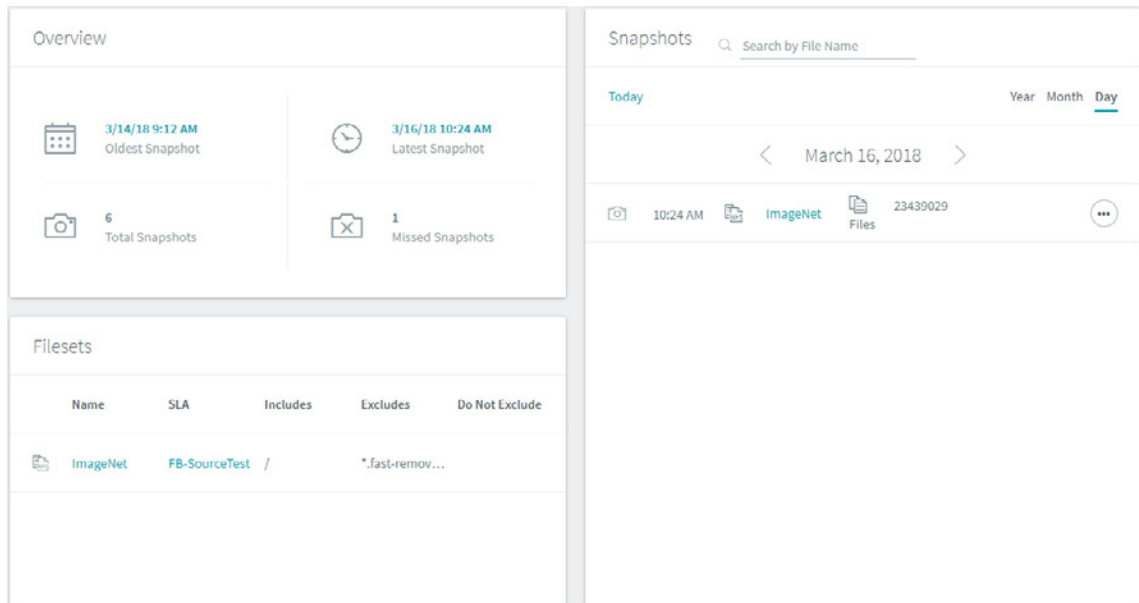


Figure 35. Rubrik GUI for snapshot management and recovery

Clicking into the snapshot enabled us to browse into the files and directories it contained:

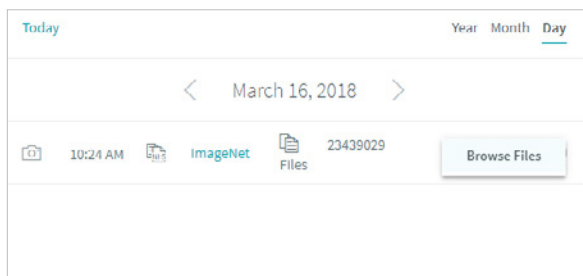


FIGURE 36. Navigating to directory for recovery in Rubrik GUI

Next, we navigated to the deleted directory we wanted, clicked on the button and selected the **Restore** option:

