Configuring Archive Node connections to archival storage

StorageGRID 11.5

NetApp
August 30, 2024

Table of Contents

Configuring Archive Node connections to archival storage

When you configure an Archive Node to connect with an external archive, you must select the target type.

The StorageGRID system supports the archiving of object data to the cloud through an S3 interface or to tape through Tivoli Storage Manager (TSM) middleware.

> Once the type of archival target is configured for an Archive Node, the target type cannot be changed.

- Archiving to the cloud through the S3 API
- Archiving to tape trough TSM middleware
- Configuring Archive Node retrieve settings
- Configuring Archive Node replication

Archiving to the cloud through the S3 API

You can configure an Archive Node to connect directly to Amazon Web Services (AWS) or to any other system that can interface to the StorageGRID system through the S3 API.

> Moving objects from an Archive Node to an external archival storage system through the S3 API has been replaced by ILM Cloud Storage Pools, which offer more functionality. The **Cloud Tiering - Simple Storage Service (S3)** option is still supported, but you might prefer to implement Cloud Storage Pools instead.

If you are currently using an Archive Node with the **Cloud Tiering - Simple Storage Service (S3)** option, consider migrating your objects to a Cloud Storage Pool. See the instructions for managing objects with information lifecycle management.

Related information

Manage objects with ILM

Configuring connection settings for the S3 API

If you are connecting to an Archive Node using the S3 interface, you must configure the connection settings for the S3 API. Until these settings are configured, the ARC service remains in a Major alarm state as it is unable to communicate with the external archival storage system.

Moving objects from an Archive Node to an external archival storage system through the S3 API has been replaced by ILM Cloud Storage Pools, which offer more functionality. The **Cloud Tiering - Simple Storage Service (S3)** option is still supported, but you might prefer to implement Cloud Storage Pools instead.

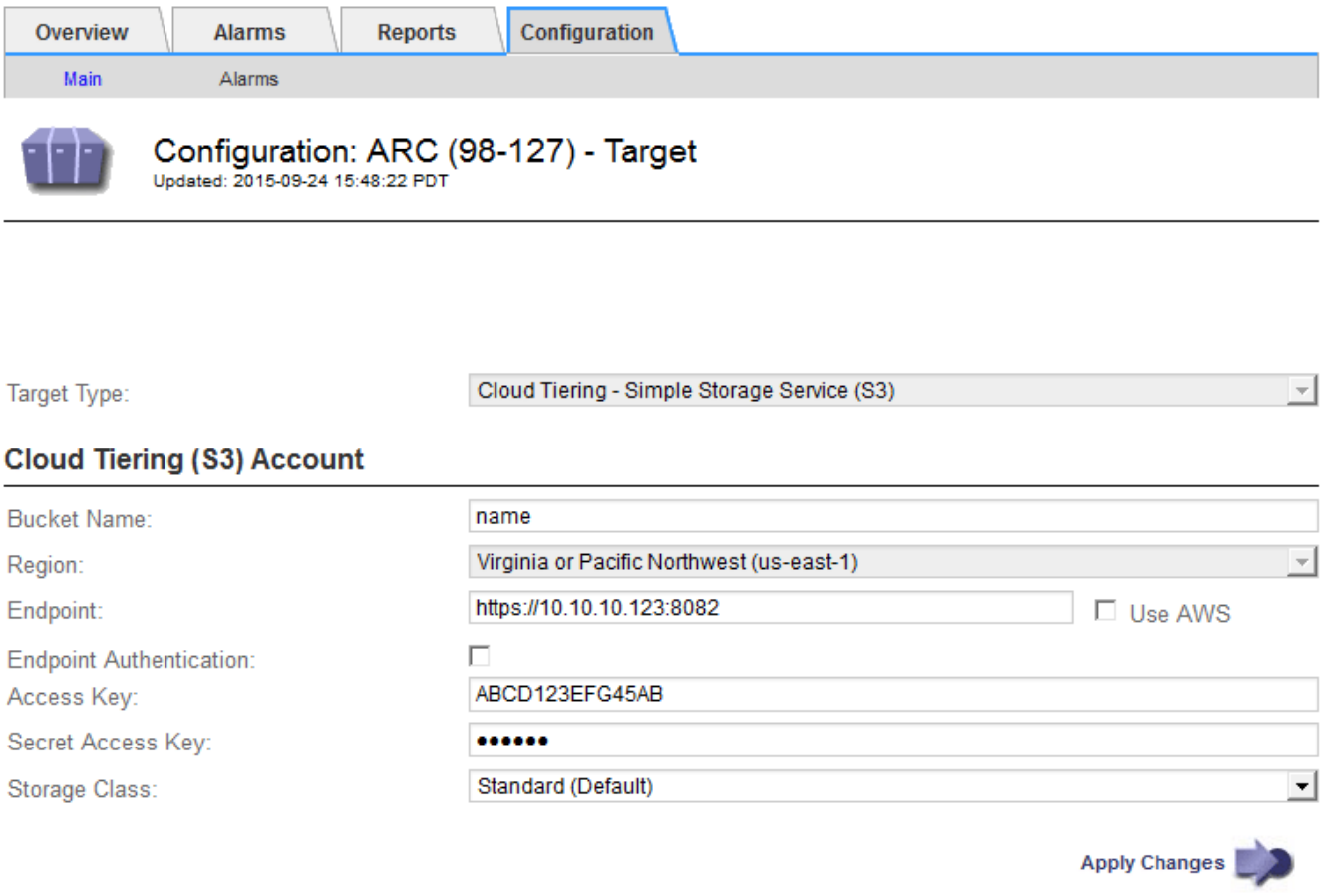
If you are currently using an Archive Node with the **Cloud Tiering - Simple Storage Service (S3)** option, consider migrating your objects to a Cloud Storage Pool. See the instructions for managing objects with information lifecycle management.

