■ NetApp

Manage nodes

ONTAP 9

NetApp November 01, 2023

This PDF was generated from https://docs.netapp.com/us-en/ontap/system-admin/add-nodes-cluster-concept.html on November 01, 2023. Always check docs.netapp.com for the latest.

Table of Contents

nage nodes	. 1
Add nodes to the cluster	. 1
Remove nodes from the cluster	. 3
Access a node's log, core dump, and MIB files by using a web browser	. 5
Access the system console of a node	. 7
Manage node root volumes and root aggregates	. 8
Start or stop a node overview	. 11
Manage a node by using the boot menu	14

Manage nodes

Add nodes to the cluster

After a cluster is created, you can expand it by adding nodes to it. You add only one node at a time.

What you'll need

- If you are adding nodes to a multiple-node cluster, more than half of the existing nodes in the cluster must be healthy (indicated by cluster show).
- If you are adding nodes to a two-node switchless cluster, you must have installed and configured the cluster management and interconnect switches before adding additional nodes.

The switchless cluster functionality is supported only in a two-node cluster.

When a cluster contains or grows to more than two nodes, cluster HA is not required and is disabled automatically.

- If you are adding a second node to a single-node cluster, the second node must have been installed, and the cluster network must have been configured.
- If the cluster has the SP automatic configuration enabled, the subnet specified for the SP to use must have available resources for the joining node.

A node that joins the cluster uses the specified subnet to perform automatic configuration for the SP.

- You must have gathered the following information for the new node's node management LIF:
 - Port
 - IP address
 - Netmask
 - Default gateway

About this task

Nodes must be in even numbers so that they can form HA pairs. After you start to add a node to the cluster, you must complete the process. The node must be part of the cluster before you can start to add another node.

Steps

1. Power on the node that you want to add to the cluster.

The node boots, and the Node Setup wizard starts on the console.

```
Welcome to node setup.

You can enter the following commands at any time:
    "help" or "?" - if you want to have a question clarified,
    "back" - if you want to change previously answered questions, and
    "exit" or "quit" - if you want to quit the setup wizard.
        Any changes you made before quitting will be saved.

To accept a default or omit a question, do not enter a value.

Enter the node management interface port [eOc]:
```

2. Exit the Node Setup wizard: exit

The Node Setup wizard exits, and a login prompt appears, warning that you have not completed the setup tasks.

- 3. Log in to the admin account by using the admin user name.
- 4. Start the Cluster Setup wizard:

cluster setup

```
::> cluster setup

Welcome to the cluster setup wizard.

You can enter the following commands at any time:
  "help" or "?" - if you want to have a question clarified,
  "back" - if you want to change previously answered questions, and
  "exit" or "quit" - if you want to quit the cluster setup wizard.
  Any changes you made before quitting will be saved.

You can return to cluster setup at any time by typing "cluster setup".
  To accept a default or omit a question, do not enter a value....

Use your web browser to complete cluster setup by accessing https://lo.63.11.29

Otherwise, press Enter to complete cluster setup using the command line interface:
```



For more information on setting up a cluster using the setup GUI, see the System Manager online help.

5. Press Enter to use the CLI to complete this task. When prompted to create a new cluster or join an existing

one, enter join.

```
Do you want to create a new cluster or join an existing cluster? {create, join}: join
```

- 6. Follow the prompts to set up the node and join it to the cluster:
 - To accept the default value for a prompt, press Enter.
 - To enter your own value for a prompt, enter the value, and then press Enter.
- 7. Repeat the preceding steps for each additional node that you want to add.

After you finish

After adding nodes to the cluster, you should enable storage failover for each HA pair.

Remove nodes from the cluster

You can remove unwanted nodes from a cluster, one node at a time. After you remove a node, you must also remove its failover partner. If you are removing a node, then its data becomes inaccessible or erased.

