Stage 2. Relocate and retire node1

AFF and FAS Controller Upgrade

NetApp
February 22, 2022

Table of Contents

Stage 2. Relocate and retire node1

Stage 2. Relocate and retire node1

During Stage 2, you relocate the node1 aggregates and LIFs to node2, record node1 information, and then retire node1.

Steps

1. Relocating non-root aggregates and NAS data LIFs owned by node1 to node2
2. Moving NAS data LIFs owned by node1 to node2
3. Recording node1 information
4. Retire node1

Relocate non-root aggregates from node1 to node2

Before you can replace node1 with node3, you need to move the non-root aggregates from node1 to node2 by using the storage aggregate relocation command and then verifying the relocation.

Steps

1. Relocate the non-root aggregates by completing the following substeps:
   a. Set the privilege level to advanced:

      ```
      set -privilege advanced
      ```

   b. Enter the following command:

      ```
      storage aggregate relocation start -node <node1> -destination <node2>
      -aggregate-list * -ndo-controller-upgrade true
      ```

   c. When prompted, enter `y`.

      Relocation will occur in the background. It could take anywhere from a few seconds to a couple of minutes to relocate an aggregate. The time includes both client outage and nonoutage portions. The command does not relocate any offline or restricted aggregates.

   d. Return to the admin level by entering the following command:

      ```
      set -privilege admin
      ```

2. Check the relocation status by entering the following command on node1:

   ```
   storage aggregate relocation show -node <node1>
   ```

   The output will display `Done` for an aggregate after it has been relocated.

   Wait until all non-root aggregates owned by node1 have been relocated to node2 before proceeding to the next step.

3. Take one of the following actions:

If relocation…	Then..
Of all aggregates is successful	Go to Step 4.

Of any aggregates fails or is vetoed	a. Check the EMS logs for the corrective action.
	b. Perform the corrective action.
	c. Relocate any failed or vetoed aggregates:
	`storage aggregate relocation start -node <node1> - destination <node2> -aggregate-list * -ndo-controller-upgrade true`
	d. When prompted, enter `y`.
	e. Return to the admin level:
	`set -privilege admin`
	If necessary, you can force the relocation using one of the following methods:
	◦ Override veto checks:
	`storage aggregate relocation start -override -vetoes true -ndo-controller-upgrade`
	◦ Override destination checks:
	`storage aggregate relocation start -override -destination-checks true -ndo-controller -upgrade`
	Refer to References to link to the *Disk and aggregate management with the CLI* content and the *ONTAP 9 Commands: Manual Page Reference* for more information about storage aggregate relocation commands.

4. Verify that all the non-root aggregates are online and their state on node2:

`storage aggregate show -node <node2> -state online -root false`

The following example shows that the non-root aggregates on node2 are online:

```
cluster::> storage aggregate show -node node2 state online -root false
Aggregate     Size Available Used% State   #Vols  Nodes            RAID
Status
--------- -------- --------- ----- ------- ------ ----------------
------------
aggr_1
           744.9GB 744.8GB      0% online       5 node2
raid_dp,

normal
aggr_2     825.0GB 825.0GB      0% online       1 node2
raid_dp,

normal
2 entries were displayed.
```

If the aggregates have gone offline or become foreign on node2, bring them online by using the following command on node2, once for each aggregate:

`storage aggregate online -aggregate <aggr_name>`

5. Verify that all the volumes are online on node2 by entering the following command on node2 and examining its output:

   `volume show -node <node2> -state offline`

   If any volumes are offline on node2, bring them online by using the following command on node2, once for each volume:

   `volume online -vserver <vserver-name> -volume <volume-name>`

   The `vserver-name` to use with this command is found in the output of the previous `volume show` command.

6. Enter the following command on node2:

   `storage failover show -node <node2>`

   The output should display the following message:

   ```
   Node owns partner's aggregates as part of the nondisruptive controller
   upgrade procedure.
   ```

7. Verify that node1 does not own any non-root aggregates that are online:

   `storage aggregate show -owner-name <node1> -ha-policy sfo -state online`

   The output should not display any online non-root aggregates, which have already been relocated to

node2.

Move NAS data LIFs owned by node1 to node2

Before you can replace node1 with node3, you need to move the NAS data LIFs owned by node1 to node2 if you have a two-node cluster, or to a third node if your cluster has more than two nodes. The method you use depends on whether the cluster is configured for NAS or SAN.

About this task

Remote LIFs handle traffic to SAN LUNs during the upgrade procedure. Moving SAN LIFs is not necessary for cluster or service health during the upgrade. You must verify that the LIFs are healthy and located on appropriate ports after you bring node3 online.