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 created a bucket on the target archival storage system:
  - The bucket must be dedicated to a single Archive Node. It cannot be used by other Archive Nodes or other applications.
  - The bucket must have the appropriate region selected for your location.
  - The bucket should be configured with versioning suspended.
- Object Segmentation must be enabled and the Maximum Segment Size must be less than or equal to 4.5 GiB (4,831,838,208 bytes). S3 API requests that exceed this value will fail if S3 is used as the external archival storage system.

Steps

1. Select **Support** > **Tools** > **Grid Topology**.
2. Select **Archive Node** > **ARC** > **Target**.
3. Select **Configuration** > **Main**.



4. Select **Cloud Tiering - Simple Storage Service (S3)** from the Target Type drop-down list.

   Configuration settings are unavailable until you select a Target Type.

5. Configure the cloud tiering (S3) account through which the Archive Node will connect to the target external S3 capable archival storage system.

   Most of the fields on this page are self-explanatory. The following describes fields for which you might need guidance.

   - **Region**: Only available if **Use AWS** is selected. The region you select must match the bucket's region.
   - **Endpoint** and **Use AWS**: For Amazon Web Services (AWS), select **Use AWS**. **Endpoint** is then automatically populated with an endpoint URL based on the Bucket Name and Region attributes. For example:

     ```
     https://bucket.region.amazonaws.com
     ```

     For a non-AWS target, enter the URL of the system hosting the bucket, including the port number. For example:

     ```
     https://system.com:1080
     ```

   - **End Point Authentication**: Enabled by default. If the network to the external archival storage system is trusted, you can unselect the check box to disable endpoint SSL certificate and hostname verification for the targeted external archival storage system. If another instance of a StorageGRID system is the target archival storage device and the system is configured with publicly signed certificates, you can keep the check box selected.

- **Storage Class**: Select **Standard (Default)** for regular storage. Select **Reduced Redundancy** only for objects that can be easily recreated. **Reduced Redundancy** provides lower cost storage with less reliability. If the targeted archival storage system is another instance of the StorageGRID system, **Storage Class** controls how many interim copies of the object are made at ingest on the target system, if dual commit is used when objects are ingested there.

6. Click **Apply Changes**.

   The specified configuration settings are validated and applied to your StorageGRID system. Once configured, the target cannot be changed.

Related information

Manage objects with ILM

Modifying connection settings for S3 API

After the Archive Node is configured to connect to an external archival storage system through the S3 API, you can modify some settings should the connection change.

What you'll need

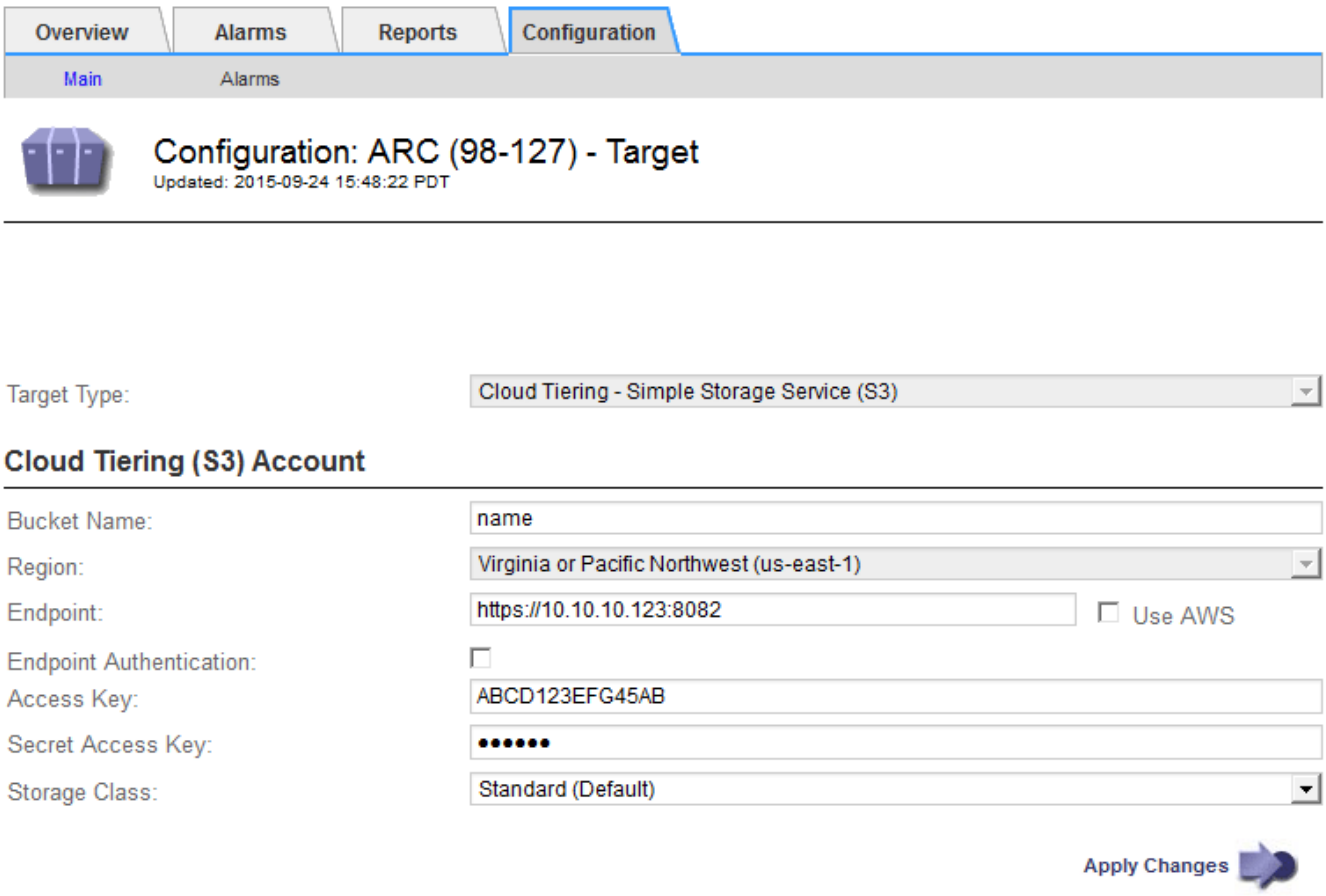
- You must be signed in to the Grid Manager using a supported browser.
- You must have specific access permissions.