Before you begin

The following conditions must be satisfied before removing nodes from the cluster:

- More than half of the nodes in the cluster must be healthy.
- All of the data on the node that you want to remove must have been evacuated.
 - This might include purging data from an encrypted volume.
- All non-root volumes have been moved from aggregates owned by the node.
- All non-root aggregates have been deleted from the node.
- If the node owns Federal Information Processing Standards (FIPS) disks or self-encrypting disks (SEDs), disk encryption has been removed by returning the disks to unprotected mode.
 - You might also want to sanitize FIPS drives or SEDs.
- Data LIFs have been deleted or relocated from the node.
- Cluster management LIFs have been relocated from the node and the home ports changed.
- All intercluster LIFs have been removed.
 - When you remove intercluster LIFs a warning is displayed that can be ignored.
- Storage failover has been disabled for the node.
- All LIF failover rules have been modified to remove ports on the node.
- All VLANs on the node have been deleted.
- If you have LUNs on the node to be removed, you should modify the Selective LUN Map (SLM) reportingnodes list before you remove the node.

If you do not remove the node and its HA partner from the SLM reporting-nodes list, access to the LUNs previously on the node can be lost even though the volumes containing the LUNs were moved to another

node.

It is recommended that you issue an AutoSupport message to notify NetApp technical support that node removal is underway.

Note: You must not perform operations such as cluster remove-node, cluster unjoin, and node rename when an automated ONTAP upgrade is in progress.

About this task

- If you are running a mixed-version cluster, you can remove the last low-version node by using one of the advanced privilege commands beginning with ONTAP 9.3:
 - o ONTAP 9.3: cluster unjoin -skip-last-low-version-node-check
 - ° ONTAP 9.4 and later: cluster remove-node -skip-last-low-version-node-check
- If you unjoin 2 nodes from a 4-node cluster, cluster HA is automatically enabled on the two remaining nodes.



All system and user data, from all disks that are connected to the node, must be made inaccessible to users before removing a node from the cluster. If a node was incorrectly unjoined from a cluster, contact NetApp Support for assistance with options for recovery.

Steps

1. Change the privilege level to advanced:

set -privilege advanced

2. Verify if a nodes on the cluster holds epsilon:

```
cluster show -epsilon true
```

a. If a node holds epsilon, change its eligibility to false.

```
cluster modify -node <node name> -eligibility false
```

3. If the node you want to remove is the current master node, then enable another node in the cluster to be elected as the master node by changing the master node's cluster eligibility to false:

```
cluster modify -eligibility false
```

The master node is the node that holds processes such as "mgmt", "vldb", "vifmgr", "bcomd", and "crs". The cluster ring show advanced command shows the current master node.

```
cluster::*> cluster modify -node node1 -eligibility false
```

- 4. Log into the remote node management LIF or the cluster-management LIF on a node other than the one that is being removed.
- 5. Remove the node from the cluster:

For this ONTAP version	Use this command
ONTAP 9.3	cluster unjoin
ONTAP 9.4 and later	cluster remove-node

If you have a mixed version cluster and you are removing the last lower version node, use the -skip -last-low-version-node-check parameter with these commands.

The system informs you of the following:

- You must also remove the node's failover partner from the cluster.
- After the node is removed and before it can rejoin a cluster, you must use boot menu option (4) Clean configuration and initialize all disks or option (9) Configure Advanced Drive Partitioning to erase the node's configuration and initialize all disks.

A failure message is generated if you have conditions that you must address before removing the node. For example, the message might indicate that the node has shared resources that you must remove or that the node is in a cluster HA configuration or storage failover configuration that you must disable.

If the node is the quorum master, the cluster will briefly lose and then return to quorum. This quorum loss is temporary and does not affect any data operations.

6. If a failure message indicates error conditions, address those conditions and rerun the cluster removenode or cluster unjoin command.

The node is automatically rebooted after it is successfully removed from the cluster.

