COHESITY | yugabyteDB

Version 1.0

February 2025

Use Cohesity DataPlatform as Backup Storage to Protect YugabyteDB Universes

ABSTRACT

Cohesity's web-scale architecture provides the ideal platform to use SmartFilesServices as a backup storage for YugabyteDB Protection. This guide helps you implement YugabyteDB using Cohesity as an immutable, globally deduplicated and compressed web-scale storage target.

Table of Contents

Figures

Cohesity YugabyteDB Solution

YugabyteDB

YugabyteDB is a distributed SQL database that combines the principles of distributed systems with traditional database concepts. Its unified architecture natively distributes data and supports multiple APIs (PostgreSQL and Cassandra). The design of YugabyteDB ensures data management and processing across multiple nodes or servers. This approach provides resiliency, consistency, high availability, and fault tolerance. YugabyteDB supports both scale-out RDBMS and internet-scale OLTP workloads. The database offers low query latency and extreme resilience against failures. Global data distribution is another key aspect of its architecture.

You can use YugabyteDB to schedule and manage backups of your universe data (PostgreSQL and Cassandra). This includes the following features:

- On-demand backup and restore.
- Scheduled backups. Schedule backups at regular intervals, along with retention periods.
- Incremental backups. Create a schedule to take full backups periodically and incremental backups between those full backups.
- Configurable performance parameters. Tune parallelization and buffers for faster backup and restore performance. In most cases, this results in 5x or more speed improvements in backups and restores.
- Point-in-time recovery. Recover universe data from a specific point in time.
- Flexible storage. Store backups in the cloud or in your data center.

Cohesity Yugabyte Solution

You can use YugabyteDB Anywhere to schedule and manage backups of your universe data. Being storage-agnostic, Yugabyte supports various backup storage types. When Yugabyte customers use Cohesity as the storage for their backups, they benefit immediately from Cohesity's many features. Our solution uses Cohesity S3 Views as a backup storage for YugabyteDB Anywhere Backup Universes.

Combining YugabyteDB Anywhere Backup Universes with Cohesity provides a comprehensive, highly scalable, and flexible backup solution that would fit the data protection needs of any size organization.

Benefits of Using Cohesity as Backup Target for YugabyteDB Universe

Once you start using Cohesity as a backup target for the YugabyteDB universe, you can immediately take advantage of Cohesity's powerful features, making it an excellent choice as a backup target.

Figure 1: Cohesity Yugabyte Data Protection

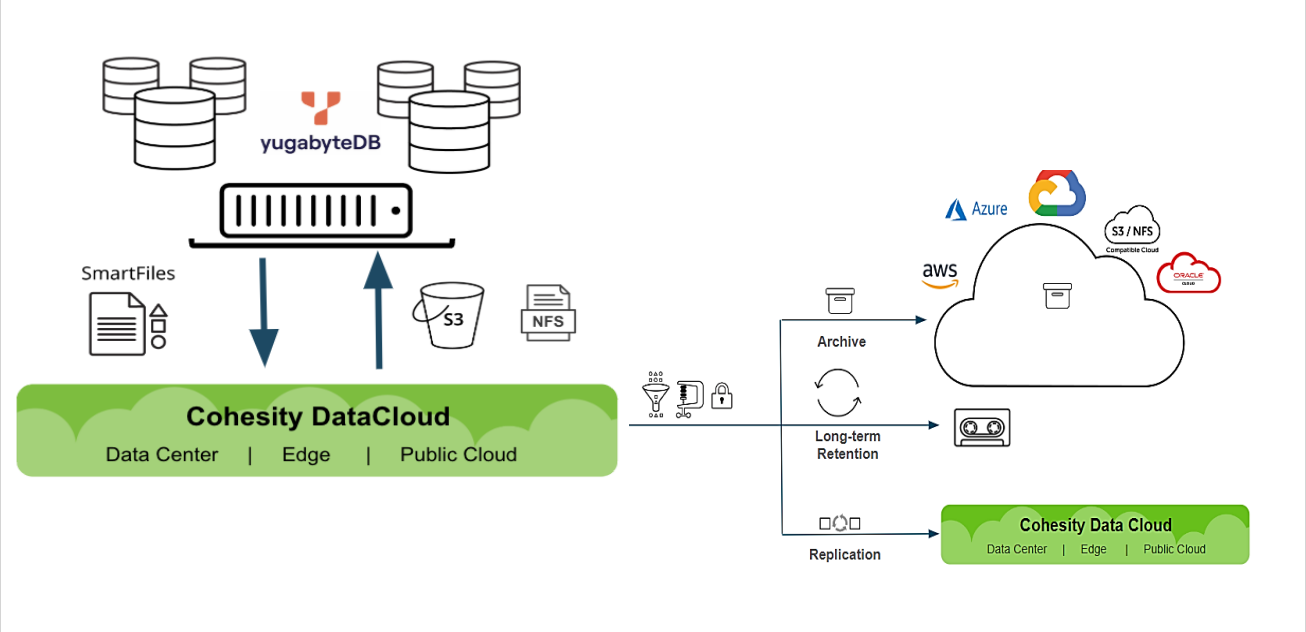


Table 1: Cohesity Platform Features and Benefits

Feature	Benefits
SmartFiles (Files and Object Services)	SmartFiles is an enterprise-class, software-defined, data-centric, multiprotocol file and object solution.
Data Protection	Policy-based data protection and flexible recovery with global search and recovery options to the same or alternate location.
DataLock	DataLock enables your storage to be WORM (write once read many)-compliant. Data such as backups and archives are immutable and cannot be tampered with or deleted.
S3 Object Lock	Cohesity supports S3 Object Lock, so you can prevent objects from being deleted or overwritten as soon as an object lands on Cohesity platform even before the protection schedule runs.

Feature	Benefits
Ransomware Protection	Machine Driven Intelligence to identify a clean copy of data, Anomaly Detection & Reporting.
Storage Efficiency	Maximizes storage capacity with Cohesity's advanced data-reduction technologies, global deduplication, and compression.
Secure Key Management	A Key Management Service (KMS) helps you manage the keys that encrypt data and secure it from unauthorized access.
Scale Out Approach	Multi-Node scale-out architecture enables parallelism, performance, infinite scale, fault tolerance, and resiliency.
Data Isolation	Cohesity FortKnox ensures virtually air-gapped data copy, management isolation, and network isolation enabling ransomware attacks.
Cloud Archival	Use CloudArchive for long-term retention and disaster recovery.
Disaster Recovery	Recover from any disaster using Cohesity Replication, Cloud Retrieve, Cloud Spin, and Cloud Edition.
Cloud Tier	Use automated, policy-based tiering to lower-cost storage for reduced TCO.