About this task

If you change the Cloud Tiering (S3) account, you must ensure that the user access credentials have read/write access to the bucket, including all objects that were previously ingested by the Archive Node to the bucket.

Steps

1. Select **Support** > **Tools** > **Grid Topology**.
2. Select ***Archive Node*** > **ARC** > **Target**.
3. Select **Configuration** > **Main**.



4. Modify account information, as necessary.

   If you change the storage class, new object data is stored with the new storage class. Existing object continue to be stored under the storage class set when ingested.

   Bucket Name, Region, and Endpoint, use AWS values and cannot be changed.

5. Click **Apply Changes**.

Modifying the Cloud Tiering Service state

You can control the Archive Node's ability read and write to the targeted external archival storage system that connects through the S3 API by changing the state of the Cloud Tiering Service.

What you'll need

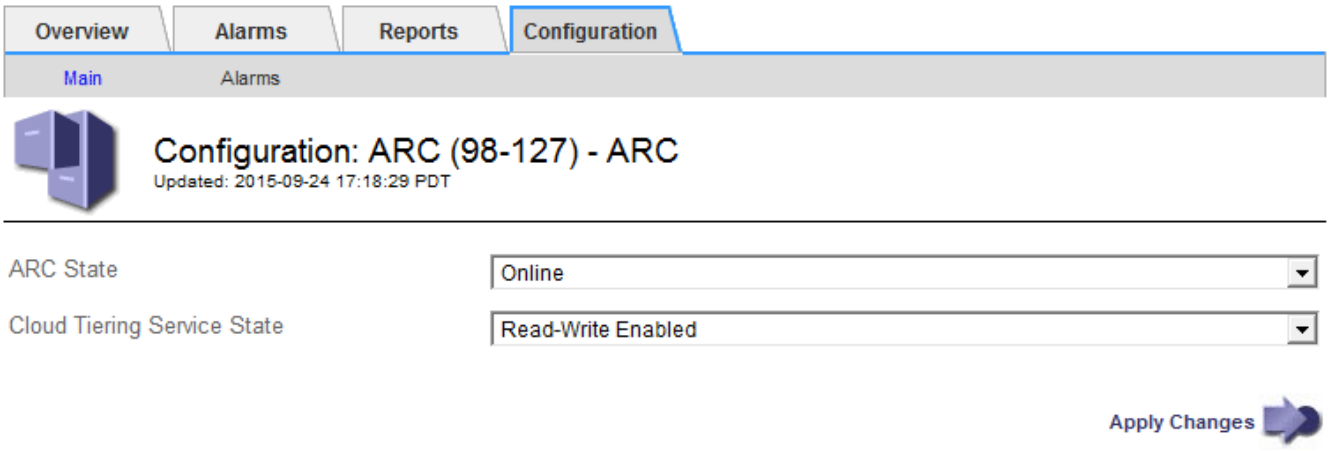
- You must be signed in to the Grid Manager using a supported browser.
- You must have specific access permissions.
- The Archive Node must be configured.

About this task

You can effectively take the Archive Node offline by changing the Cloud Tiering Service State to **Read-Write Disabled**.

Steps

1. Select **Support** > **Tools** > **Grid Topology**.
2. Select ***Archive Node*** > **ARC**.
3. Select **Configuration** > **Main**.



4. Select a **Cloud Tiering Service State**.
5. Click **Apply Changes**.

Resetting the Store Failure Count for S3 API connection

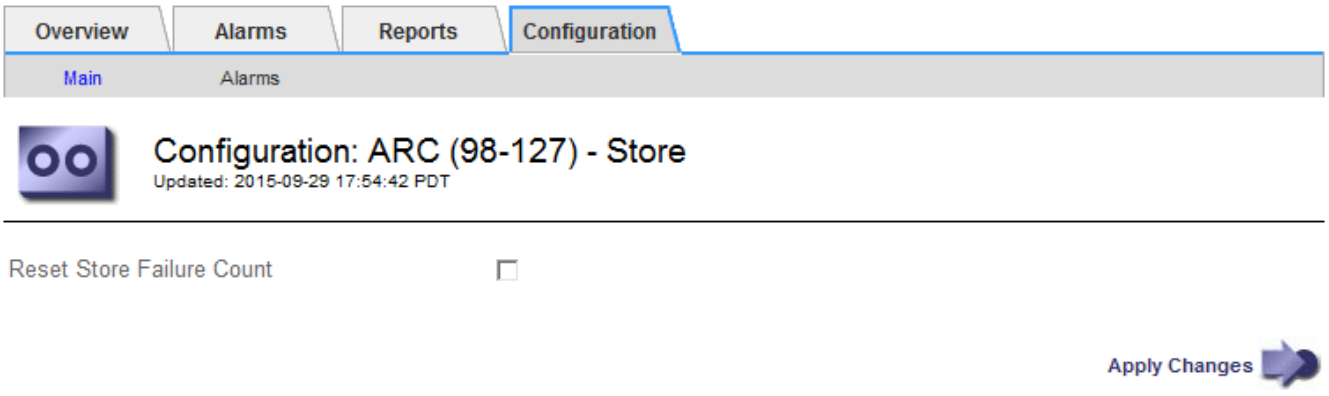
If your Archive Node connects to an archival storage system through the S3 API, you can reset the Store Failure Count, which can be used to clear the ARVF (Store Failures) alarm.