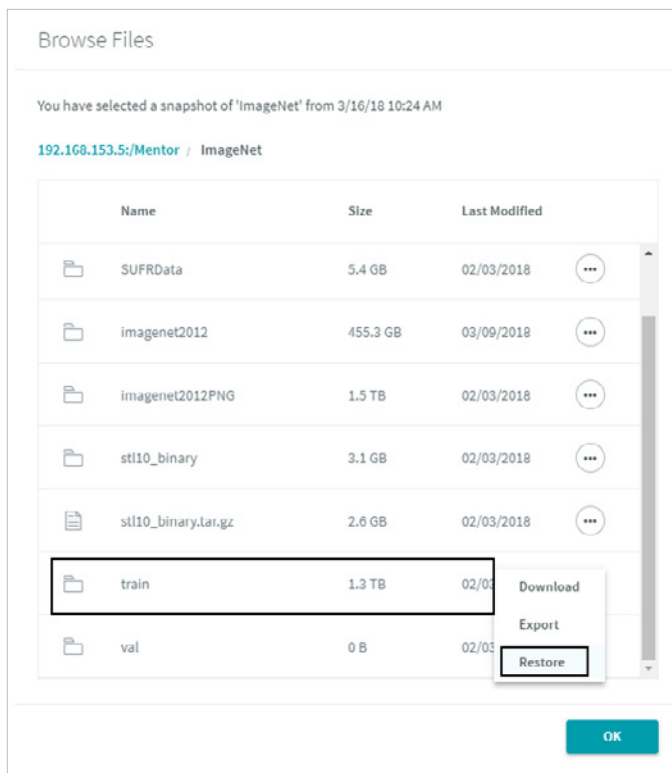


FIGURE 37. Selecting deleted directory for recovery

There are several restore options, and we elected to restore this directory back to its original location:

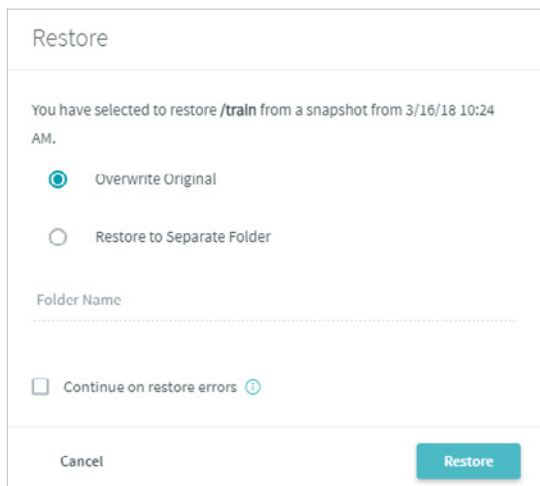


Figure 38. Available Rubrik recovery options

A single Rubrik node handled this recovery job so that the other cluster nodes would be free to handle other active backup and restore jobs in parallel. We found that the entire dataset of over 14 million files and 1.3TB of data was rebuilt on FlashBlade in less than 7 hours. Of course, smaller restore jobs would finish in a fraction of the time.

Activity Detail		
Status	Activity	Date
✓	admin started a restore job for fileset 'ImageNet' on NFS host '192.168.153.5' from '/train' to '/' using snapshot 011312/f-da94-48c9-a310-b06d6892c/d2 taken on Fri Mar 16 17:24:41 UTC 2018	3/16 4:52 pm

Activity Detail		
Status	Activity	Date
✓	Completed writing data to '192.168.153.5:/[REDACTED]' from node 'RVM15C5009529' for fileset 'ImageNet' in 23973 seconds. Restored at a rate of 592 files per second and 54.18 MB/s.	3/16 11:31 pm
✓	Successful restore of '/train' to '192.168.153.5:/[REDACTED]' based on snapshot taken at 'Fri Mar 16 17:24:41 UTC 2018'. Successfully restored 1.3 TB of data in 14197087 files and 21842 directories. Failed to restore 0 files and 0 directories.	3/16 11:31 pm

Figure 39. Rubrik GUI summarizing recovery operation

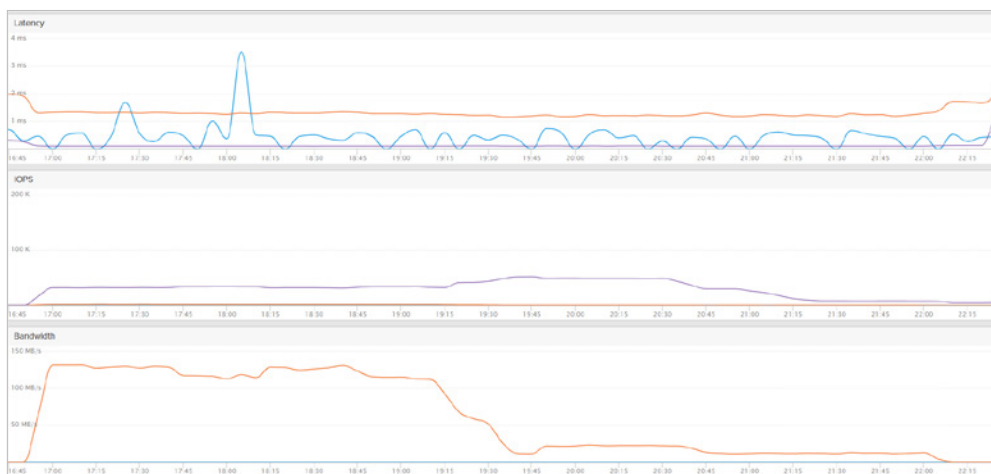


Figure 40. FlashBlade GUI showing recovered data being written to the original file system