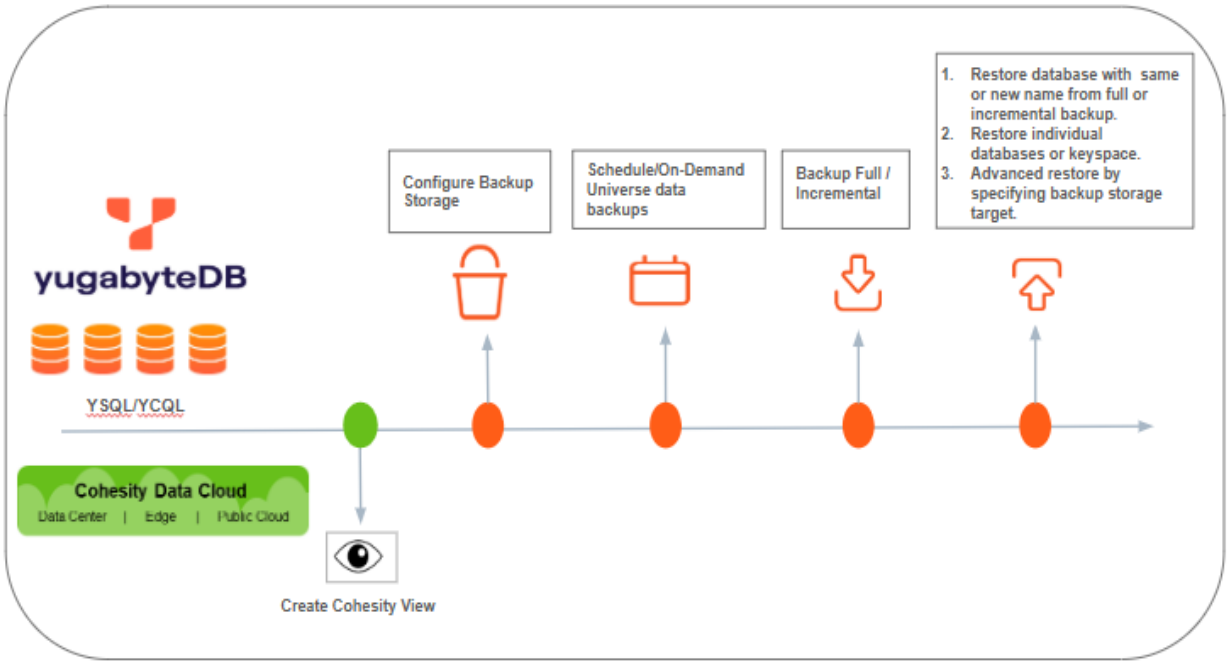
Cohesity Backup Target for YugabyteDB Universe Workflow Overview

To protect data using YugabyteDB universe, you need to configure and associate backup storage, where YugabyteDB Universe data would be stored, backup storage requires an underlying storage location. Cohesity provides this required storage in a web-scale, globally deduplicated, and compressed format by providing a Cohesity View via S3.

To use a Cohesity View as a backup storage for YugabyteDB backup universe, you need to perform the following tasks:

1. Create a Cohesity S3 View.
2. Configure S3 backup storage in YugabyteDB universe.
3. Configure your YugabyteDB backup jobs to use the backup storage that you created.

Figure 2: Cohesity - YugabyteDB Workflow



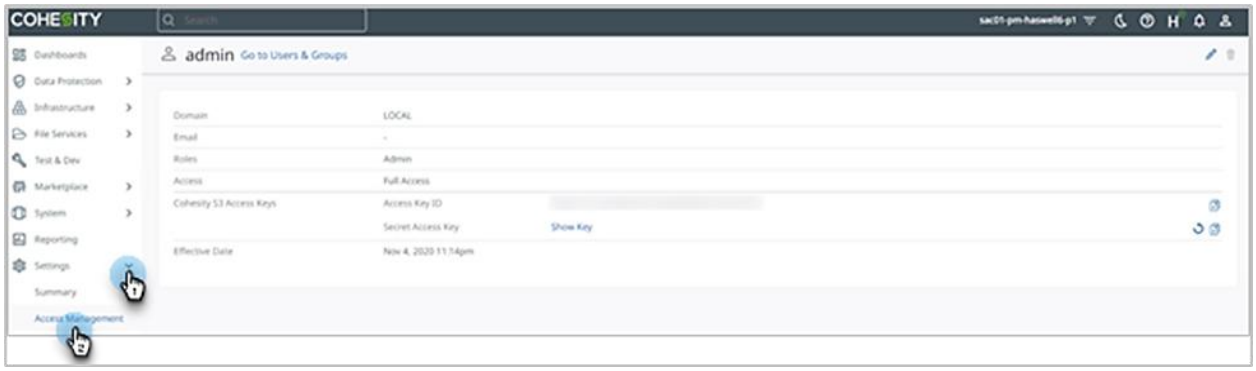
Create Cohesity S3 View for YugabyteDB Universe

To use Cohesity SmartFiles S3 View as backup storage for YugabyteDB backup universe, you will create a Cohesity View using the category of object services, choose a QoS policy, and configure the View for S3. For this solution, Cohesity recommends having inline deduplication and inline compression enabled on the Storage Domain in which you create the View. For details, see Create or Edit Storage Domains in the online Help.

To access Cohesity SmartFiles S3 Object Service as a backup storage, one would need access key and secret key of S3 bucket owner, please refer to manager access keys for more details.

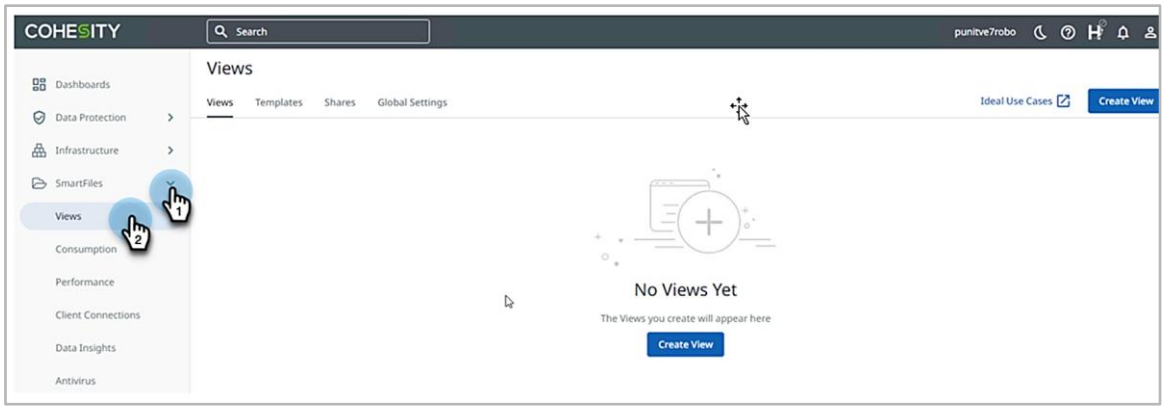
1. Access and secret keys are required to access S3 bucket created on Cohesity Cluster. Go to the following location to copy the access key and secret key.