What you'll need

- You must be signed in to the Grid Manager using a supported browser.
- You must have specific access permissions.

Steps

1. Select **Support** > **Tools** > **Grid Topology**.
2. Select ***Archive Node*** > **ARC** > **Store**.
3. Select **Configuration** > **Main**.



4. Select **Reset Store Failure Count**.
5. Click **Apply Changes**.

The Store Failures attribute resets to zero.

Migrating objects from Cloud Tiering - S3 to a Cloud Storage Pool

If you are currently using the **Cloud Tiering - Simple Storage Service (S3)** feature to tier object data to an S3 bucket, consider migrating your objects to a Cloud Storage Pool instead. Cloud Storage Pools provide a scalable approach that takes advantage of all of the Storage Nodes in your StorageGRID system.

What you'll need

- You must be signed in to the Grid Manager using a supported browser.
- You must have specific access permissions.
- You have already stored objects in the S3 bucket configured for Cloud Tiering.

> Before migrating object data, contact your NetApp account representative to understand and manage any associated costs.

About this task

From an ILM perspective, a Cloud Storage Pool is similar to a storage pool. However, while storage pools consist of Storage Nodes or Archive Nodes within the StorageGRID system, a Cloud Storage Pool consists of an external S3 bucket.

Before migrating objects from Cloud Tiering - S3 to a Cloud Storage Pool, you must first create an S3 bucket and then create the Cloud Storage Pool in StorageGRID. Then, you can create a new ILM policy and replace the ILM rule used to store objects in the Cloud Tiering bucket with a cloned ILM rule that stores the same objects in the Cloud Storage Pool.

> When objects are stored in a Cloud Storage Pool, copies of those objects cannot also be stored within StorageGRID. If the ILM rule you are currently using for Cloud Tiering is configured to store objects in multiple locations at the same time, consider whether you still want to perform this optional migration because you will lose that functionality. If you continue with this migration, you must create new rules instead of cloning the existing ones.

Steps

1. Create a Cloud Storage Pool.

   Use a new S3 bucket for the Cloud Storage Pool to ensure it contains only the data managed by the Cloud Storage Pool.

2. Locate any ILM rules in the active ILM policy that cause objects to be stored in the Cloud Tiering bucket.
3. Clone each of these rules.
4. In the cloned rules, change the placement location to the new Cloud Storage Pool.
5. Save the cloned rules.
6. Create a new policy that uses the new rules.
7. Simulate and activate the new policy.

   When the new policy is activated and ILM evaluation occurs, the objects are moved from the S3 bucket configured for Cloud Tiering to the S3 bucket configured for the Cloud Storage Pool. The usable space on

the grid is not affected. After the objects are moved to the Cloud Storage Pool, they are removed from the Cloud Tiering bucket.

Related information

Manage objects with ILM

Archiving to tape through TSM middleware

You can configure an Archive Node to target a Tivoli Storage Manager (TSM) server that provides a logical interface for storing and retrieving object data to random or sequential access storage devices, including tape libraries.

The Archive Node's ARC service acts as a client to the TSM server, using Tivoli Storage Manager as middleware for communicating with the archival storage system.

TSM management classes

Management classes defined by the TSM middleware outline how the TSM's backup and archive operations function, and can be used to specify rules for content that are applied by the TSM server. Such rules operate independently of the StorageGRID system's ILM policy, and must be consistent with the StorageGRID system's requirement that objects are stored permanently and are always available for retrieval by the Archive Node. After object data is sent to a TSM server by the Archive Node, the TSM lifecycle and retention rules are applied while the object data is stored to tape managed by the TSM server.

The TSM management class is used by the TSM server to apply rules for data location or retention after objects are sent to the TSM server by the Archive Node. For example, objects identified as database backups (temporary content that can be overwritten with newer data) could be treated differently than application data (fixed content that must be retained indefinitely).

Configuring connections to TSM middleware

Before the Archive Node can communicate with Tivoli Storage Manager (TSM) middleware, you must configure a number of settings.

