COHESITY

RSYNC Backup and Recovery - Cohesity Remote Adapter

Solution guide to back up any unix based operating system using rsync and Cohesity remote adapter

Version 0.1

June 2022

COHESITY

Table Of Contents

COHESITY

In Scope

- Shell script
- First full Backup (Initial Sync) of unix filesystems to Cohesity NFS
- Incremental forever (Changed Size and Last Modified time) of unix filesystem to Cohesity NFS live view
- Preserve ACL's, ownership, extended attributes
- Preserve mtime, ctime, atime
- Exclusions and Inclusion based on pattern/input-file
- Automation using RA
- Indexing backup for easy search option
- Download file/folder option from UI
- Clone view and mount as NFS workflow to recover back to source
- Capture and Display errors on UI

Out of Scope

- Full server recovery or BMR
- Performance aspects, if any (as we are dependent on rsync)
- Hard Links handling

Script Location

The script can be downloaded from the following github link:

https://github.com/manojmittalcohesity/cohesitysolutionsengineering/blob/main/rsyncbackup.sh

COHESITY

Pre Requisites

- Any Unix based operating system
- Commands/utilities required
    - rsync version 3.0.8 or later
    - sudo
    - mount
    - umount
    - Bourne shell (default shell)
- Setup passwordless ssh access from Cohesity to the unix host
- Ensure the user being used for password less access has sudo access to execute following commands:
    - rsync
    - mount
    - Umount

Sample sudoers file content below:

```
bash-5.0# cat /usr/local/etc/sudoers
mittal  ALL=(ALL) NOPASSWD: /usr/bin/rsync, /usr/sbin/mount, /usr/sbin/umount
manoj   ALL=(ALL) NOPASSWD: ALL
bash-5.0#
```

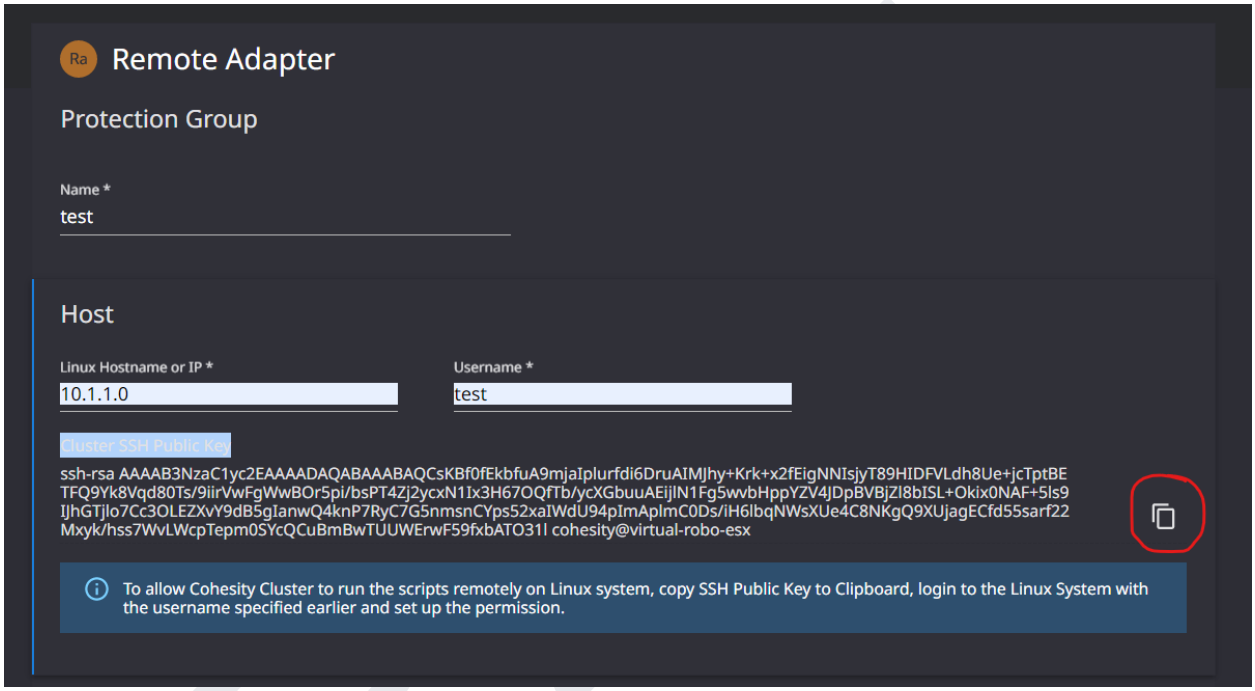
- PATH variable should be updated in script. This is needed so that script can locate and use required commands and utilities. Currently the script is hardcoded to use the following PATH variable
    - /usr/sbin:/usr/bin:/usr/local/bin
- Create an NFS view on Cohesity for each host that needs to be protected.
- Create a mount path on each host which would be used to mount cohesity view. The script takes care of automatically mounting and unmounting cohesity as NFS to this path. The path needs to be provided as a parameter to remote adapter job
- Follow the following KB to ensure cohesity views can be mounted as NFS on HP-UX and Solaris hosts

https://support.cohesity.com/s/article/HP-UX-v-11-31-NFS-UDP-initialization

COHESITY

Setup Passwordless ssh access from Cohesity to Unix host

- Log in to Cohesity UI and create a remote adapter group. Provide group a name, unix hostname, username