   Cohesity Cluster UI > **Settings** > **Access Management** > **User** (which is used to create the S3 View).

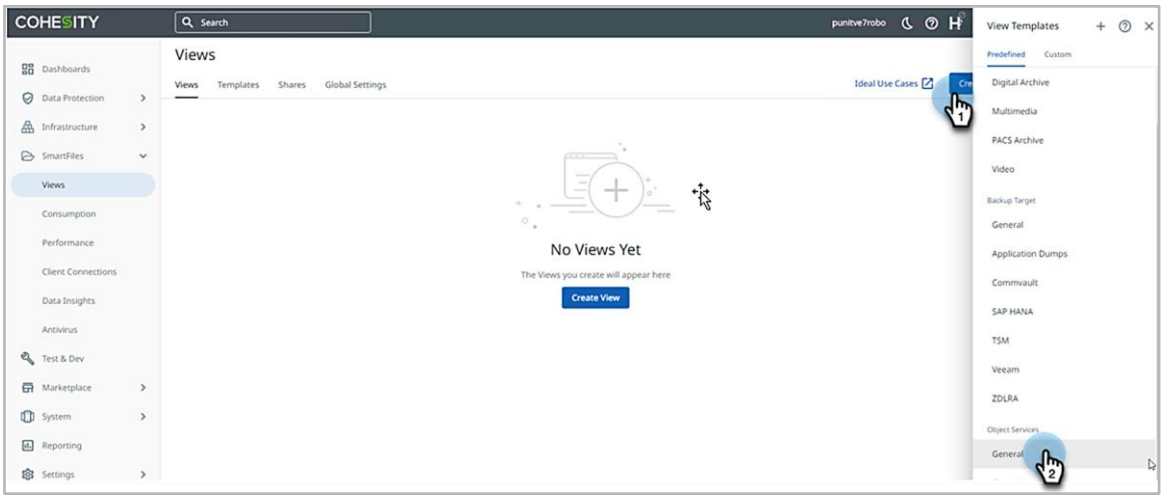


   Create a Cohesity View, select the optimal QoS policy, select Object Service, set Object key Pattern, and Read/Write protocol as S3.

2. Log in to Cohesity and navigate to **SmartFiles** > **Views**.

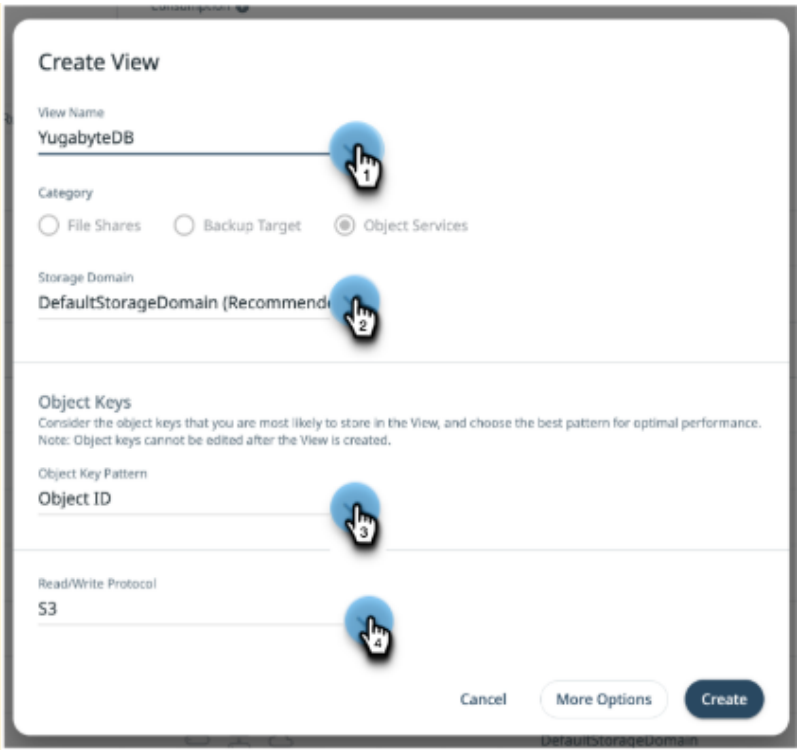


On the **Views** page, click **"Create View"** and click **on General** in Object Services**.**



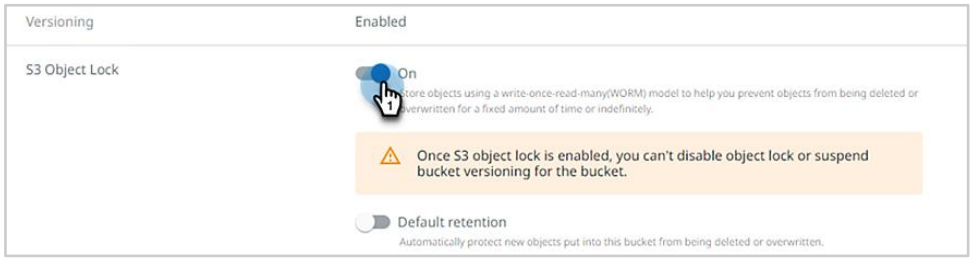
Name the View in the **Create View** form, choose the **Storage Domain**, and **Category** as **Object Services** is selected by default, and select Cohesity recommended Object Key Pattern "**Object ID"** for Performance and scalability under **Object Key Pattern**. Under **Read/Write Protocol**, select **S3** and Click **More Options**.

NOTE: Cohesity recommends using the Object ID as object key pattern for S3 backup storage Object ID key partner is available from Cohesity cluster version 7.0 and later.