- 7. If you are repurposing the node, erase the node configuration and initialize all disks:
 - a. During the boot process, press Ctrl-C to display the boot menu when prompted to do so.
 - b. Select the boot menu option (4) Clean configuration and initialize all disks.
- 8. Return to admin privilege level:

set -privilege admin

9. Repeat the preceding steps to remove the failover partner from the cluster.

After you finish

If you removed nodes to have a single-node cluster, you should modify the cluster ports to serve data traffic by modifying the cluster ports to be data ports, and then creating data LIFs on the data ports.

Access a node's log, core dump, and MIB files by using a web browser

The Service Processor Infrastructure (spi) web service is enabled by default to enable a web browser to access the log, core dump, and MIB files of a node in the cluster. The files remain accessible even when the node is down, provided that the node is taken over by its partner.

What you'll need

The cluster management LIF must be up.

You can use the management LIF of the cluster or a node to access the spi web service. However, using the cluster management LIF is recommended.

The network interface show command displays the status of all LIFs in the cluster.

- You must use a local user account to access the spi web service, domain user accounts are not supported.
- If your user account does not have the "admin" role (which has access to the spi web service by default), your access-control role must be granted access to the spi web service.

The vserver services web access show command shows what roles are granted access to which web services.

• If you are not using the "admin" user account (which includes the http access method by default), your user account must be set up with the http access method.

The security login show command shows user accounts' access and login methods and their access-control roles.

 If you want to use HTTPS for secure web access, SSL must be enabled and a digital certificate must be installed.

The system services web show command displays the configuration of the web protocol engine at the cluster level.

About this task

The spi web service is enabled by default, and the service can be disabled manually (vserver services web modify -vserver * -name spi -enabled false).

The "admin" role is granted access to the spi web service by default, and the access can be disabled manually (services web access delete -vserver cluster name -name spi -role admin).

Steps

1. Point the web browser to the spi web service URL in one of the following formats:

```
  http://cluster-mgmt-LIF/spi/
  https://cluster-mgmt-LIF/spi/
  cluster-mgmt-LIF is the IP address of the cluster management LIF.
```

2. When prompted by the browser, enter your user account and password.

After your account is authenticated, the browser displays links to the /mroot/etc/log/, /mroot/etc/crash/, and /mroot/etc/mib/ directories of each node in the cluster.

Access the system console of a node

If a node is hanging at the boot menu or the boot environment prompt, you can access it only through the system console (also called the *serial console*). You can access the system console of a node from an SSH connection to the node's SP or to the cluster.

About this task

Both the SP and ONTAP offer commands that enable you to access the system console. However, from the SP, you can access only the system console of its own node. From the cluster, you can access the system console of any node in the cluster.

Steps

1. Access the system console of a node:

If you are in the	Enter this command
SP CLI of the node	system console
ONTAP CLI	system node run-console

- 2. Log in to the system console when you are prompted to do so.
- 3. To exit the system console, press Ctrl-D.

Examples of accessing the system console

The following example shows the result of entering the system console command at the "SP node2" prompt. The system console indicates that node2 is hanging at the boot environment prompt. The boot_ontap command is entered at the console to boot the node to ONTAP. Ctrl-D is then pressed to exit the console and return to the SP.

(Ctrl-D is pressed to exit the system console.)

```
Connection to 123.12.123.12 closed.
SP node2>
```

The following example shows the result of entering the system node run-console command from ONTAP to access the system console of node2, which is hanging at the boot environment prompt. The boot_ontap command is entered at the console to boot node2 to ONTAP. Ctrl-D is then pressed to exit the console and return to ONTAP.

(Ctrl-D is pressed to exit the system console.)

```
Connection to 123.12.123.12 closed. cluster1::>
```

Manage node root volumes and root aggregates

A node's root volume is a FlexVol volume that is installed at the factory or by setup software. It is reserved for system files, log files, and core files. The directory name is /mroot, which is accessible only through the systemshell by technical support. The minimum size for a node's root volume depends on the platform model.