What you'll need

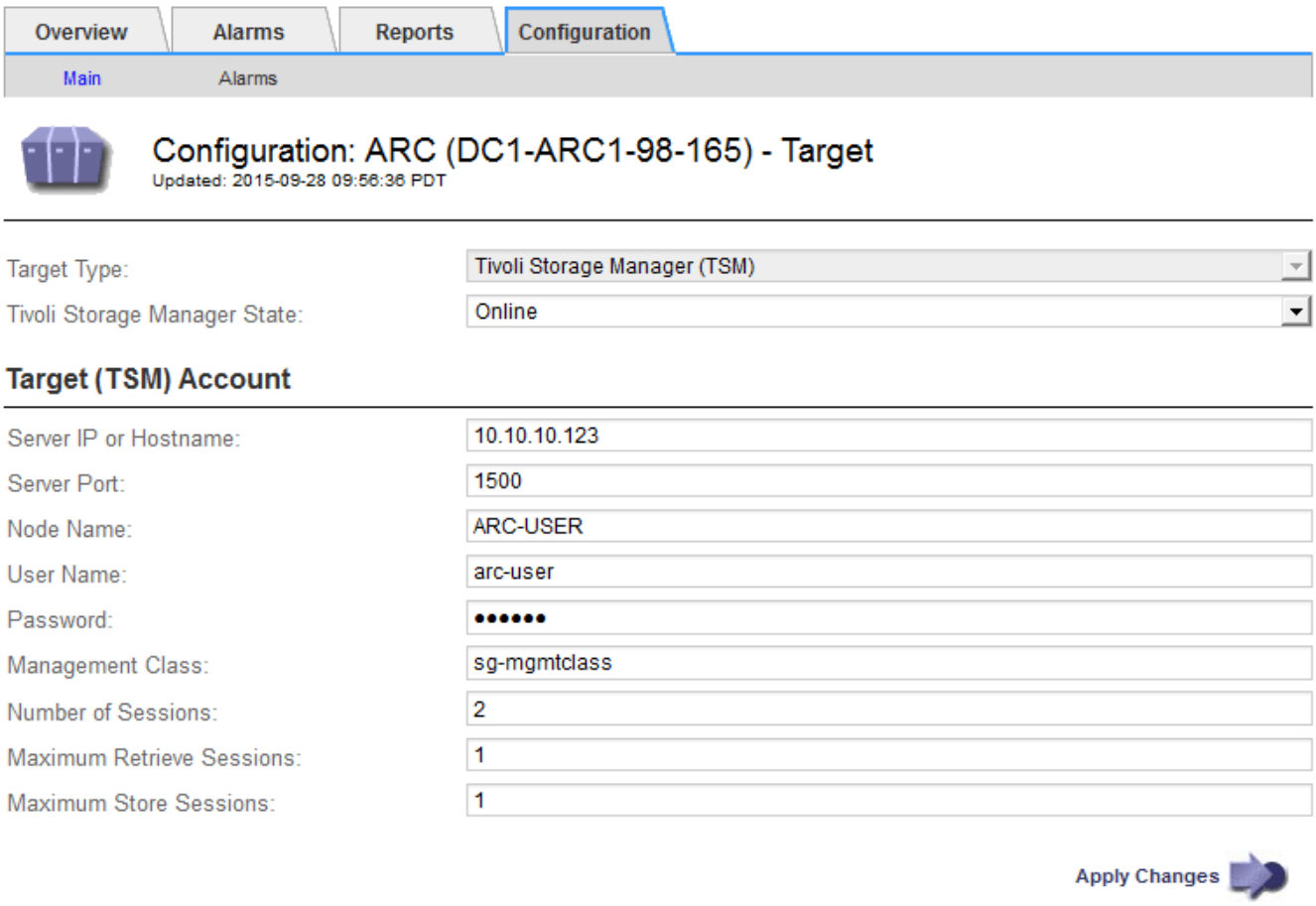
- You must be signed in to the Grid Manager using a supported browser.
- You must have specific access permissions.

About this task

Until these settings are configured, the ARC service remains in a Major alarm state as it is unable to communicate with the Tivoli Storage Manager.

Steps

1. Select **Support** > **Tools** > **Grid Topology**.
2. Select ***Archive Node*** > **ARC** > **Target**.
3. Select **Configuration** > **Main**.



4. From the **Target Type** drop-down list, select **Tivoli Storage Manager (TSM)**.
5. For the **Tivoli Storage Manager State**, select **Offline** to prevent retrievals from the TSM middleware server.

   By default, the Tivoli Storage Manager State is set to Online, which means that the Archive Node is able to retrieve object data from the TSM middleware server.

6. Complete the following information:
   - **Server IP or Hostname**: Specify the IP address or fully qualified domain name of the TSM middleware server used by the ARC service. The default IP address is 127.0.0.1.
   - **Server Port**: Specify the port number on the TSM middleware server that the ARC service will connect to. The default is 1500.
   - **Node Name**: Specify the name of the Archive Node. You must enter the name (arc-user) that you registered on the TSM middleware server.
   - **User Name**: Specify the user name the ARC service uses to log in to the TSM server. Enter the default user name (arc-user) or the administrative user you specified for the Archive Node.
   - **Password**: Specify the password used by the ARC service to log in to the TSM server.
   - **Management Class**: Specify the default management class to use if a management class is not specified when the object is being saved to the StorageGRID system, or the specified management class is not defined on the TSM middleware server.
   - **Number of Sessions**: Specify the number of tape drives on the TSM middleware server that are dedicated to the Archive Node. The Archive Node concurrently creates a maximum of one session per mount point plus a small number of additional sessions (less than five).

You must change this value to be the same as the value set for MAXNUMMP (maximum number of mount points) when the Archive Node was registered or updated. (In the register command, the default value of MAXNUMMP used is 1, if no value is set.)

You must also change the value of MAXSESSIONS for the TSM server to a number that is at least as large as the Number of Sessions set for the ARC service. The default value of MAXSESSIONS on the TSM server is 25.

- **Maximum Retrieve Sessions**: Specify the maximum number of sessions that the ARC service can open to the TSM middleware server for retrieve operations. In most cases, the appropriate value is Number of Sessions minus Maximum Store Sessions. If you need to share one tape drive for storage and retrieval, specify a value equal to the Number of Sessions.
- **Maximum Store Sessions**: Specify the maximum number of concurrent sessions that the ARC service can open to the TSM middleware server for archive operations.

  This value should be set to one except when the targeted archival storage system is full and only retrievals can be performed. Set this value to zero to use all sessions for retrievals.

7. Click **Apply Changes**.

Optimizing an Archive Node for TSM middleware sessions

You can optimize the performance of an Archive Node that connects to Tivoli Server Manager (TSM) by configuring the Archive Node's sessions.