- Copy the cluster ssh public key as shown in the screenshot above
- Log into the unix host using the username with which remote adapter group is configured
- Update the users authorized_keys file (under ~/.ssh/ directory) with the cluster ssh public key that was copied from cohesity UI
- Change the permissions of the file to minimum (400) and save the file
- This completes passwordless access from Cohesity to the unix host

RSYNC options used by the script

The script uses the various rsync command options to do its job. Following is an explanation of each of these options used by RSYNC command in the script

-a: Archive mode. In this mode, script takes care of backing up files recursively, preserves permissions, symlinks, modification times, ACL's, device and special files

--stats: This tells rsync to print a verbose set of statistics on the file transfer, allowing you to tell how effective rsync's delta-transfer algorithm is for your data

--human-readable: Output numbers in a more human-readable format

--out-format: output updates using the specified FORMAT

--delete: Delete extraneous files from destination directory. This option ensures that if any directory is deleted on source, it is also deleted on the live view

--relative: Use relative paths. This indicates the full path names specified on the command line are sent to the server rather than only the last parts of the file names.

--no-whole-file: This option ensures that in case a large file is modified, only changed blocks are sent rather than the whole file improving the efficiency of transfer

--exclude: This is a hardcoded path in the script which ensures that the rsync destination directory is always excluded

--exclude-from: This is a path to the text file containing list of files and directories that needs to be excluded or included. More details on how to create this file here

Using Exclusions and Inclusions

- Create a file on the host being protected, name it anything you like. An example would be /home/cohesityrsyncexclusioninclusion
- Edit the file and add contents as per the examples given below as per your requirements

- The script automatically polls any NFS filesystems mounted during runtime and excludes them. There is no need to specify them in the user specified exclude file

Example 1: Exclude a complete directory

/home/xyz/d/ ---------- Excludes "d" directory under /home/xyz/

Example 2: Exclude all hidden files in a directory

/home/xyz/.* ---------- Excludes all hidden files under /home/mittal/

Example 3: Exclude all hidden files in a directory but backup one named hidden file

+ /home/xyz/.cshrc —---------- This will include ".cshrc" file under /home/xyz/ directory
- /home/xyz/.* ---------- This will exclude all hidden files under xyz but backup ".cshrc"

More example on how to to use exclusions and inclusions can be found at below link:

https://sites.google.com/site/rsync2u/home/rsync-tutorial/the-exclude-from-option

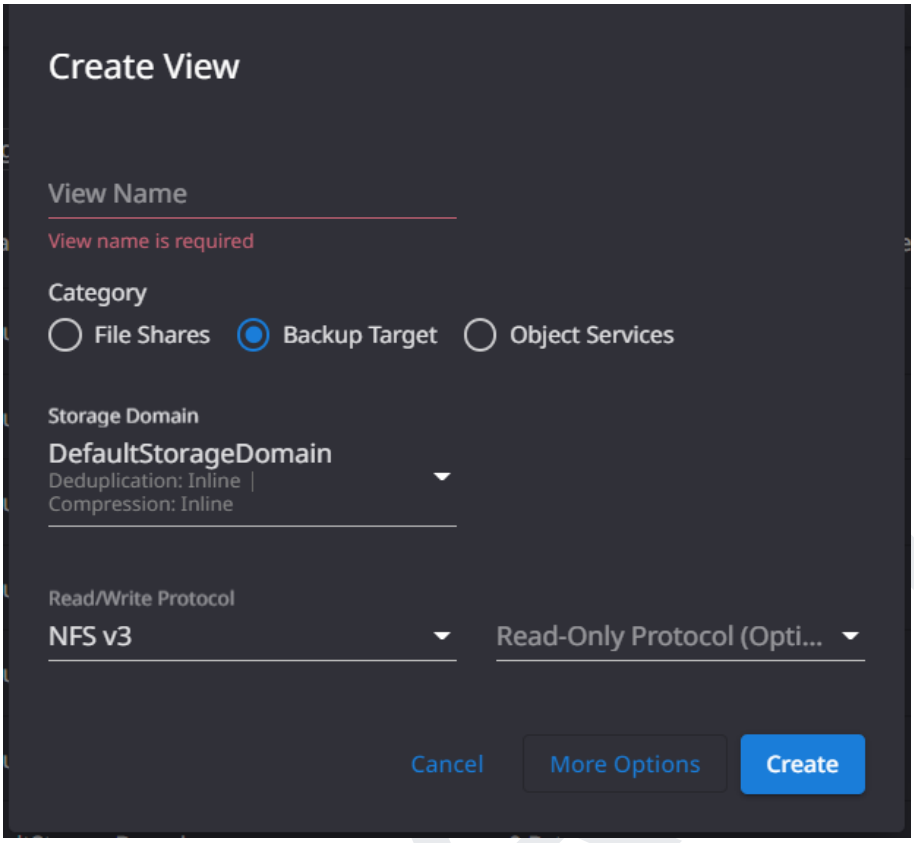
Script Parameters

This sections explains various parameters that needs to be passed on to the script through Cohesity's remote adapter protection group that runs the script

