Configure expanded system

StorageGRID 11.8

NetApp
March 19, 2024

Table of Contents

Configure expanded system . . . 1
- Configuration steps after expansion . . . 1
- Verify that Storage Node is active . . . 2
- Copy Admin Node database . . . 2
- Copy Prometheus metrics . . . 4
- Copy audit logs . . . 5
- Rebalance erasure-coded data after adding Storage Nodes . . . 7

Configure expanded system

Configuration steps after expansion

After completing an expansion, you must perform additional integration and configuration steps.

About this task

You must complete the configuration tasks listed below for the grid nodes or sites you are adding in your expansion. Some tasks might be optional, depending on the options selected when installing and administering your system, and how you want to configure the nodes and sites added during the expansion.

Steps

1. If you added a site:
   - Create a storage pool for the site and each storage grade you selected for the new Storage Nodes.
   - Confirm that the ILM policy meets the new requirements. If rule changes are required, create new rules and update the ILM policy. If the rules are already correct, activate a new policy with no rule changes to ensure StorageGRID uses the new nodes.
   - Confirm that Network Time Protocol (NTP) servers are accessible from that site. See Manage NTP servers.

   > Make sure that at least two nodes at each site can access at least four external NTP sources. If only one node at a site can reach the NTP sources, timing issues will occur if that node goes down. In addition, designating two nodes per site as primary NTP sources ensures accurate timing if a site is isolated from the rest of the grid.

2. If you added one or more Storage Nodes to an existing site:
   - View storage pool details to confirm that each node you added is included in the expected storage pools and used in the expected ILM rules.
   - Confirm that the ILM policy meets the new requirements. If rule changes are required, create new rules and update the ILM policy. If the rules are already correct, activate a new policy with no rule changes to ensure StorageGRID uses the new nodes.
   - Verify that the Storage Node is active and able to ingest objects.
   - If you were unable to add the recommended number of Storage Nodes, rebalance erasure-coded data. See Rebalance erasure-coded data after adding Storage Nodes.
3. If you added a Gateway Node:
   - If high availability (HA) groups are used for client connections, optionally add the Gateway Node to an HA group. Select **CONFIGURATION** > **Network** > **High availability groups** to review the list of existing HA groups and to add the new node. See Configure high availability groups.
4. If you added an Admin Node:
   a. If single sign-on (SSO) is enabled for your StorageGRID system, create a relying party trust for the new Admin Node. You can't sign in to the node until you create this relying party trust. See Configure single sign-on.
   b. If you plan to use the Load Balancer service on Admin Nodes, optionally add the new Admin Node to an HA group. Select **CONFIGURATION** > **Network** > **High availability groups** to review the list of existing HA groups and to add the new node. See Configure high availability groups.

c. Optionally, copy the Admin Node database from the primary Admin Node to the expansion Admin Node if you want to keep the attribute and audit information consistent on each Admin Node. See Copy the Admin Node database.

d. Optionally, copy the Prometheus database from the primary Admin Node to the expansion Admin Node if you want to keep the historical metrics consistent on each Admin Node. See Copy Prometheus metrics.

e. Optionally, copy the existing audit logs from the primary Admin Node to the expansion Admin Node if you want to keep the historical log information consistent on each Admin Node. See Copy audit logs.

5. To check if expansion nodes were added with an untrusted Client Network or to change whether a node's Client Network is untrusted or trusted, go to **CONFIGURATION** > **Security** > **Firewall control**.

   If the Client Network on the expansion node is untrusted, then connections to the node on the Client Network must be made using a load balancer endpoint. See Configure load balancer endpoints and Manage firewall controls.

6. Configure the DNS.

   If you have been specifying DNS settings separately for each grid node, you must add custom per-node DNS settings for the new nodes. See Modify DNS configuration for single grid node.

To ensure proper operation, specify two or three DNS servers. If you specify more than three, it is possible that only three will be used because of known OS limitations on some platforms. If you have routing restrictions in your environment, you can customize the DNS server list for individual nodes (typically all nodes at a site) to use a different set of up to three DNS servers.

If possible, use DNS servers that each site can access locally to ensure that an islanded site can resolve the FQDNs for external destinations.

Verify that Storage Node is active

After an expansion operation that adds new Storage Nodes completes, the StorageGRID system should automatically start using the new Storage Nodes. You must use the StorageGRID system to verify that the new Storage Node is active.