## CONCLUSION

We have clearly demonstrated that the paradigm of backup and recovery has dramatically shifted – from what has traditionally been a pain point in both large and small enterprises to an easy to setup and manage, holistic solution. FlashBlade provides the density, scalability, and unparalleled performance to handle your most demanding unstructured workloads, while Rubrik allows for exceedingly simple and smooth recovery of entire directories, all the way down to individual files.

This partnership and integration allows our joint customers to focus on improving their business processes rather than struggling daily to keep their proprietary data available to the right individuals and rapidly recoverable in the event of disaster, ransomware, or accidental keystroke. This kind of solution provides just what the IT industry has been asking for: consolidation, protection, and ease of management for exponentially increasing datasets, in denser footprints, with access to deeper data analytics for the new stack.

## ADDENDUM: USING A LINUX HOST

Using a Linux host as an intermediary for backing up a FlashBlade NFS share is an alternative option provided by Rubrik. Because the Linux host is a single point of failure, wherever possible this method is not recommended over the native NAS option shown earlier in this guide. However, there certainly are use cases where NFS volumes are mounted on a Linux host and the ability to provide data protection at that layer will be an attractive functionality for some users.

An example of how to mount a FlashBlade NFS file system to a linux host (CentOS7 in this case) is:

### Syntax:

```
mkdir [local Linux Host mount directory]
mount -t nfs -o nolock,hard,intr,nfsvers=3,tcp,bg [data IP for FlashBlade]:/[FlashBlade_FileSystem] [Linux Host mount directory]
```

### Examples:

```
[root@lab-fb-b01-11-01 ~]# mkdir /mnt/idx01
[root@lab-fb-b01-11-01 ~]# mount -t nfs -o nolock,hard,intr,nfsvers=3,tcp,bg 10.21.108.15:/idx01 /mnt/idx01
```

With NFS mounted on our Linux host, we can now use Rubrik to protect it. To start, click on the **Linux Hosts** option on the main **Rubrik Protection** menu.

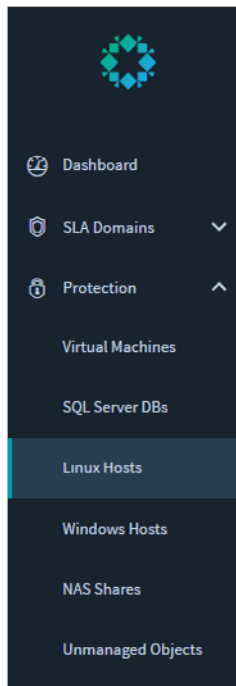


FIGURE 41. Select Linux Hosts under Protection Menu

Next, click on the **Add Linux Hosts** button on the top right.

**Add Linux Hosts**

FIGURE 42. Click Add Linux Hosts button

The Rubrik Backup Service agent will need to be downloaded and installed to the Linux host. Supported Linux distributions as of the publishing of this paper are shown in the screenshot below. In our example, we are using CentOS7.

### Add Linux Hosts

The **Rubrik Backup Service** must be installed on the Linux hosts. Please enter comma separated IP addresses or hostnames to connect hosts. Compatible with: RHEL 5,6,7, CentOS 5,6,7, and Oracle Linux 5,6,7