-V : Name of the NFS view created on Cohesity that should be mounted on the unix host
-P : Path on the unix host where the script would attempt to mount the Cohesity view
-I : IP or hostname of the cohesity cluster node which would be used to mount the view
-S : Source path that needs to be backed up (Ex: / or /home)
-E : Path to the exclude/include file. For more details, refer the section on exclusions and inclusions

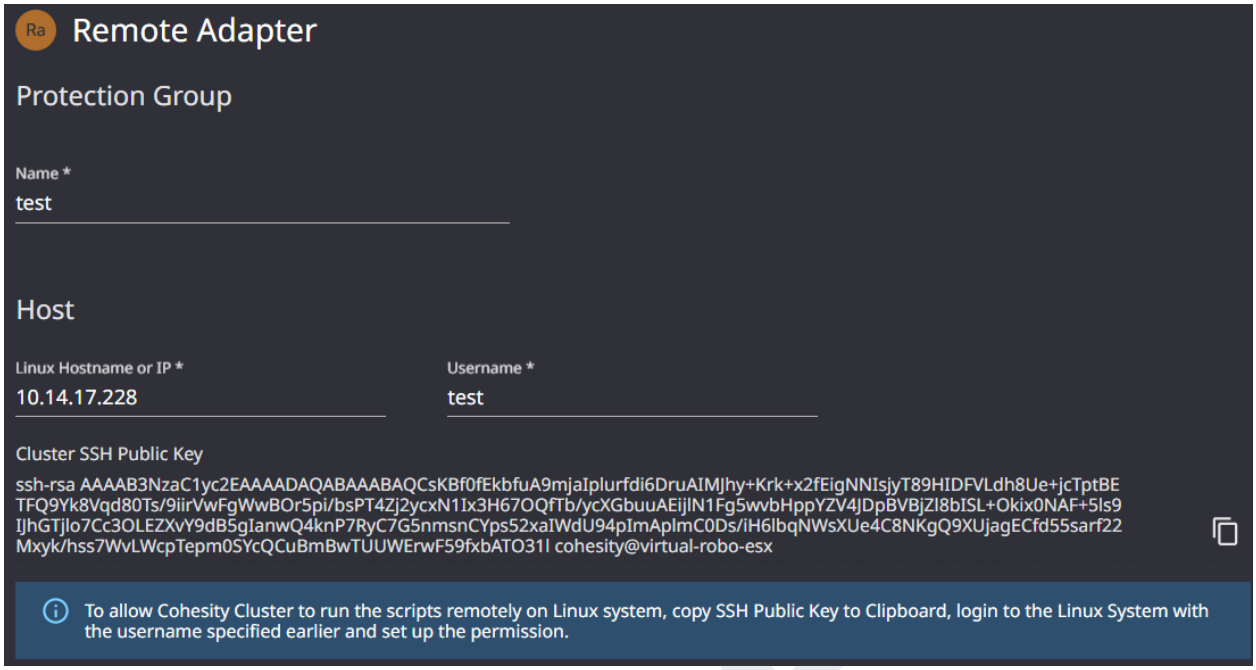
Steps