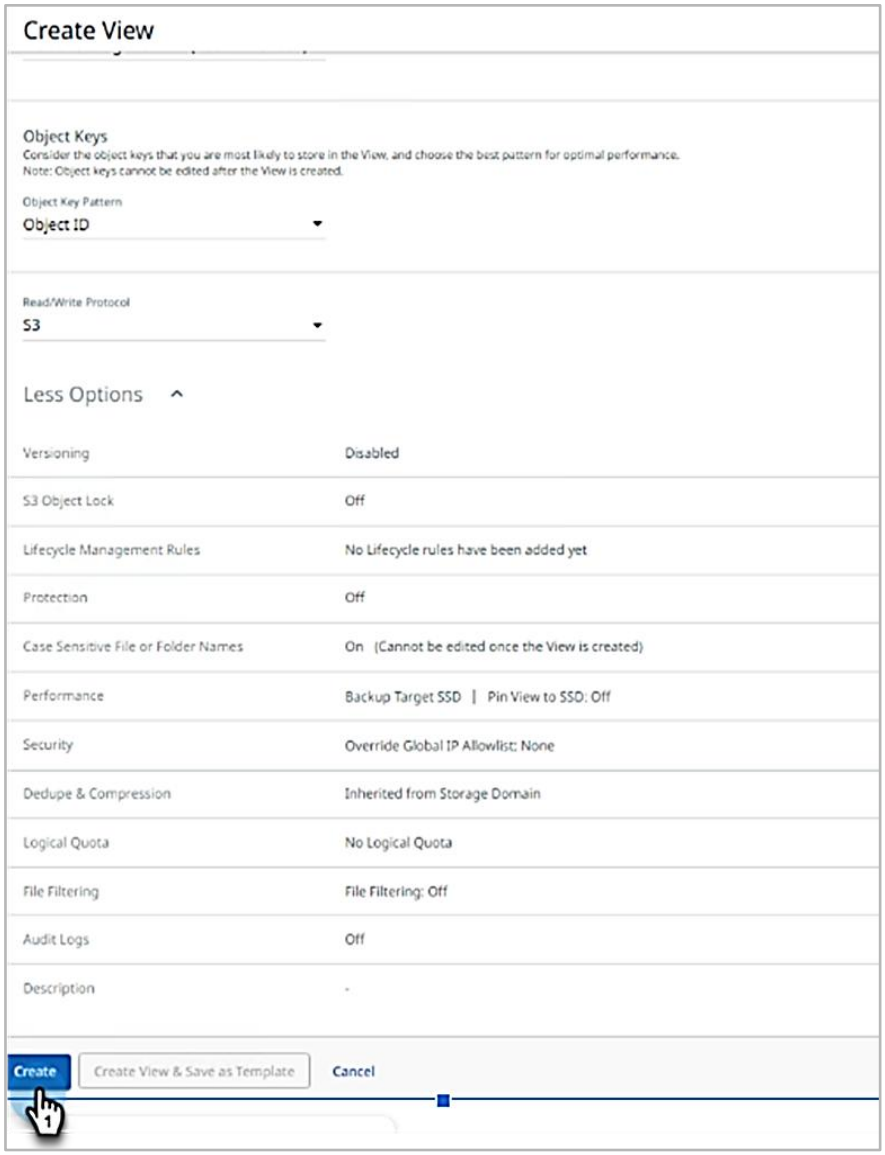


(**Optional)** You can enable the S3 Object Lock for Immutability if you want to use it with S3 Object Lock-enabled bucket. This will work only with versioned S3 Views. If you want to avoid using the Object Lock, skip the step below.

In the same form, under S3 Object Lock, toggle the button to enable the versioning on the SmartFiles S3 View by default.



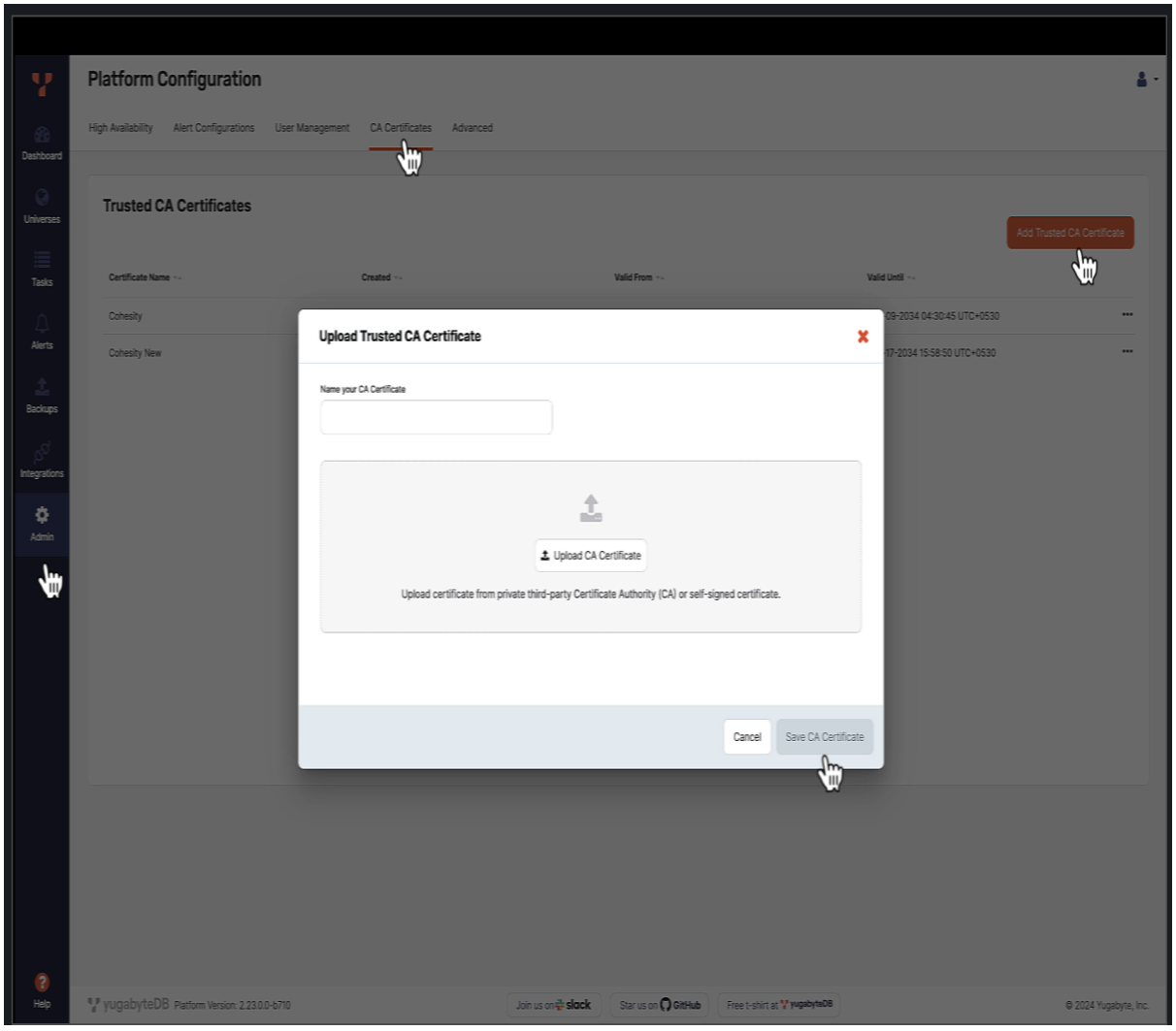
2. In the same form, Under **Performance**, click **Edit** (✎) and select the QoS Policy as Backup Target SSD.
3. Click **Create View** at the bottom of the form. View will be listed in the **View** Tab.