FIGURE 43. Download the Rubrik Backup Service binary to the Linux host

Installing the Rubrik backup package is achieved through a simple **rpm -Uvh rubrik-agent.x86\_64.rpm** command. Expected output is shown below:

```
[root@lab-fb-b01-11-01 ~]# pwd
/root
[root@lab-fb-b01-11-01 ~]# ls *.rpm
rubrik-agent.x86_64.rpm
[root@lab-fb-b01-11-01 ~]# rpm -Uvh rubrik-agent.x86_64.rpm
Preparing...                               [100%]
Updating / installing...
 1:rubrik-agent-3.1.0.0A1-1.0               [100%]
Generate uuid
Generate credentials
/etc/rubrik
Generating RSA private key, 2048 bit long modulus
.....++++
e is 65537 (0x10001)
Using configuration from /tmp/tmp.cCbBIKQZR/conf
Check that the request matches the signature
Signature ok
The Subject's Distinguished Name is as follows
commonName      :ASN.1 12:'076662c5-09a2-4dcc-bfc3-ed00076a6476.agent.rubrik.local'
countryName     :PRINTABLE:'US'
stateOrProvinceName :ASN.1 12:'California'
localityName    :ASN.1 12:'Palo Alto'
organizationName :ASN.1 12:'Rubrik, Inc.'
Certificate is to be certified until Oct  3 16:58:43 2044 GMT (10000 days)

Write out database with 1 new entries
Data Base Updated
Status daemon: bootstrap_agent_main [DOWN]
Starting daemon: bootstrap_agent_main
[0518 09:58:43.215813 36172 common.cpp:23] Binary built on Feb 22 2017 at 23:27:11
[ OK ]
Status daemon: bootstrap_agent_main [UP]
Status daemon: backup_agent_main [DOWN]
Starting daemon: backup_agent_main
[0518 09:58:43.333513 36205 common.cpp:23] Binary built on Feb 22 2017 at 23:27:11
[ OK ]
Status daemon: backup_agent_main [UP]
[root@lab-fb-b01-11-01 ~]#
```

FIGURE 44. Install the Rubrik Backup Service to the [Linux host](#)

Now that the Backup Service is installed and running, adding the Linux host can be accomplished by adding the IP address or hostname:

Add Linux Hosts

The **Rubrik Backup Service** must be installed on the Linux hosts. Please enter comma separated IP addresses or hostnames to connect hosts. Compatible with: RHEL 5,6,7, CentOS 5,6,7, and Oracle Linux 5,6,7

IPs or Hostnames (Use commas to separate)

10.21.107.183

Cancel

Add

FIGURE 45. Enter the IP address or hostname of the Linux Host within the Rubrik GUI

With the host added, we are now able to add a fileset including our mounted NFS source from FlashBlade and assign the same SLA as used in the earlier section to back up the file system.

All Linux Hosts					
Search by Name					
	Name	Fileset	SLA Domain	Status	OS Type
<input type="checkbox"/>	10.21.107.183	--	--	<input checked="" type="checkbox"/> Connected	Linux

FIGURE 46. Adding a fileset and SLA domain to the Linux host

© 2018 Pure Storage, Inc. All rights reserved.

Pure Storage, the “P” Logo, and FlashBlade are trademarks or registered trademarks of Pure Storage, Inc. in the U.S. and other countries. Rubrik is a registered trademark of Rubrik in the U.S. and other countries.

The Pure Storage product described in this documentation is distributed under a license agreement and may be used only in accordance with the terms of the agreement. The license agreement restricts its use, copying, distribution, decompilation, and reverse engineering. No part of this documentation may be reproduced in any form by any means without prior written authorization from Pure Storage, Inc. and its licensors, if any.

THE DOCUMENTATION IS PROVIDED “AS IS” AND ALL EXPRESS OR IMPLIED CONDITIONS, REPRESENTATIONS AND WARRANTIES, INCLUDING ANY IMPLIED WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, OR NON-INFRINGEMENT ARE DISCLAIMED, EXCEPT TO THE EXTENT THAT SUCH DISCLAIMERS ARE HELD TO BE LEGALLY INVALID. PURE STORAGE SHALL NOT BE LIABLE FOR INCIDENTAL OR CONSEQUENTIAL DAMAGES IN CONNECTION WITH THE FURNISHING, PERFORMANCE, OR USE OF THIS DOCUMENTATION. THE INFORMATION CONTAINED IN THIS DOCUMENTATION IS SUBJECT TO CHANGE WITHOUT NOTICE.

ps-wp28p-flashblade-data-protection-with-rubrik-02