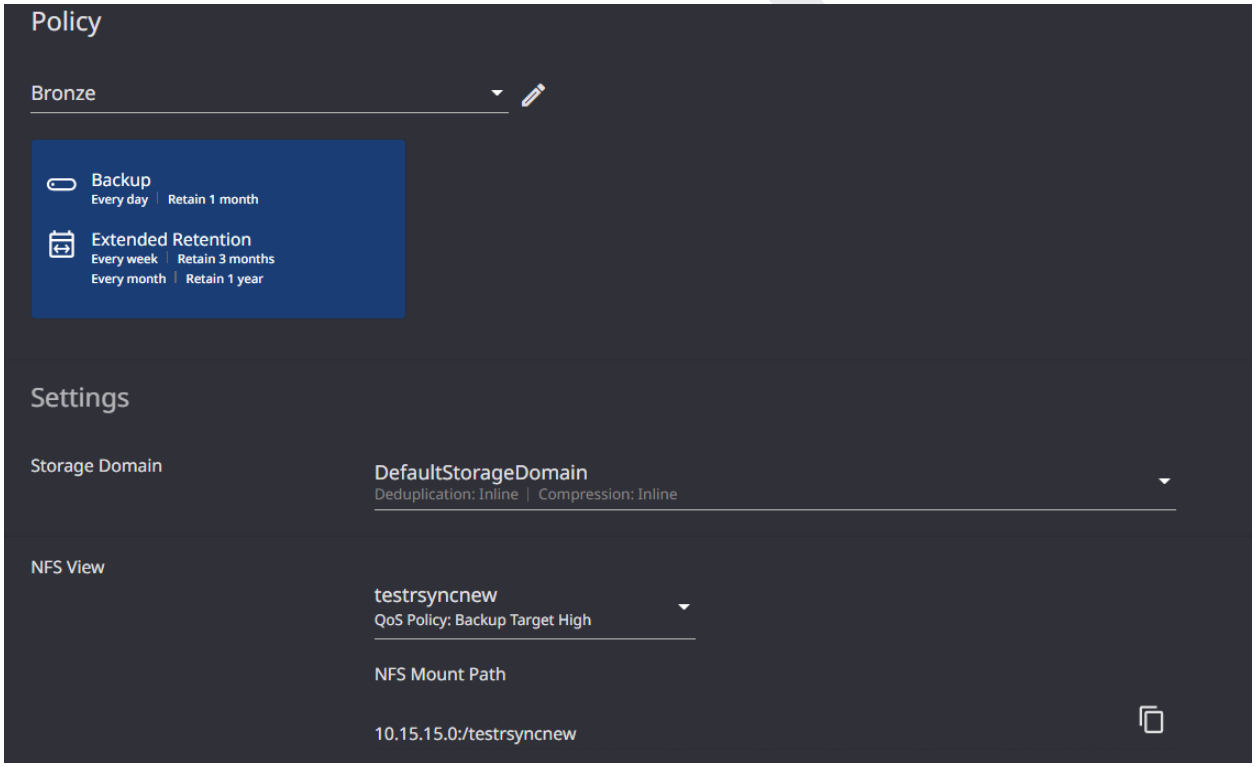
- Create one NFS view per unix host that needs to be protected on Cohesity

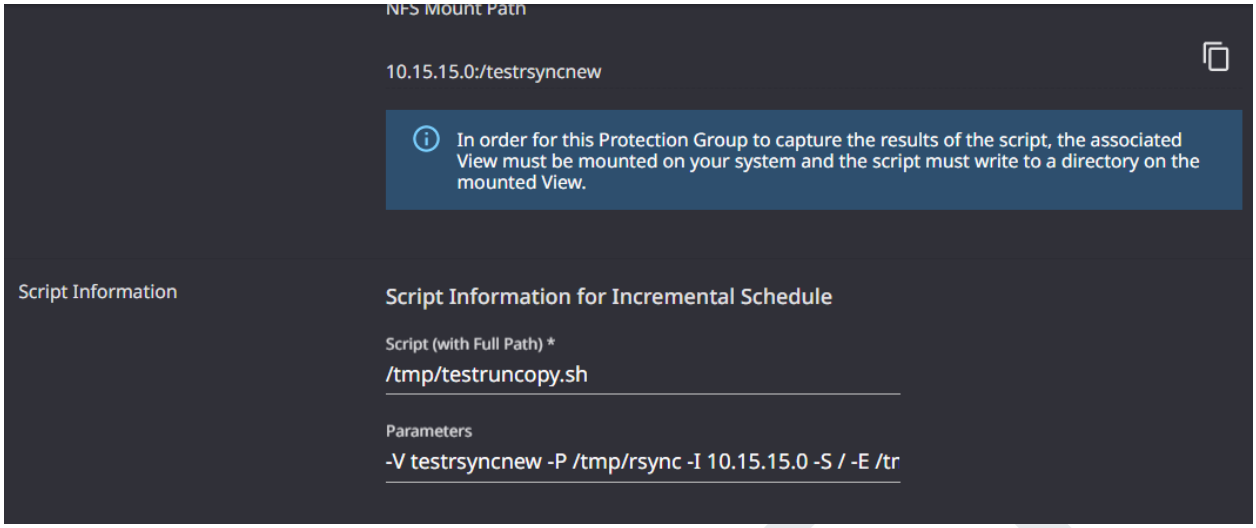


- Copy the rsync script to the unix system and ensure it has adequate privileges. The user which is configured on Cohesity Remote adapter job should have execute privilege to the script
- Identify a user which has super user privilege on the system to backup all the files
- Create a remote adapter job on Cohesity