What you'll need

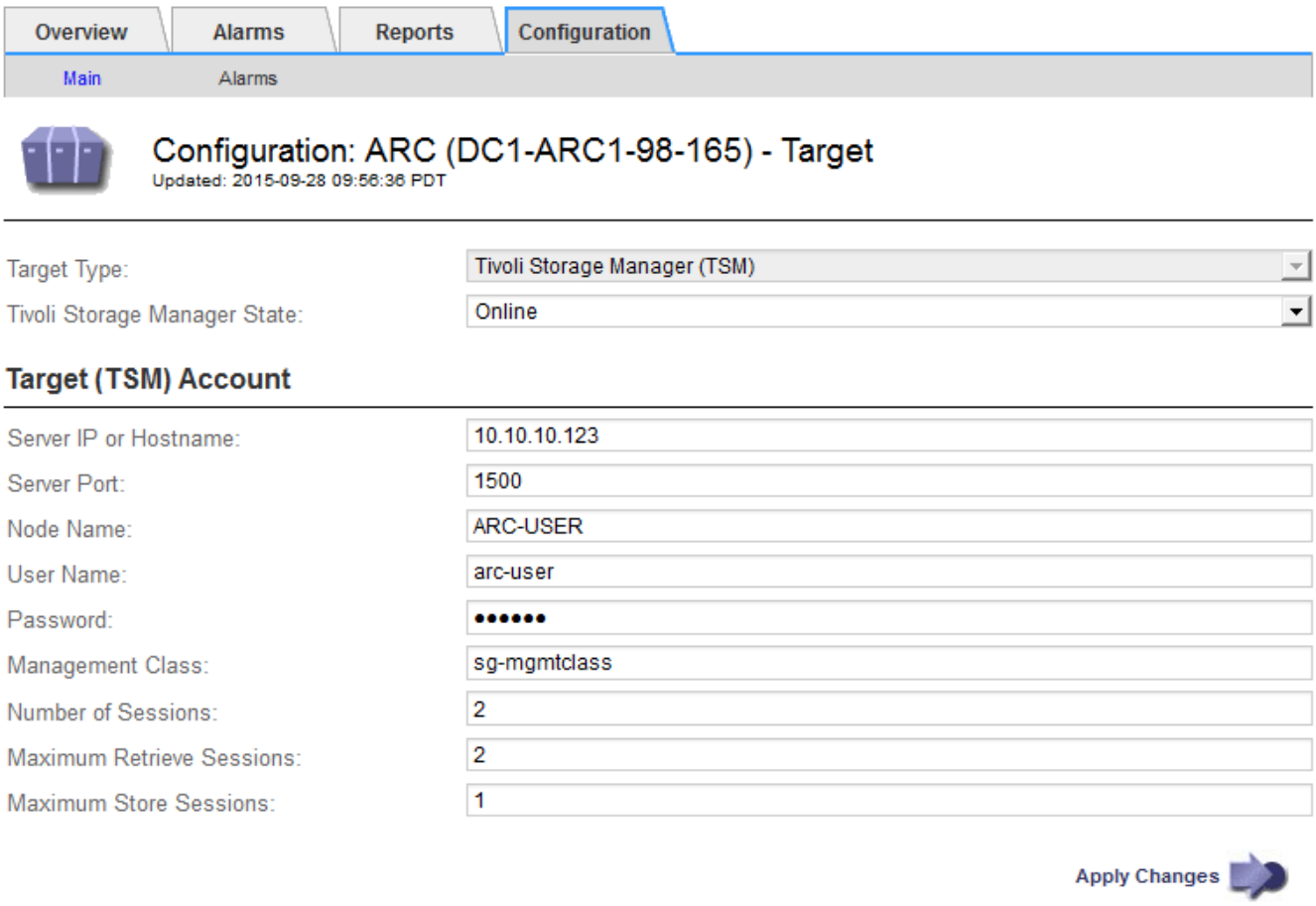
- You must be signed in to the Grid Manager using a supported browser.
- You must have specific access permissions.

About this task

Typically, the number of concurrent sessions that the Archive Node has open to the TSM middleware server is set to the number of tape drives the TSM server has dedicated to the Archive Node. One tape drive is allocated for storage while the rest are allocated for retrieval. However, in situations where a Storage Node is being rebuilt from Archive Node copies or the Archive Node is operating in Read-only mode, you can optimize TSM server performance by setting the maximum number of retrieve sessions to be the same as number of concurrent sessions. The result is that all drives can be used concurrently for retrieval, and, at most, one of these drives can also be used for storage if applicable.

Steps

1. Select **Support** > **Tools** > **Grid Topology**.
2. Select ***Archive Node*** > **ARC** > **Target**.
3. Select **Configuration** > **Main**.
4. Change **Maximum Retrieve Sessions** to be the same as **Number of Sessions**.