Steps

1. Sign in to the Grid Manager using a supported web browser.
2. Select **NODES** > ***Expansion Storage Node*** > **Storage**.
3. Position your cursor over the **Storage Used - Object Data** graph to view the value for **Used**, which is the amount of the Total usable space that has been used for object data.
4. Verify that the value of **Used** is increasing as you move your cursor to the right on the graph.

Copy Admin Node database

When adding Admin Nodes through an expansion procedure, you can optionally copy the database from the primary Admin Node to the new Admin Node. Copying the database allows you to retain historical information about attributes, alerts, and alerts.

Before you begin

- You have completed the required expansion steps to add an Admin Node.
- You have the `Passwords.txt` file.
- You have the provisioning passphrase.

About this task

The StorageGRID software activation process creates an empty database for the NMS service on the expansion Admin Node. When the NMS service starts on the expansion Admin Node, it records information for servers and services that are currently part of the system or added later. This Admin Node database includes the following information:

- Alert history
- Alarm history
- Historical attribute data, which is used in the charts and text reports available from the **SUPPORT** > **Tools** > **Grid topology** page

To ensure that the Admin Node database is consistent between nodes, you can copy the database from the primary Admin Node to the expansion Admin Node.

> Copying the database from the primary Admin Node (the*source Admin Node*) to an expansion Admin Node can take up to several hours to complete. During this period, the Grid Manager is inaccessible.

Use these steps to stop the MI service and the Management API service on both the primary Admin Node and the expansion Admin Node before copying the database.

Steps

1. Complete the following steps on the primary Admin Node:
   a. Log in to the Admin Node:
      i. Enter the following command: `ssh admin@grid_node_IP`
      ii. Enter the password listed in the `Passwords.txt` file.
      iii. Enter the following command to switch to root: `su -`
      iv. Enter the password listed in the `Passwords.txt` file.
   b. Run the following command: `recover-access-points`
   c. Enter the provisioning passphrase.
   d. Stop the MI service: `service mi stop`
   e. Stop the Management Application Program Interface (mgmt-api) service: `service mgmt-api stop`
2. Complete the following steps on the expansion Admin Node:
   a. Log in to the expansion Admin Node:
      i. Enter the following command: `ssh admin@grid_node_IP`
      ii. Enter the password listed in the `Passwords.txt` file.
      iii. Enter the following command to switch to root: `su -`
      iv. Enter the password listed in the `Passwords.txt` file.

b. Stop the MI service: `service mi stop`

c. Stop the mgmt-api service: `service mgmt-api stop`

d. Add the SSH private key to the SSH agent. Enter:`ssh-add`

e. Enter the SSH Access Password listed in the `Passwords.txt` file.

f. Copy the database from the source Admin Node to the expansion Admin Node: `/usr/local/mi/bin/mi-clone-db.sh` *`Source_Admin_Node_IP`*

g. When prompted, confirm that you want to overwrite the MI database on the expansion Admin Node.

   The database and its historical data are copied to the expansion Admin Node. When the copy operation is done, the script starts the expansion Admin Node.

h. When you no longer require passwordless access to other servers, remove the private key from the SSH agent. Enter:`ssh-add -D`

3. Restart the services on the primary Admin Node: `service servermanager start`

Copy Prometheus metrics

After adding a new Admin Node, you can optionally copy the historical metrics maintained by Prometheus from the primary Admin Node to the new Admin Node. Copying the metrics ensures that historical metrics are consistent between Admin Nodes.

Before you begin

- The new Admin Node is installed and running.
- You have the `Passwords.txt` file.
- You have the provisioning passphrase.

About this task

When you add an Admin Node, the software installation process creates a new Prometheus database. You can keep the historical metrics consistent between nodes by copying the Prometheus database from the primary Admin Node (the *source Admin Node*) to the new Admin Node.

> Copying the Prometheus database might take an hour or more. Some Grid Manager features will be unavailable while services are stopped on the source Admin Node.