COHESITY

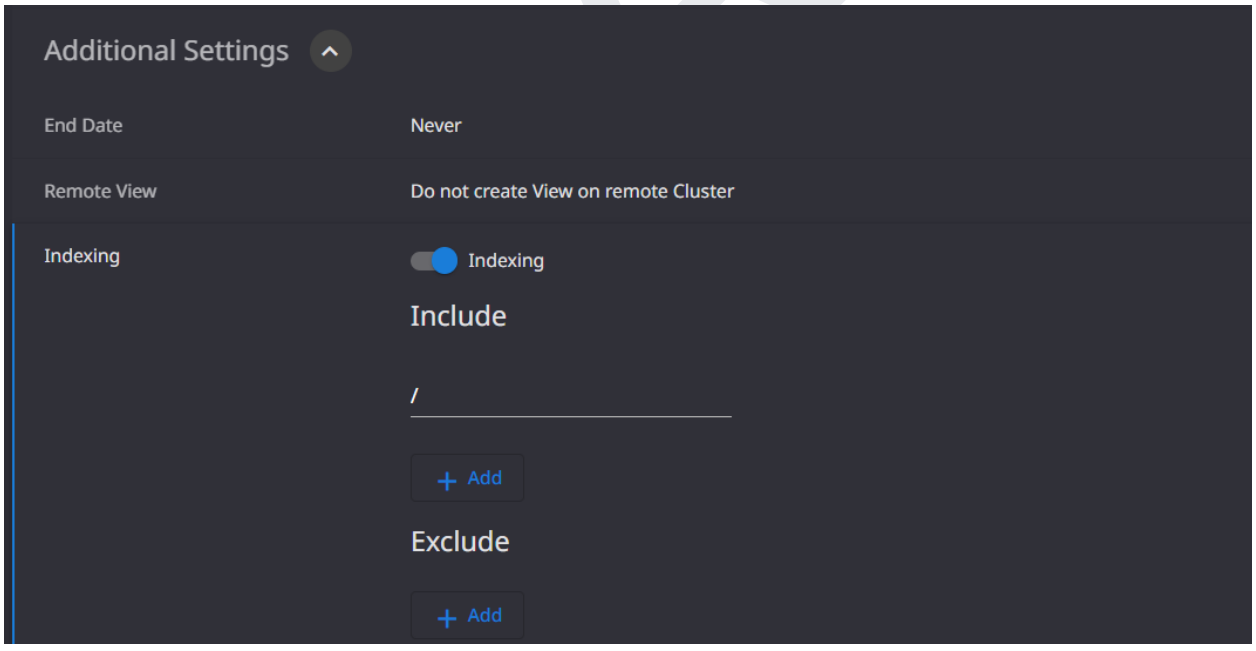

Ra
Remote Adapter
Protection Group
Name *
test
Host
Linux Hostname or IP *
10.14.17.228
Username *
test
Cluster SSH Public Key
ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAABAQCsKBf0fEkbfuA9mjaIplurfdi6DruAIMJhy+Krk+x2fEigNNIsjyT89HIDFVLdh8Ue+jcTptBE
TFQ9Yk8Vqd80Ts/9iirVwFgWwBOr5pi/bsPT4Zj2ycxN1Ix3H67OQfTb/ycXGbuuAEijlN1Fg5wvbHppYZV4JDpBVBjZl8bISL+Okix0NAF+5ls9
IJhGTjlo7Cc3OLEZXvY9dB5gIanwQ4knP7RyC7G5nmsnCYps52xaIWdU94pImAplmC0Ds/iH6lbqNWsXUe4C8NKgQ9XUjagECfd55sarf22
Mxyk/hss7WvLWcpTepm0SYcQCuBmBwTUUWErwF59fxbATO31I cohesity@virtual-robo-esx
To allow Cohesity Cluster to run the scripts remotely on Linux system, copy SSH Public Key to Clipboard, login to the Linux System with the username specified earlier and set up the permission.


Policy
Bronze
Backup
Every day | Retain 1 month
Extended Retention
Every week | Retain 3 months
Every month | Retain 1 year
Settings
Storage Domain
DefaultStorageDomain
Deduplication: Inline | Compression: Inline
NFS View
testrsyncnew
QoS Policy: Backup Target High
NFS Mount Path
10.15.15.0:/testrsyncnew



As shown in the above screenshot, please pass all the parameters. EXplanation of all the parameters can be found here

