Migrating data into StorageGRID

StorageGRID 11.5

NetApp
January 04, 2024

Table of Contents

Migrating data into StorageGRID

You can migrate large amounts of data to the StorageGRID system while simultaneously using the StorageGRID system for day-to-day operations.

The following section is a guide to understanding and planning a migration of large amounts of data into the StorageGRID system. It is not a general guide to data migration, and it does not include detailed steps for performing a migration. Follow the guidelines and instructions in this section to ensure that data is migrated efficiently into the StorageGRID system without interfering with day-to-day operations, and that the migrated data is handled appropriately by the StorageGRID system.

Confirming capacity of the StorageGRID system

Before migrating large amounts of data into the StorageGRID system, confirm that the StorageGRID system has the disk capacity to handle the anticipated volume.

If the StorageGRID system includes an Archive Node and a copy of migrated objects has been saved to nearline storage (such as tape), ensure that the Archive Node's storage has sufficient capacity for the anticipated volume of migrated data.

As part of the capacity assessment, look at the data profile of the objects you plan to migrate and calculate the amount of disk capacity required. For details about monitoring the disk capacity of your StorageGRID system, see the instructions for monitoring and troubleshooting StorageGRID.

Related information

Monitor & troubleshoot

Managing Storage Nodes

Determining the ILM policy for migrated data

The StorageGRID system's ILM policy determines how many copies are made, the locations to which copies are stored, and for how long these copies are retained. An ILM policy consists of a set of ILM rules that describe how to filter objects and manage object data over time.

Depending on how migrated data is used and your requirements for migrated data, you might want to define unique ILM rules for migrated data that are different from the ILM rules used for day-to-day operations. For example, if there are different regulatory requirements for day-to-day data management than there are for the data that is included in the migration, you might want a different number of copies of the migrated data on a different grade of storage.

You can configure rules that apply exclusively to migrated data if it is possible to uniquely distinguish between migrated data and object data saved from day-to-day operations.

If you can reliably distinguish between the types of data using one of the metadata criteria, you can use this criteria to define an ILM rule that applies only to migrated data.

Before beginning data migration, ensure that you understand the StorageGRID system's ILM policy and how it will apply to migrated data, and that you have made and tested any changes to the ILM policy.

> An ILM policy that has been incorrectly specified can cause unrecoverable data loss. Carefully review all changes you make to an ILM policy before activating it to make sure the policy will work as intended.

Related information

Manage objects with ILM

Impact of migration on operations

A StorageGRID system is designed to provide efficient operation for object storage and retrieval, and to provide excellent protection against data loss through the seamless creation of redundant copies of object data and metadata.

However, data migration must be carefully managed according to the instructions in this chapter to avoid having an impact on day-to-day system operations, or, in extreme cases, placing data at risk of loss in case of a failure in the StorageGRID system.

Migration of large quantities of data places additional load on the system. When the StorageGRID system is heavily loaded, it responds more slowly to requests to store and retrieve objects. This can interfere with store and retrieve requests which are integral to day-to-day operations. Migration can also cause other operational issues. For example, when a Storage Node is nearing capacity, the heavy intermittent load due to batch ingest can cause the Storage Node to cycle between read-only and read-write, generating notifications.

If the heavy loading persists, queues can develop for various operations that the StorageGRID system must perform to ensure full redundancy of object data and metadata.

Data migration must be carefully managed according to the guidelines in this document to ensure safe and efficient operation of the StorageGRID system during migration. When migrating data, ingest objects in batches or continuously throttle ingest. Then, continuously monitor the StorageGRID system to ensure that various attribute values are not exceeded.

Scheduling data migration

Avoid migrating data during core operational hours. Limit data migration to evenings, weekends, and other times when system usage is low.

If possible, do not schedule data migration during periods of high activity. However, if it is not practical to completely avoid the high activity period, it is safe to proceed as long as you closely monitor the relevant attributes and take action if they exceed acceptable values.

Related information

Monitoring data migration

Monitoring data migration

Data migration must be monitored and adjusted as necessary to ensure data is placed according to the ILM policy within the required timeframe.

This table lists the attributes you must monitor during data migration, and the issues that they represent.

If you use traffic classification policies with rate limits to throttle ingest, you can monitor the observed rate in conjunction with the statistics described in the following table and reduce the limits if necessary.

Monitor	Description
Number of objects waiting for ILM evaluation	1. Select **Support** > **Tools** > **Grid Topology**. 2. Select ***deployment*** > **Overview** > **Main**. 3. In the ILM Activity section, monitor the number of objects shown for the following attributes: ◦ **Awaiting - All (XQUZ)**: The total number of objects awaiting ILM evaluation. ◦ **Awaiting - Client (XCQZ)**: The total number of objects awaiting ILM evaluation from client operations (for example, ingest). 4. If the number of objects shown for either of these attributes exceeds 100,000, throttle the ingest rate of objects to reduce the load on the StorageGRID system.
Targeted archival system's storage capacity	If the ILM policy saves a copy of the migrated data to a targeted archival storage system (tape or the cloud), monitor the capacity of the targeted archival storage system to ensure that there is sufficient capacity for the migrated data.
Archive Node > **ARC** > **Store**	If an alarm for the **Store Failures (ARVF)** attribute is triggered, the targeted archival storage system might have reached capacity. Check the targeted archival storage system and resolve any issues that triggered an alarm.

Creating custom notifications for migration alarms

You might want StorageGRID to send alert notifications or alarm (legacy system) notifications to the system administrator responsible for monitoring migration if certain values exceed recommended thresholds.

What you'll need

- You must be signed in to the Grid Manager using a supported browser.
- You must have specific access permissions.
- You must have configured email settings for alert (or alarm) notifications.

Steps

1. Create a custom alert rule or a Global Custom alarm for each Prometheus metric or StorageGRID attribute

you want to monitor during data migration.

Alerts are triggered based on Prometheus metric values. Alarms are triggered based on attribute values. See the instructions for monitoring and troubleshooting StorageGRID for more information.

2. Disable the custom alert rule or the Global Custom alarm after data migration is complete.

   Note that Global Custom alarms override Default alarms.

Related information

Monitor & troubleshoot

Copyright information

Trademark information