Refer to Appendix B to verify the communication between the YugabyteDB and Cohesity Cluster.

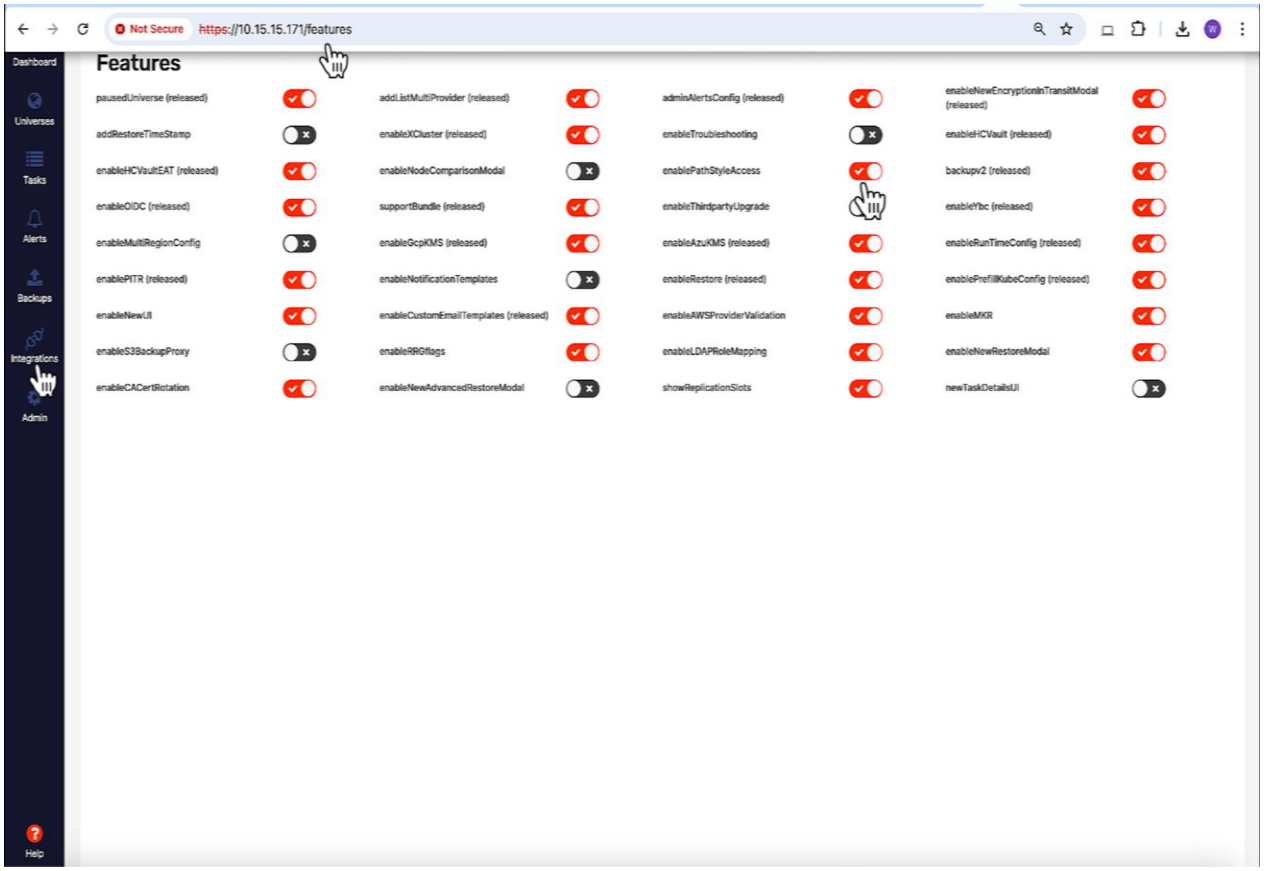
Configure Cohesity S3 Storage for YugabyteDB Universe

1. Login to YugabyteDB Admin Console and navigate to **Admin** > **CA Certificates** > **Add Trusted CA Certificate** > **Upload CA Certificate** > **Save CA Certificate** herein it is mandatory to upload Cohesity Cluster CA Certificate on YugabyteDB Anywhere platform to establish trust for secure S3 communication.