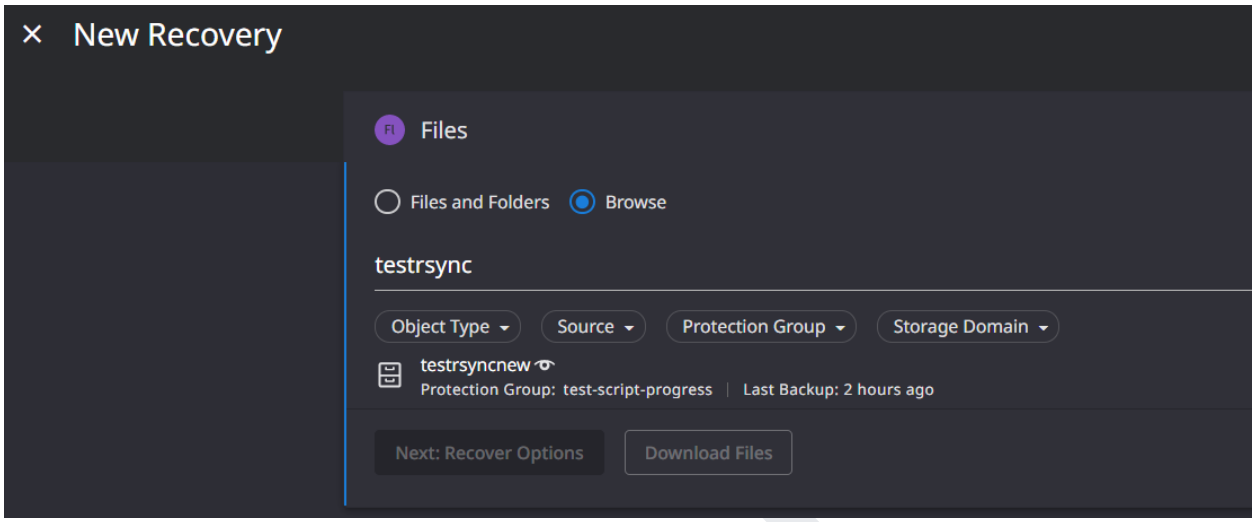
- Copy ssh public key generated on cohesity UI to the user's home directory (~/.ssh/authorized_keys). Follow instructions here
- Make sure indexing is turned on the job with include set to "/"

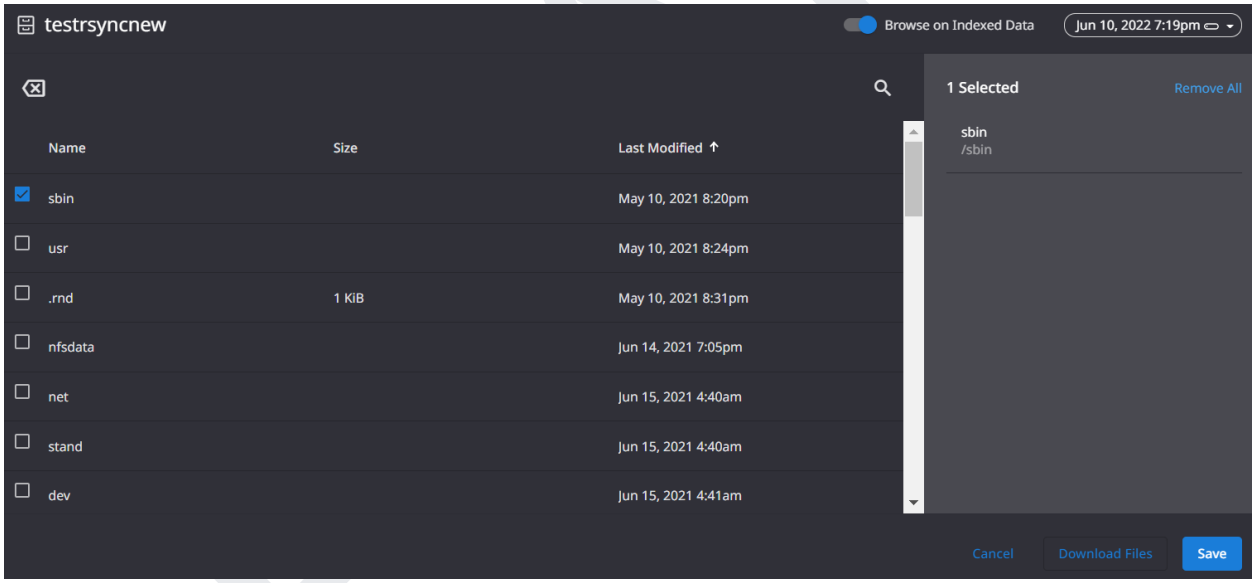


- Save and run the job
- If the job is successful, if should backup all the files to Cohesity mount and at the end of the job Cohesity will initiate a snapshot of the view
- View retention is defined by the policy used on the remote adapter job