Steps

1. Log in to the source Admin Node:
   a. Enter the following command: `ssh admin@`*`grid_node_IP`*
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.
2. From the source Admin Node, stop the Prometheus service: `service prometheus stop`
3. Complete the following steps on the new Admin Node:
   a. Log in to the new Admin Node:

i. Enter the following command: `ssh admin@grid_node_IP`

ii. Enter the password listed in the `Passwords.txt` file.

iii. Enter the following command to switch to root: `su -`

iv. Enter the password listed in the `Passwords.txt` file.

b. Stop the Prometheus service: `service prometheus stop`

c. Add the SSH private key to the SSH agent. Enter:`ssh-add`

d. Enter the SSH Access Password listed in the `Passwords.txt` file.

e. Copy the Prometheus database from the source Admin Node to the new Admin Node: `/usr/local/prometheus/bin/prometheus-clone-db.sh Source_Admin_Node_IP`

f. When prompted, press **Enter** to confirm that you want to destroy the new Prometheus database on the new Admin Node.

The original Prometheus database and its historical data are copied to the new Admin Node. When the copy operation is done, the script starts the new Admin Node. The following status appears:

```
Database cloned, starting services
```

g. When you no longer require passwordless access to other servers, remove the private key from the SSH agent. Enter:

```
ssh-add -D
```

4. Restart the Prometheus service on the source Admin Node.

```
service prometheus start
```

Copy audit logs

When you add a new Admin Node through an expansion procedure, its AMS service only logs events and actions that occur after it joins the system. As required, you can copy audit logs from a previously installed Admin Node to the new expansion Admin Node so that it is in sync with the rest of the StorageGRID system.

Before you begin

- You have completed the required expansion steps to add an Admin Node.
- You have the `Passwords.txt` file.

About this task

To make historical audit messages available on a new Admin Node, you must copy the audit log files manually from an existing Admin Node to the expansion Admin Node.

By default, audit information is sent to the audit log on Admin Nodes. You can skip these steps if either of the following applies:

- You configured an external syslog server and audit logs are now being sent to the syslog server instead of to Admin Nodes.
- You explicitly specified that audit messages should be saved only on the local nodes that generated them.

See Configure audit messages and log destinations for details.

Steps

1. Log in to the primary Admin Node:
   a. Enter the following command: `ssh admin@_primary_Admin_Node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.

   When you are logged in as root, the prompt changes from `$` to `#`.
2. Stop the AMS service to prevent it from creating a new file: `service ams stop`
3. Navigate to the audit export directory:

   ```
   cd /var/local/log
   ```

4. Rename the source `audit.log` file to ensure that it does not overwrite the file on the expansion Admin Node you are copying it to:

   ```
   ls -l
   mv audit.log _new_name_.txt
   ```

5. Copy all audit log files to the destination location on the expansion Admin Node:

   ```
   scp -p * IP_address:/var/local/log
   ```

6. If prompted for the passphrase for `/root/.ssh/id_rsa`, enter the SSH Access Password for the Primary Admin Node listed in the `Passwords.txt` file.
7. Restore the original `audit.log` file:

   ```
   mv new_name.txt audit.log
   ```

8. Start the AMS service:

   ```
   service ams start
   ```

9. Log out from the server:

   ```
   exit
   ```

10. Log in to the expansion Admin Node:
    a. Enter the following command: `ssh admin@expansion_Admin_Node_IP`
    b. Enter the password listed in the `Passwords.txt` file.
    c. Enter the following command to switch to root: `su -`
    d. Enter the password listed in the `Passwords.txt` file.

    When you are logged in as root, the prompt changes from `$` to `#`.

11. Update the user and group settings for the audit log files:

```
cd /var/local/log
chown ams-user:bycast *
```

12. Log out from the server:

```
exit
```

Rebalance erasure-coded data after adding Storage Nodes

After you add Storage Nodes, you can use the EC rebalance procedure to redistribute erasure-coded fragments among the existing and new Storage Nodes.

Before you begin

- You have completed the expansion steps to add the new Storage Nodes.
- You have reviewed the considerations for rebalancing erasure-coded data.
- You understand that replicated object data will not be moved by this procedure and that the EC rebalance procedure does not consider the replicated data usage on each Storage Node when determining where to move erasure-coded data.
- You have the `Passwords.txt` file.