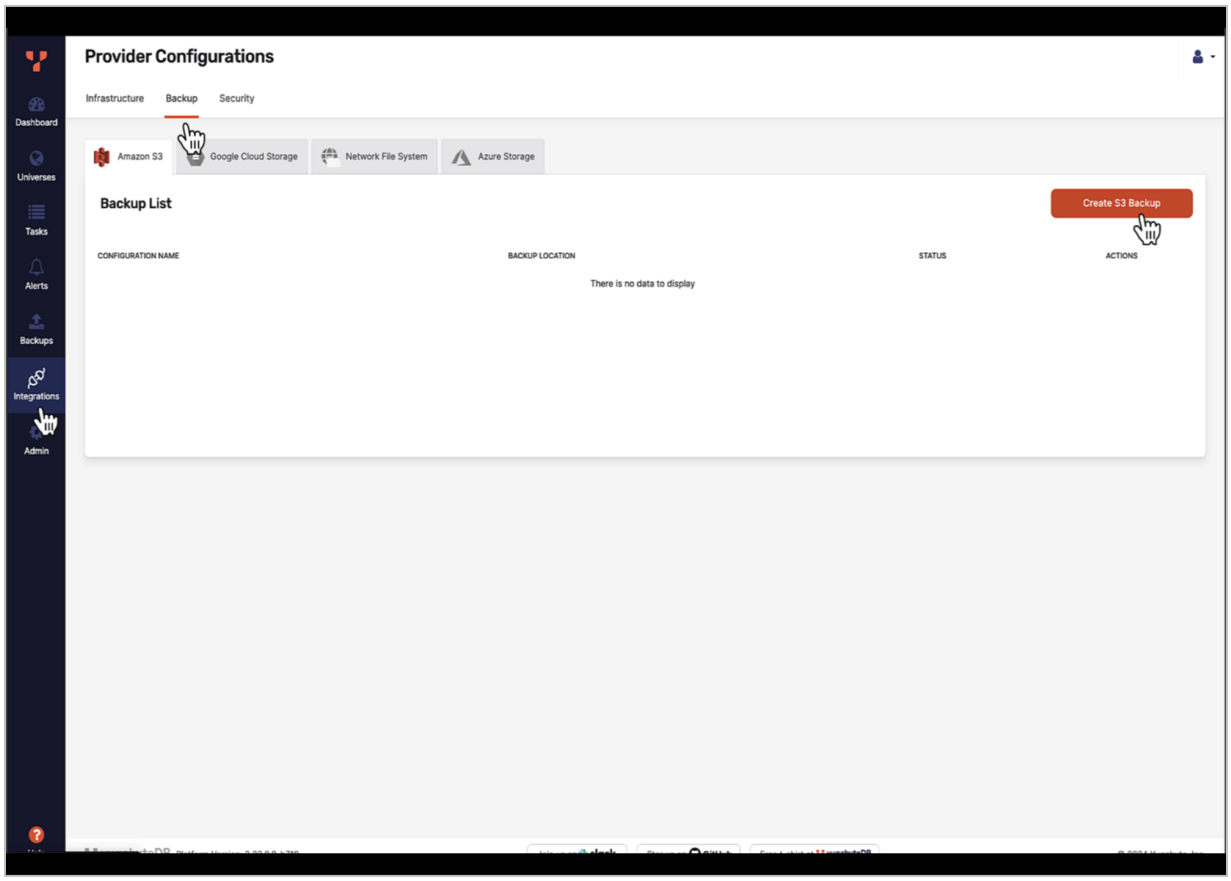


2. On YugabyteDB Anywhere **enablePathStyleAccess** feature. Feature flag can be enabled by adding "/features" after hostname or IP address to the admin console URL. Navigate to enablePathStyleAccess and switch the toggle button to ON.

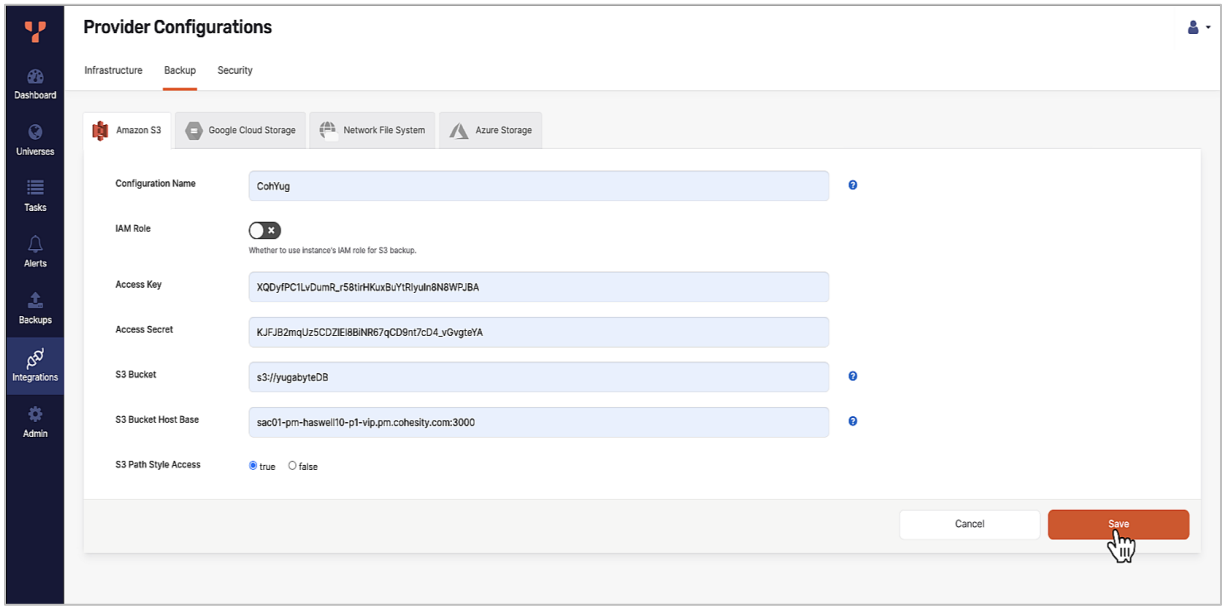
NOTE: Do not refresh GUI, instead proceed with the next step by clicking Integration.



3. To Configure Cohesity S3 storage as backup target for YugabyteDB Universe, Navigate to **Integrations > Backup** and click **Create S3 Backup.**



4. Once you click **Create S3 Backup,** the configuration form appears as shown below.



- **Configuration Name**: field to provide a meaningful name for your storage configuration.
- **Access Key and Access Secret** fields: Access and secret key are required to access S3 buckets created on Cohesity Cluster. Refer to Manage Cohesity S3 Access keys for getting Access for the Cohesity user who is the owner of S3 bucket.

  Cohesity Cluster UI > Settings > Access Management > User (which is used to create the S3 View).
- **S3 Bucket**: Bucket name refers to S3 bucket created on Cohesity which you plan to use as backup storage.
- **S3 Bucket Host Base**: The host base field refers to Cohesity Cluster URL affixed with a port e.g. https://cohesity.com:3000.
- **Save**: Save of backup configuration only succeeds if Cohesity S3 bucket is fully accessible.