Rules governing node root volumes and root aggregates overview

A node's root volume contains special directories and files for that node. The root aggregate contains the root volume. A few rules govern a node's root volume and root aggregate.

- The following rules govern the node's root volume:
 - Unless technical support instructs you to do so, do not modify the configuration or content of the root volume.

Do not store user data in the root volume.

Storing user data in the root volume increases the storage giveback time between nodes in an HA pair.

- You can move the root volume to another aggregate. See Relocate root volumes to new aggregates.
- The root aggregate is dedicated to the node's root volume only.

ONTAP prevents you from creating other volumes in the root aggregate.

NetApp Hardware Universe

Free up space on a node's root volume

A warning message appears when a node's root volume has become full or almost full. The node cannot operate properly when its root volume is full. You can free up space on a node's root volume by deleting core dump files, packet trace files, and root volume Snapshot copies.

Steps

1. Display the node's core dump files and their names:

```
system node coredump show
```

2. Delete unwanted core dump files from the node:

```
system node coredump delete
```

3. Access the nodeshell:

system node run -node nodename

nodename is the name of the node whose root volume space you want to free up.

4. Switch to the nodeshell advanced privilege level from the nodeshell:

priv set advanced

- 5. Display and delete the node's packet trace files through the nodeshell:
 - a. Display all files in the node's root volume:

```
ls /etc
```

b. If any packet trace files (*.trc) are in the node's root volume, delete them individually:

```
rm /etc/log/packet traces/file name.trc
```

- 6. Identify and delete the node's root volume Snapshot copies through the nodeshell:
 - a. Identify the root volume name:

vol status

The root volume is indicated by the word "root" in the "Options" column of the vol status command output.

In the following example, the root volume is vol0:

node1*> vol status

Volume State vol0 online

Status raid dp, flex

Options root, nvfail=on

64-bit

b. Display root volume Snapshot copies:

```
snap list root_vol_name
```

c. Delete unwanted root volume Snapshot copies:

```
snap delete root_vol_namesnapshot_name
```

7. Exit the nodeshell and return to the clustershell:

exit

Relocate root volumes to new aggregates

The root replacement procedure migrates the current root aggregate to another set of disks without disruption.

About this task

Storage failover must be enabled to relocate root volumes. You can use the storage failover modify -node nodename -enable true command to enable failover.

You can change the location of the root volume to a new aggregate in the following scenarios:

- · When the root aggregates are not on the disk you prefer
- When you want to rearrange the disks connected to the node
- When you are performing a shelf replacement of the EOS disk shelves

Steps

1. Set the privilege level to advanced:

```
set privilege advanced
```

2. Relocate the root aggregate:

system node migrate-root -node nodename -disklist disklist -raid-type raidtype

• -node

Specifies the node that owns the root aggregate that you want to migrate.

· -disklist

Specifies the list of disks on which the new root aggregate will be created. All disks must be spares and owned by the same node. The minimum number of disks required is dependent on the RAID type.

· -raid-type

Specifies the RAID type of the root aggregate. The default value is raid-dp.

3. Monitor the progress of the job:

```
job show -id jobid -instance
```

Results

If all of the pre-checks are successful, the command starts a root volume replacement job and exits. Expect the node to restart.

Start or stop a node overview

You might need to start or stop a node for maintenance or troubleshooting reasons. You can do so from the ONTAP CLI, the boot environment prompt, or the SP CLI.

Using the SP CLI command system power off or system power cycle to turn off or power-cycle a node might cause an improper shutdown of the node (also called a *dirty shutdown*) and is not a substitute for a graceful shutdown using the ONTAP system node halt command.

Reboot a node at the system prompt

You can reboot a node in normal mode from the system prompt. A node is configured to boot from the boot device, such as a PC CompactFlash card.