COHESITY

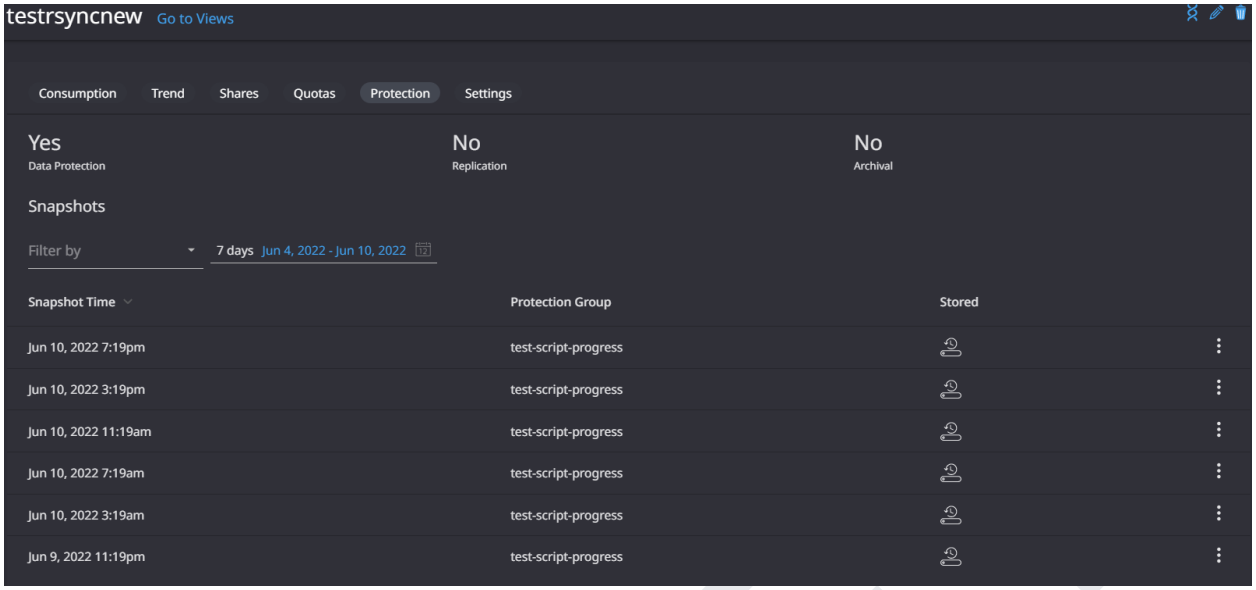
- The live view is used for regular rsync delta whenever this job is run, so there is no requirement for the script to have intelligence to retire data
- View snapshots will be used as different recovery points
- During recovery, one can directly search and download any file directly from Cohesity UI