5. Click **Apply Changes**.

Configuring the archive state and counters for TSM

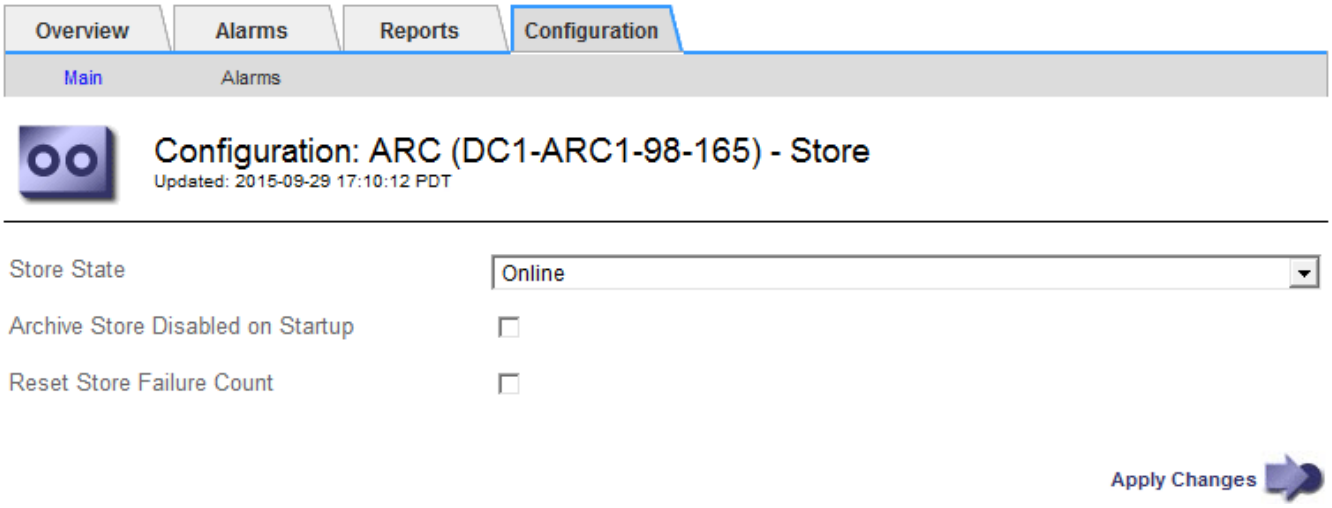
If your Archive Node connects to a TSM middleware server, you can configure an Archive Node's archive store state to Online or Offline. You can also disable the archive store when the Archive Node first starts up, or reset the failure count being tracked for the associated alarm.

What you'll need

- You must be signed in to the Grid Manager using a supported browser.
- You must have specific access permissions.

Steps

1. Select **Support** > **Tools** > **Grid Topology**.
2. Select ***Archive Node*** > **ARC** > **Store**.
3. Select **Configuration** > **Main**.



4. Modify the following settings, as necessary:
   - Store State: Set the component state to either:
     - Online: The Archive Node is available to process object data for storage to the archival storage system.
     - Offline: The Archive Node is not available to process object data for storage to the archival storage system.
   - Archive Store Disabled on Startup: When selected, the Archive Store component remains in the Read-only state when restarted. Used to persistently disable storage to the targeted the archival storage system. Useful when the targeted the archival storage system is unable to accept content.
   - Reset Store Failure Count: Reset the counter for store failures. This can be used to clear the ARVF (Stores Failure) alarm.
5. Click **Apply Changes**.

Related information

Managing an Archive Node when TSM server reaches capacity

Managing an Archive Node when TSM server reaches capacity

The TSM server has no way to notify the Archive Node when either the TSM database or the archival media storage managed by the TSM server is nearing capacity. The Archive Node continues to accept object data for transfer to the TSM server after the TSM server stops accepting new content. This content cannot be written to media managed by the TSM server. An alarm is triggered if this happens. This situation can be avoided through proactive monitoring of the TSM server.

What you'll need

- You must be signed in to the Grid Manager using a supported browser.
- You must have specific access permissions.

About this task

To prevent the ARC service from sending further content to the TSM server, you can take the Archive Node offline by taking its **ARC** > **Store** component offline. This procedure can also be useful in preventing alarms when the TSM server is unavailable for maintenance.

Steps

1. Select **Support** > **Tools** > **Grid Topology**.
2. Select ***Archive Node*** > **ARC** > **Store**.
3. Select **Configuration** > **Main**.

4. Change **Store State** to `Offline`.
5. Select **Archive Store Disabled on Startup**.
6. Click **Apply Changes**.

Setting Archive Node to read-only if TSM middleware reaches capacity

If the targeted TSM middleware server reaches capacity, the Archive Node can be optimized to only perform retrievals.

What you'll need

- You must be signed in to the Grid Manager using a supported browser.
- You must have specific access permissions.

Steps

1. Select **Support** > **Tools** > **Grid Topology**.
2. Select ***Archive Node*** > **ARC** > **Target**.
3. Select **Configuration** > **Main**.
4. Change Maximum Retrieve Sessions to be the same as the number of concurrent sessions listed in Number of Sessions.
5. Change Maximum Store Sessions to 0.

> Changing Maximum Store Sessions to 0 is not necessary if the Archive Node is Read-only. Store sessions will not be created.

6. Click **Apply Changes**.

Configuring Archive Node retrieve settings

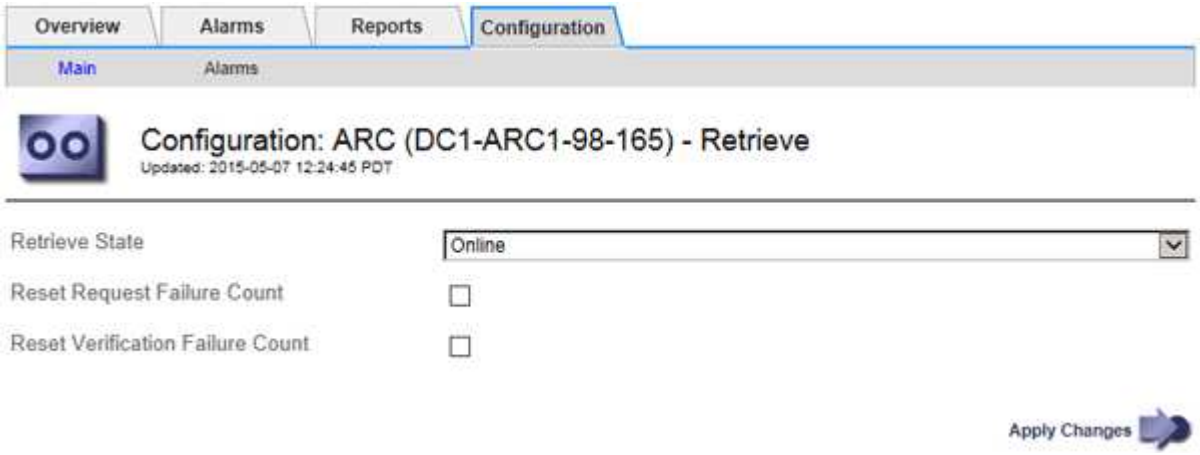
You can configure the retrieve settings for an Archive Node to set the state to Online or Offline, or reset the failure counts being tracked for the associated alarms.