Steps

- 1. If the cluster contains four or more nodes, verify that the node to be rebooted does not hold epsilon:
 - a. Set the privilege level to advanced:

set -privilege advanced

b. Determine which node holds epsilon:

cluster show

The following example shows that "node1" holds epsilon:

cluster1::*> cluster	show		
Node	Health	Eligibility	Epsilon
node1	true	true	true
node2	true	true	false
node3	true	true	false
node4	true	true	false
4 entries were displ	ayed.		

c. If the node to be rebooted holds epsilon, then remove epsilon from the node:

```
cluster modify -node node name -epsilon false
```

d. Assign epsilon to a different node that will remain up:

```
cluster modify -node node name -epsilon true
```

e. Return to the admin privilege level:

```
set -privilege admin
```

2. Use the system node reboot command to reboot the node.

If you do not specify the <code>-skip-lif-migration</code> parameter, the command attempts to migrate data and cluster management LIFs synchronously to another node prior to the reboot. If the LIF migration fails or times out, the rebooting process is aborted, and ONTAP displays an error to indicate the LIF migration failure.

```
cluster1::> system node reboot -node node1 -reason "software upgrade"
```

The node begins the reboot process. The ONTAP login prompt appears, indicating that the reboot process is complete.

Boot ONTAP at the boot environment prompt

You can boot the current release or the backup release of ONTAP when you are at the boot environment prompt of a node.

Steps

1. Access the boot environment prompt from the storage system prompt by using the system node halt command.

The storage system console displays the boot environment prompt.

2. At the boot environment prompt, enter one of the following commands:

To boot	Enter
The current release of ONTAP	boot_ontap
The ONTAP primary image from the boot device	boot_primary
The ONTAP backup image from the boot device	boot_backup

If you are unsure about which image to use, you should use boot ontap in the first instance.

Shut down a node

You can shut down a node if it becomes unresponsive or if support personnel direct you to do so as part of troubleshooting efforts.

Steps

- 1. If the cluster contains four or more nodes, verify that the node to be shut down does not hold epsilon:
 - a. Set the privilege level to advanced:

set -privilege advanced

b. Determine which node holds epsilon:

cluster show

The following example shows that "node1" holds epsilon:

```
cluster1::*> cluster show
Node
                Health Eligibility
                                 Epsilon
_______
node1
                true
                      true
                                 true
                true true
node2
                                 false
node3
                                 false
                true
                     true
node4
                true true
                                 false
4 entries were displayed.
```

c. If the node to be shut down holds epsilon, then remove epsilon from the node:

```
cluster modify -node node name -epsilon false
```

d. Assign epsilon to a different node that will remain up:

```
cluster modify -node node name -epsilon true
```

e. Return to the admin privilege level:

```
set -privilege admin
```

2. Use the system node halt command to shut down the node.

If you do not specify the <code>-skip-lif-migration</code> parameter, the command attempts to migrate data and cluster management LIFs synchronously to another node prior to the shutdown. If the LIF migration fails or times out, the shutdown process is aborted, and ONTAP displays an error to indicate the LIF migration failure.

You can manually trigger a core dump with the shutdown by using both the -dump parameter.

The following example shuts down the node named "node1" for hardware maintenance:

Manage a node by using the boot menu

You can use the boot menu to correct configuration problems on a node, reset the admin password, initialize disks, reset the node configuration, and restore the node configuration information back to the boot device.



If an HA pair is using encrypting SAS or NVMe drives (SED, NSE, FIPS), you must follow the instructions in the topic Returning a FIPS drive or SED to unprotected mode for all drives within the HA pair prior to initializing the system (boot options 4 or 9). Failure to do this may result in future data loss if the drives are repurposed.

Steps

1. Reboot the node to access the boot menu by using the system node reboot command at the system prompt.

The node begins the reboot process.

2. During the reboot process, press Ctrl-C to display the boot menu when prompted to do so.

The node displays the following options for the boot menu:

- (1) Normal Boot.
- (2) Boot without /etc/rc.
- (3) Change password.
- (4) Clean configuration and initialize all disks.
- (5) Maintenance mode boot.
- (6) Update flash from backup config.
- (7) Install new software first.
- (8) Reboot node.
- (9) Configure Advanced Drive Partitioning.
- (10) Set onboard key management recovery secrets.
- (11) Configure node for external key management.