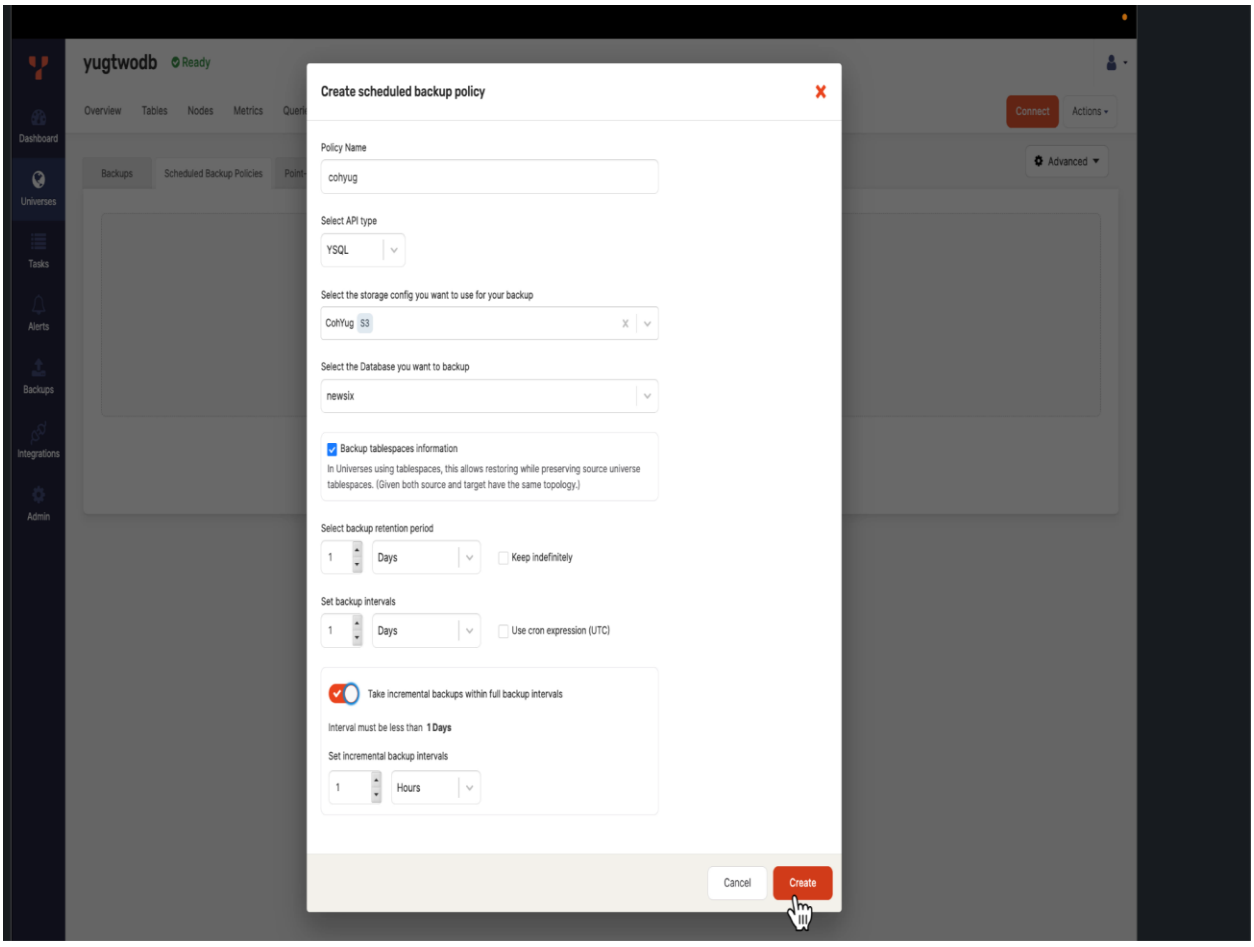
Also, refer to Configure backup storage for more details.

Schedule/On-Demand Data Backups from YugabyteDB Universe

You can use YugabyteDB Anywhere to perform regularly scheduled/On-demand backups of YugabyteDB universe data for all tables in a database (YSQL) or keyspace (YCQL) or only the specified tables (YCQL only).

If you are using YBA version 2.16 or later to manage universes with YugabyteDB version 2.16 or later, you can additionally create incremental backups and configure backup performance parameters. Please refer to Back up universe data for triggering on-demand full and incremental backups

To schedule backups, backups must be enabled for the universe. On the universe Tables tab, click Actions to verify that backups are enabled. If disabled, click Enable Backup. Please refer to Schedule data backups for a detailed explanation on policy based scheduling, retention, backup options, disabling backup and deleting schedules.



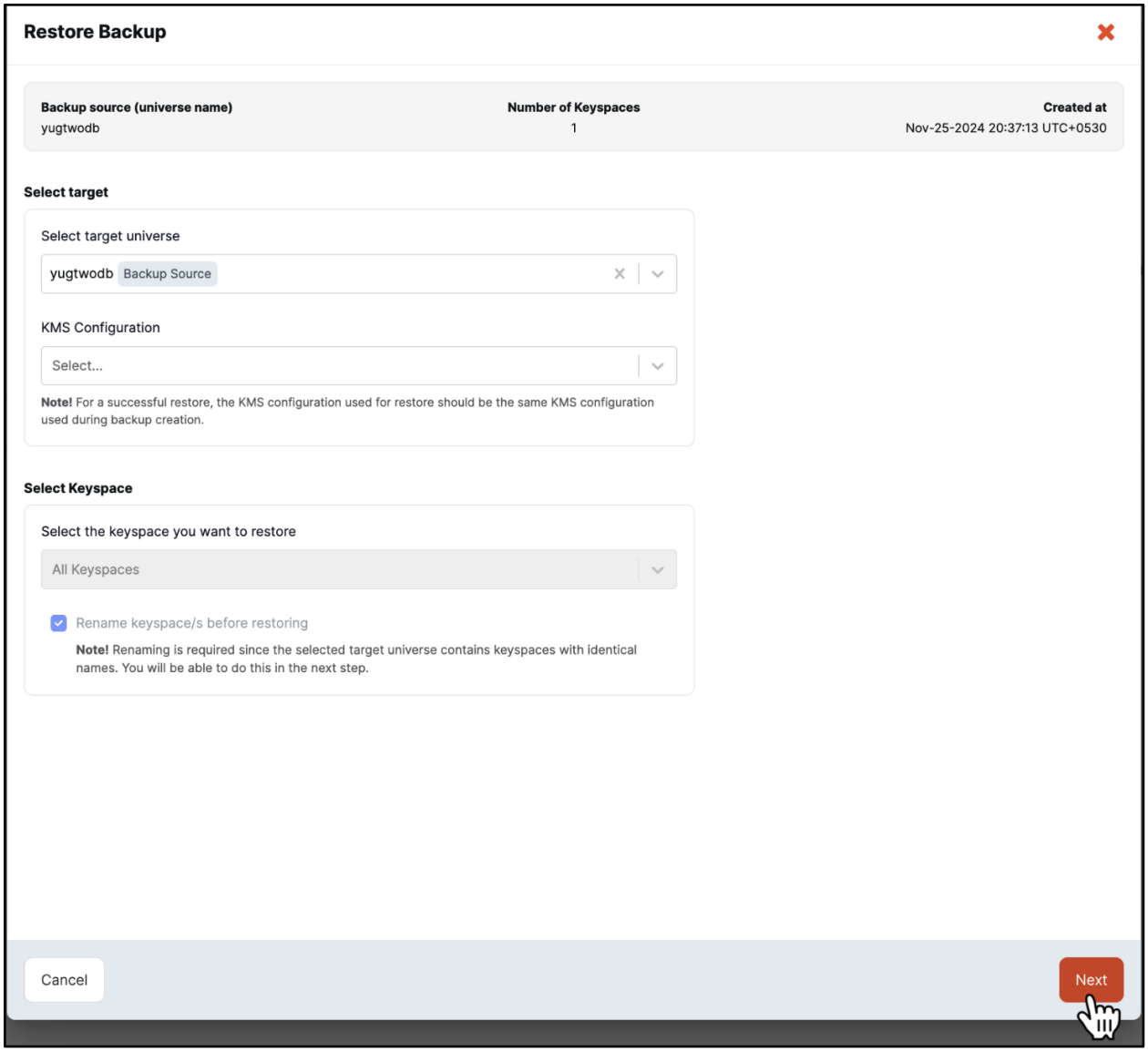
Restore YugabyteDB Universe Data

YugabyteDB universe supports the below types of restores from backups stored on Cohesity storage.