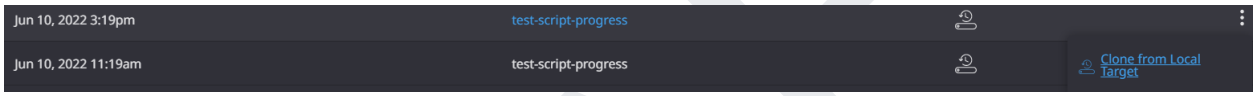


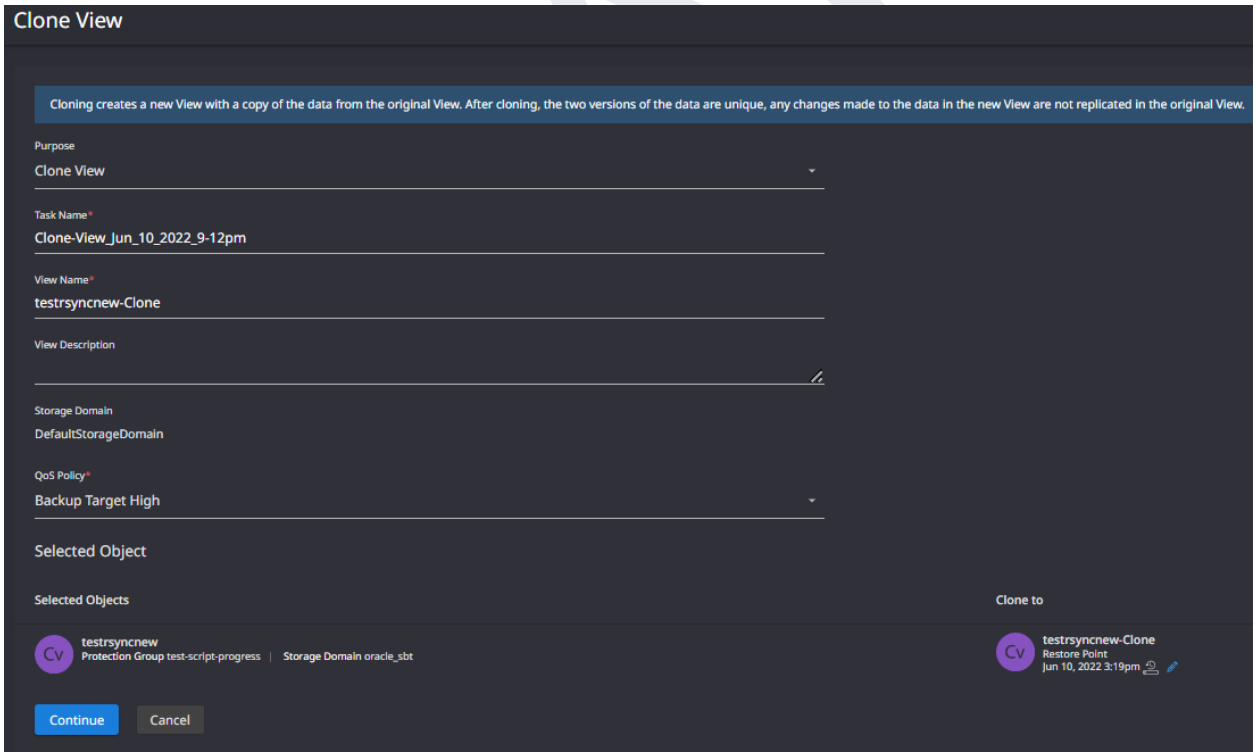


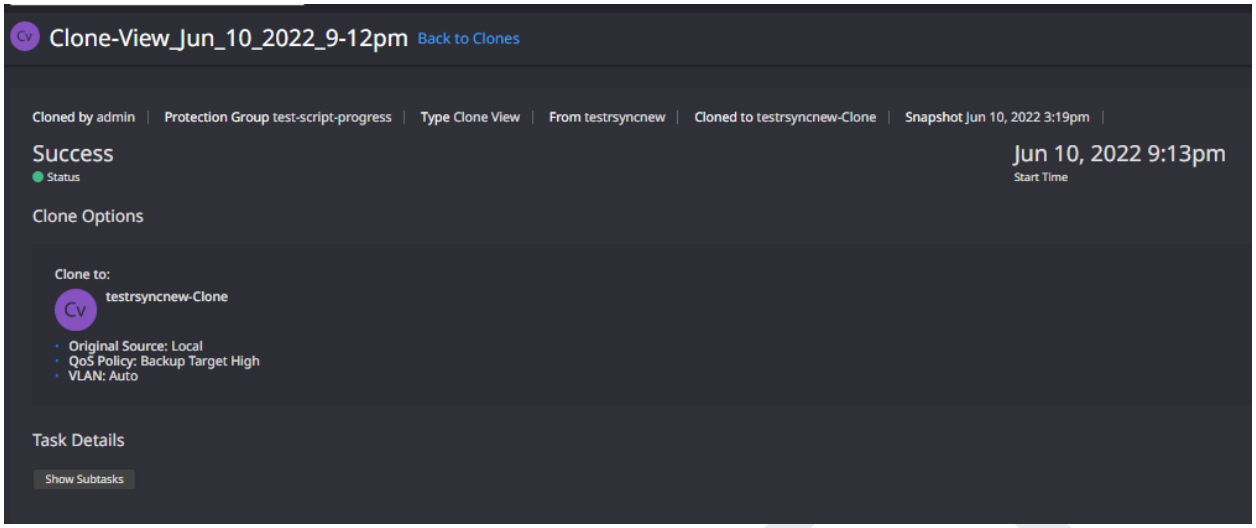
- For recovery back to source, one needs to clone required view snapshot and mount the clone to target system. Once done, any regular unix copy command can be used to perform recovery

COHESITY


testrsyncnew
Go to Views
Consumption
Trend
Shares
Quotas
Protection
Settings
Yes
Data Protection
No
Replication
No
Archival
Snapshots
Filter by
7 days Jun 4, 2022 - Jun 10, 2022
Snapshot Time
Protection Group
Stored
Jun 10, 2022 7:19pm test-script-progress
Jun 10, 2022 3:19pm test-script-progress
Jun 10, 2022 11:19am test-script-progress
Jun 10, 2022 7:19am test-script-progress
Jun 10, 2022 3:19am test-script-progress
Jun 9, 2022 11:19pm test-script-progress


Jun 10, 2022 3:19pm test-script-progress
Jun 10, 2022 11:19am test-script-progress
Clone from Local Target


Clone View
Cloning creates a new View with a copy of the data from the original View. After cloning, the two versions of the data are unique, any changes made to the data in the new View are not replicated in the original View.
Purpose
Clone View
Task Name*
Clone-View_Jun_10_2022_9-12pm
View Name*
testrsyncnew-Clone
View Description
Storage Domain
DefaultStorageDomain
QoS Policy*
Backup Target High
Selected Object
Selected Objects
Clone to
Cv
testrsyncnew
Protection Group test-script-progress | Storage Domain oracle_sbt
Cv
testrsyncnew-Clone
Restore Point
Jun 10, 2022 3:19pm
Continue
Cancel



- This cloned view can be mounted back as NFS to any unix system and files recovered

COHESITY

About the Author

Manoj Mittal is a Product Solutions Architect at Cohesity. In his role, Manoj focuses on Data protection and Platform.

Document Version History

VERSION	DATE	DOCUMENT HISTORY
0.1	June 2022	Draft

COHESITY

ABOUT COHESITY

Cohesity ushers in a new era in data management that solves a critical challenge facing businesses today: mass data fragmentation. The vast majority of enterprise data — backups, archives, file shares, object stores, and data used for dev/test and analytics — sits in fragmented infrastructure silos that makes it hard to protect, expensive to manage, and difficult to analyze. Cohesity consolidates silos onto one web-scale platform, spanning on-premises, cloud, and the edge, and uniquely empowers organizations to run apps on that platform — making it easier than ever to back up and extract insights from data. Cohesity is a 2019 CNBC Disruptor and was named a Technology Pioneer by the World Economic Forum.

Visit our website and blog, follow us on Twitter and LinkedIn and like us on Facebook.