What happens when this procedure runs

Before starting the procedure, note the following:

- The EC rebalance procedure will not start if one or more volumes are offline (unmounted) or if they are online (mounted) but in an error state.
- The EC rebalance procedure temporarily reserves a large amount of storage. Storage alerts might be triggered, but will resolve when the rebalance is complete. If there is not enough storage for the reservation, the EC rebalance procedure will fail. Storage reservations are released when the EC rebalance procedure completes, whether the procedure failed or succeeded.
- If a volume goes offline while the EC rebalance procedure is in process, the rebalance procedure will terminate. Any data fragments that were already moved will remain in their new locations, and no data will be lost.

  You can rerun the procedure after all volumes are back online.

- When the EC rebalance procedure is running, the performance of ILM operations and S3 and Swift client operations might be impacted.

S3 and Swift API operations to upload objects (or object parts) might fail during the EC rebalance procedure if they require more than 24 hours to complete. Long-duration PUT operations will fail if the applicable ILM rule uses Balanced or Strict placement on ingest. The following error will be reported: `500 Internal Server Error`.

- During this procedure all nodes have a storage capacity limit of 80%. Nodes that exceed this limit, but still store below the target data partition, are excluded from:
  - The site imbalance value
  - Any job completion conditions

  The target data partition is calculated by dividing the total data for a site by the number of nodes.

- **Job completion conditions**. The EC rebalance procedure is considered complete when any of the following is true:
  - It can't move any more erasure-coded data.
  - The data in all nodes is within a 5% deviation of the target data partition.
  - The procedure has been running for 30 days.

Steps

1. Review the current object storage details for the site you plan to rebalance.
   a. Select **NODES**.
   b. Select the first Storage Node at the site.
   c. Select the **Storage** tab.
   d. Position your cursor over the Storage Used - Object Data chart to see the current amount of replicated data and erasure-coded data on the Storage Node.
   e. Repeat these steps to view the other Storage Nodes at the site.
2. Log in to the primary Admin Node:
   a. Enter the following command: `ssh admin@primary_Admin_Node_IP`
   b. Enter the password listed in the `Passwords.txt` file.
   c. Enter the following command to switch to root: `su -`
   d. Enter the password listed in the `Passwords.txt` file.

   When you are logged in as root, the prompt changes from `$` to `#`.
3. Start the procedure:

   `rebalance-data start --site "*site-name*"

   For "*site-name*", specify the first site where you added new Storage Node or nodes. Enclose `site-name` in quotes.

   The EC rebalance procedure starts, and a job ID is returned.
4. Copy the job ID.

5. Monitor the status of the EC rebalance procedure.
   - To view the status of a single EC rebalance procedure:

     `rebalance-data status --job-id job-id`

     For *job-id*, specify the ID that was returned when you started the procedure.

   - To view the status of the current EC rebalance procedure and any previously completed procedures:

     `rebalance-data status`

     > To get help on the rebalance-data command:
     >
     > `rebalance-data --help`

6. Perform additional steps, based on the status returned:
   - If `State` is `In progress`, the EC rebalance operation is still running. You should periodically monitor the procedure until it completes.

     Use the `Site Imbalance` value to assess how unbalanced erasure-code data usage is across the Storage Nodes at the site. This value can range from 1.0 to 0, with 0 indicating that erasure-coding data usage is completely balanced across all Storage Nodes at the site.

     The EC rebalance job is considered complete and will stop when the data in all nodes is within a 5% deviation of the target data partition.

   - If `State` is `Success`, optionally review object storage to see the updated details for the site.

     Erasure-coded data should now be more balanced among the Storage Nodes at the site.

   - If `State` is `Failure`:
     a. Confirm that all Storage Nodes at the site are connected to the grid.
     b. Check for and resolve any alerts that might be affecting these Storage Nodes.
     c. Restart the EC rebalance procedure:

        `rebalance-data start --job-id job-id`

     d. View the status of the new procedure. If `State` is still `Failure`, contact technical support.
7. If the EC rebalance procedure is generating too much load (for example, ingest operations are affected), pause the procedure.

   `rebalance-data pause --job-id job-id`

8. If you need to terminate the EC rebalance procedure (for example, so you can perform a StorageGRID software upgrade), enter the following:

   `rebalance-data terminate --job-id job-id`

   > When you terminate an EC rebalance procedure, any data fragments that have already been moved remain in their new locations. Data is not moved back to the original location.

9. If you are using erasure coding at more than one site, run this procedure for all other affected sites.

Copyright information

Trademark information