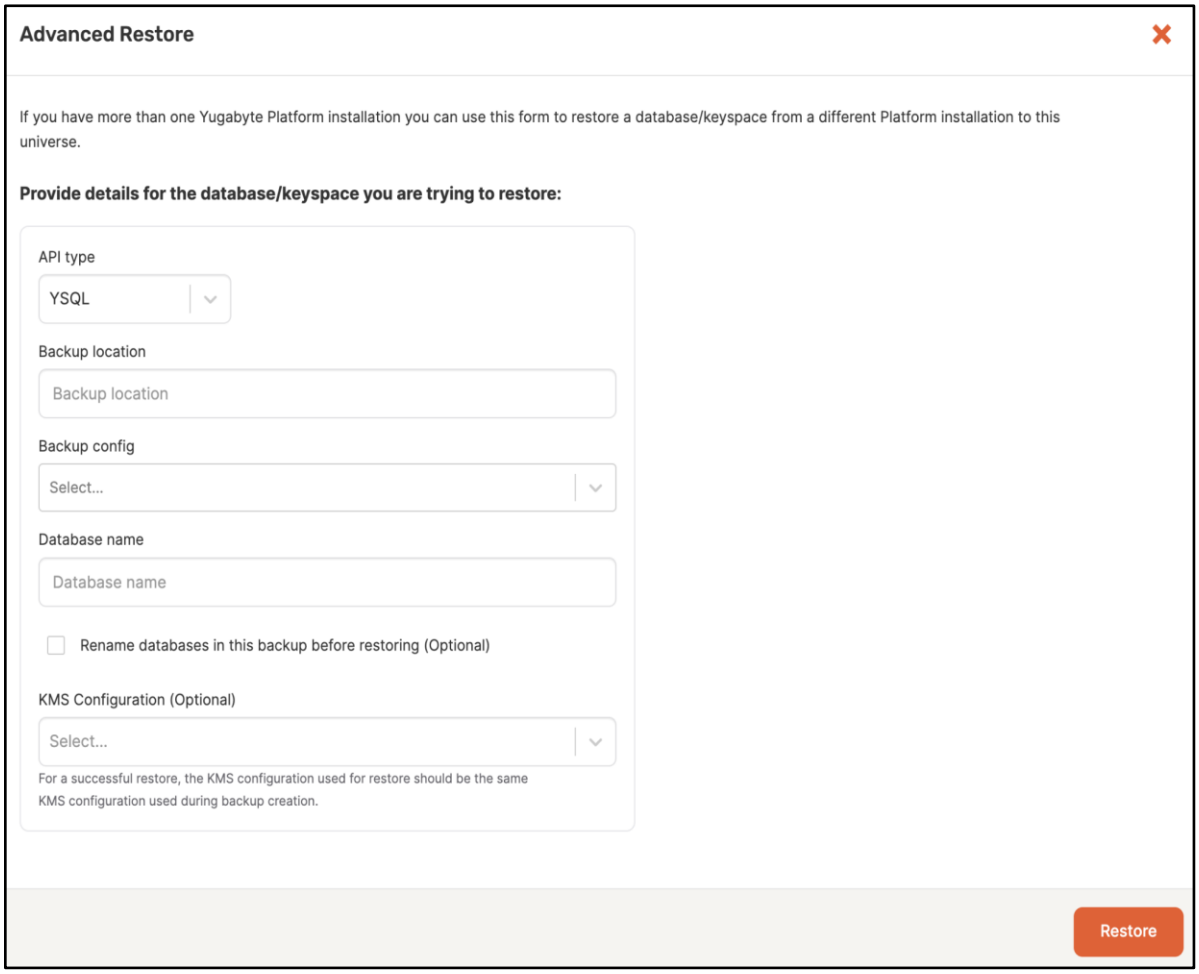
- Restore from full or incremental backups
- Restore selective databases or keyspaces and tables
- Advanced restore wherein backup is moved to different location by accessing specified backup storage location

For details on pre-requisites & restore types options and please refer to Restore universe data.

1. Restore from full or incremental backups, for description of the various options refer to Restore universe data.



2. Advanced restore by specifying backup location and configuration. For description of the various options, refer to Restore universe data.



Appendix A: Protect Your YugabyteDB Backup Storage from Ransomware Attacks

Protection from ransomware attacks is a major concern when it comes to enterprise data retention and security. Cohesity recommends the following best practices to keep your Yugabyte backup storage safe from ransomware and keep your data integrity intact.

- Keep the original backup data in an immutable state and avoid mounting the gold copy of the data by an external system.
- Make sure to enable multifactor authentication (MFA) and write once read many (WORM) capabilities for the snapshots.
- To detect attacks in real-time, continuously monitor the data, and the solution and analyze files and audit logs to detect abnormal or even smaller change rates. Note that relying exclusively on backup data-ingest change rates to detect such behaviors is insufficient.

Appendix B: Verify SmartFiles S3 View Connectivity

To verify whether the SmartFiles S3 View is accessible from the YugabyteDB Universe and database servers you use to write the backup to the S3 View, check the connectivity using the following command on telnet from the database machine to port 3000:

```
telnet <Cohesity_Cluster_Vip> 3000
```

Your Feedback

Was this document helpful? Send us your feedback!

About the Authors

Waseem Khan is a Sr. Product Solutions Architect at Cohesity. In his role, Waseem focuses on data protection for traditional and modern Databases.

Other major contributors included:

- Akshay Kumar Sr Mgr, Product Solutions
- Adaikkappan Arumugam, Sr, Dir, Product Solutions
- Touhid Shaikh, Staff Product Manager

Document Version History

VERSION	DATE	DOCUMENT HISTORY
1.0	Feb 2025	First Release

ABOUT COHESITY

Cohesity is a leader in AI-powered data security and management. Aided by an extensive ecosystem of partners, Cohesity makes it easier to protect, manage, and get value from data – across the data center, edge, and cloud. Cohesity helps organizations defend against cybersecurity threats with comprehensive data security and management capabilities, including immutable backup snapshots, AI-based threat detection, monitoring for malicious behavior, and rapid recovery at scale. Cohesity solutions are delivered as a service, self-managed, or provided by a Cohesity-powered partner. Cohesity is headquartered in San Jose, CA, and is trusted by the world's largest enterprises, including six of the Fortune 10 and 44 of the Fortune 100.

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