Selection (1-11)?



Boot menu option (2) Boot without /etc/rc is obsolete and takes no effect on the system.

3. Select one of the following options by entering the corresponding number:

То	Select
Continue to boot the node in normal mode	1) Normal Boot

То	Select
Change the password of the node, which is also the "admin" account password	3) Change Password
Initialize the node's disks and create a root volume for the node	This menu option erases all data on the disks of the node and resets your node configuration to the factory default settings. Only select this menu item after the node has been removed from a cluster (unjoined) and is not joined to another cluster. For a node with internal or external disk shelves, the root volume on the internal disks is initialized. If there are no internal disk shelves, then the root volume on the external disks is initialized. For a system running FlexArray Virtualization with internal or external disk shelves, the array LUNs are not initialized. Any native disks on either internal or external shelves are initialized. For a system running FlexArray Virtualization with only array LUNS and no internal or external disk shelves, the root volume on the storage array LUNS are initialized, see Installing FlexArray. If the node you want to initialize has disks that are partitioned for root-data partitioning, the disks must be unpartitioned before the node can be initialized, see 9) Configure Advanced Drive Partitioning and Disks and aggregates management.
Perform aggregate and disk maintenance operations and obtain detailed aggregate and disk information.	5) Maintenance mode boot You exit Maintenance mode by using the halt command.
Restore the configuration information from the node's root volume to the boot device, such as a PC CompactFlash card	6) Update flash from backup config ONTAP stores some node configuration information on the boot device. When the node reboots, the information on the boot device is automatically backed up onto the node's root volume. If the boot device becomes corrupted or needs to be replaced, you must use this menu option to restore the configuration information from the node's root volume back to the boot device.

То	Select
Install new software on the node	7) Install new software first If the ONTAP software on the boot device does not include support for the storage array that you want to use for the root volume, you can use this menu option to obtain a version of the software that supports your storage array and install it on the node. This menu option is only for installing a newer version of ONTAP software on a node that has no root volume installed. Do <i>not</i> use this menu option to upgrade ONTAP.
Reboot the node	8) Reboot node
Unpartition all disks and remove their ownership information or clean the configuration and initialize the system with whole or partitioned disks	9) Configure Advanced Drive Partitioning Beginning with ONTAP 9.2, the Advanced Drive Partitioning option provides additional management features for disks that are configured for root-data or root-data-data partitioning. The following options are available from Boot Option 9:
	<pre>(9a) Unpartition all disks and remove their ownership information. (9b) Clean configuration and initialize system with partitioned disks. (9c) Clean configuration and initialize system with whole disks. (9d) Reboot the node. (9e) Return to main boot menu.</pre>

Copyright information

Copyright © 2023 NetApp, Inc. All Rights Reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system—without prior written permission of the copyright owner.

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETAPP "AS IS" AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

LIMITED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (b)(3) of the Rights in Technical Data -Noncommercial Items at DFARS 252.227-7013 (FEB 2014) and FAR 52.227-19 (DEC 2007).

Data contained herein pertains to a commercial product and/or commercial service (as defined in FAR 2.101) and is proprietary to NetApp, Inc. All NetApp technical data and computer software provided under this Agreement is commercial in nature and developed solely at private expense. The U.S. Government has a non-exclusive, non-transferrable, nonsublicensable, worldwide, limited irrevocable license to use the Data only in connection with and in support of the U.S. Government contract under which the Data was delivered. Except as provided herein, the Data may not be used, disclosed, reproduced, modified, performed, or displayed without the prior written approval of NetApp, Inc. United States Government license rights for the Department of Defense are limited to those rights identified in DFARS clause 252.227-7015(b) (FEB 2014).

Trademark information

NETAPP, the NETAPP logo, and the marks listed at http://www.netapp.com/TM are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.