What you'll need

- You must be signed in to the Grid Manager using a supported browser.
- You must have specific access permissions.

Steps

1. Select **Support** > **Tools** > **Grid Topology**.
2. Select **Archive Node** > **ARC** > **Retrieve**.
3. Select **Configuration** > **Main**.



4. Modify the following settings, as necessary:
   - **Retrieve State**: Set the component state to either:
     - Online: The grid node is available to retrieve object data from the archival media device.
     - Offline: The grid node is not available to retrieve object data.
   - Reset Request Failures Count: Select the check box to reset the counter for request failures. This can be used to clear the ARRF (Request Failures) alarm.
   - Reset Verification Failure Count: Select the check box to reset the counter for verification failures on retrieved object data. This can be used to clear the ARRV (Verification Failures) alarm.
5. Click **Apply Changes**.

Configuring Archive Node replication

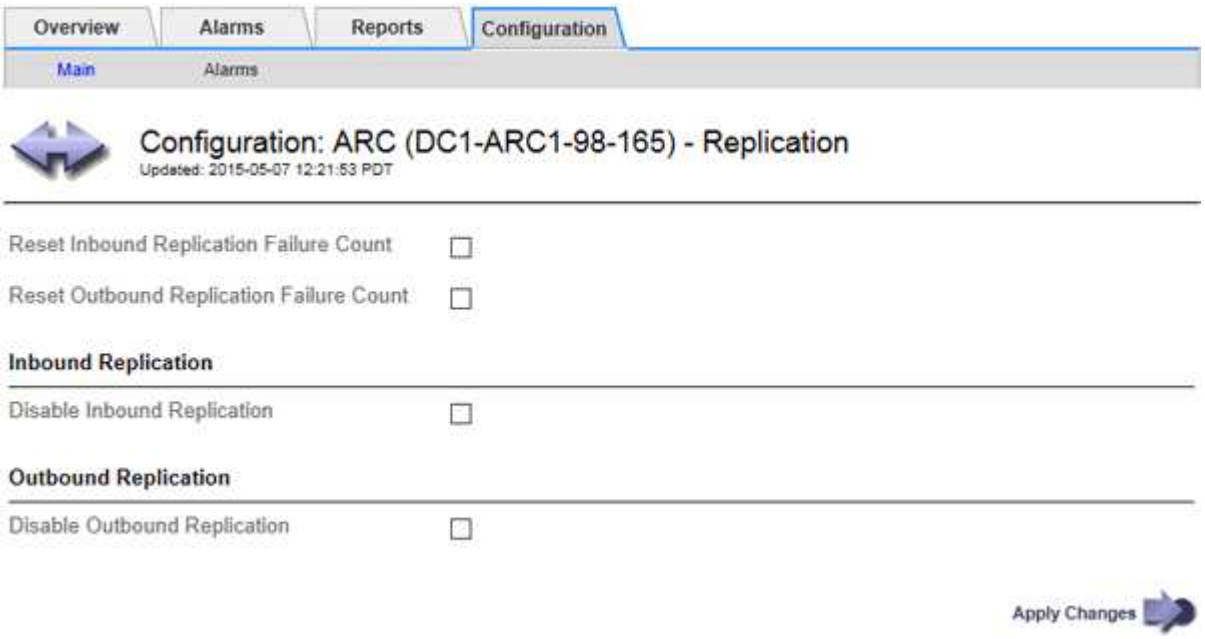
You can configure the replication settings for an Archive Node and disable inbound and outbound replication, or reset the failure counts being tracked for the associated alarms.

What you'll need

- You must be signed in to the Grid Manager using a supported browser.
- You must have specific access permissions.

Steps

1. Select **Support** > **Tools** > **Grid Topology**.
2. Select ***Archive Node*** > **ARC** > **Replication**.
3. Select **Configuration** > **Main**.



4. Modify the following settings, as necessary:
   - **Reset Inbound Replication Failure Count**: Select to reset the counter for inbound replication failures. This can be used to clear the RIRF (Inbound Replications — Failed) alarm.
   - **Reset Outbound Replication Failure Count**: Select to reset the counter for outbound replication failures. This can be used to clear the RORF (Outbound Replications — Failed) alarm.
   - **Disable Inbound Replication**: Select to disable inbound replication as part of a maintenance or testing procedure. Leave cleared during normal operation.

     When inbound replication is disabled, object data can be retrieved from the ARC service for replication to other locations in the StorageGRID system, but objects cannot be replicated to this ARC service from other system locations. The ARC service is read-only.

   - **Disable Outbound Replication**: Select the check box to disable outbound replication (including content requests for HTTP retrievals) as part of a maintenance or testing procedure. Leave unchecked during normal operation.

     When outbound replication is disabled, object data can be copied to this ARC service to satisfy ILM rules, but object data cannot be retrieved from the ARC service to be copied to other locations in the StorageGRID system. The ARC service is write-only.

5. Click **Apply Changes**.

Copyright information

Trademark information