Steps

1. List all the NAS data LIFs hosted on node1 by entering the following command and capturing the output:

   `network interface show -data-protocol nfs|cifs -curr-node <node1>`

   The system displays the NAS data LIFs on node1, as shown in the following example:

```
 cluster::> network interface show -data-protocol nfs|cifs -curr-node
node1

          Logical     Status      Network             Current   Current
Is
 Vserver  Interface   Admin/Oper  Address/Mask        Node      Port
Home
 -------- ----------  ----------  ------------------  --------- -------
----
 vs0
          a0a         up/down     10.63.0.53/24       node1     a0a
true
          data1       up/up       10.63.0.50/18       node1     e0c
true
          rads1       up/up       10.63.0.51/18       node1     e1a
true
          rads2       up/down     10.63.0.52/24       node1     e1b
true
 vs1
          lif1        up/up       192.17.176.120/24   node1     e0c
true
          lif2        up/up       172.17.176.121/24   node1     e1a
true
```

2. Take one of the following actions:

If node1…	Then…
Has interface groups of VLANs configured	Go to Step 3.
Does not have interface groups or VLANs configured	Skip Step 3 and go to Step 4.

Use the `network port vlan show` command to display information about the network ports attached to VLANs, and use the `network port ifgrp show` command to display information about the port interface groups.

3. Take the following steps to migrate any NAS data LIFs hosted on interface groups and VLANs on node1:
   a. Migrate the LIFs hosted on any interface groups and the VLANs on node1 to a port on node2 that is capable of hosting LIFs on the same network as that of the interface groups by entering the following command, once for each LIF:

      ```
      network interface migrate -vserver <Vserver_name> -lif <LIF_name>
      -destination-node <node2> –destination-port <netport|ifgrp>
      ```

   b. Modify the home port and the home node of the LIFs and VLANs in Substep a to the port and node currently hosting the LIFs by entering the following command, once for each LIF:

      ```
      network interface modify -vserver <Vserver_name> -lif <LIF_name> -home-node
      <node2> - home-port <netport|ifgrp>
      ```

4. Take one of the following actions:

If the cluster is configured for…	Then…
NAS	Complete Step 5 through Step 8.
SAN	Disable all the SAN LIFs on the node to take them down for the upgrade: `network interface modify -vserver <Vserver-name> -lif <LIF_name> -home-node <node_to_upgrade> -home-port <netport|ifgrp> -status-admin down`

5. Migrate NAS data LIFs from node1 to node2 by entering the following command, once for each data LIF:

   ```
   network interface migrate -vserver <Vserver-name> -lif <LIF_name> -destination
   -node <node2> -destination-port <data_port>
   ```

6. Enter the following command and examine its output to verify that LIFs have been moved to the correct ports and that the LIFs have the status of up by entering the following command on either node and examining the output:

   ```
   network interface show -curr-node <node2> -data-protocol nfs|cifs
   ```

7. Enter the following command to modify the home node of the migrated LIFs:

   ```
   network interface modify -vserver <Vserver-name> -lif <LIF_name> -home-node
   <node2> -home-port <port_name>
   ```

8. Verify whether the LIF is using the port as its home or current port. If the port is not home or current port then go to Step 9:

```
network interface show -home-node <node2> -home-port <port_name>

network interface show -curr-node <node_name> -curr-port <port_name>
```

9. If the LIFs are using the port as a home port or current port, then modify the LIF to use a different port:

```
network interface migrate -vserver <Vserver-name> -lif <LIF_name>
-destination-node <node_name> -destination-port <port_name>

network interface modify -vserver <Vserver-name> -lif <LIF_name> -home-node
<node_name> -home-port <port_name>
```

10. If the ports currently hosting data LIFs are not going to exist on the new hardware, remove them from the broadcast domain now:

```
network port broadcast-domain remove-ports -ipspace Default -broadcast-domain
Default -ports <node:port>
```

11. If any LIFs are down, set the administrative status of the LIFs to "up" by entering the following command, once for each LIF:

```
network interface modify -vserver <Vserver-name> -lif <LIF_name> -home-node
<nodename> -status-admin up
```

> For MetroCluster configurations, you might not be able to change the broadcast domain of a port because it is associated with a port hosting the LIF of a destination storage virtual machine (SVM). Enter the following command from the corresponding source SVM on the remote site to reallocate the destination LIF to an appropriate port:
> `metrocluster vserver resync -vserver <Vserver_name>`

12. Enter the following command and examine its output to verify that there are no data LIFs remaining on node1:

```
network interface show -curr-node <node1> -role data
```

13. If you have interface groups or VLANs configured, complete the following substeps:
    a. Remove the VLANs from the interface groups by entering the following command:

    ```
    network port vlan delete -node <nodename> -port <ifgrp_name> -vlan-id
    <VLAN_ID>
    ```

    b. Enter the following command and examine its output to see if there are any interface groups configured on the node:

    ```
    network port ifgrp show -node <nodename> -ifgrp <ifgrp_name> -instance
    ```

    The system displays interface group information for the node as shown in the following example:

```
cluster::> network port ifgrp show -node node1 -ifgrp a0a -instance
                  Node: node1
 Interface Group Name: a0a
Distribution Function: ip
        Create Policy: multimode_lacp
          MAC Address: 02:a0:98:17:dc:d4
   Port Participation: partial
        Network Ports: e2c, e2d
             Up Ports: e2c
           Down Ports: e2d
```

c. If any interface groups are configured on the node, record the names of those groups and the ports assigned to them, and then delete the ports by entering the following command, once for each port:

`network port ifgrp remove-port -node <nodename> -ifgrp <ifgrp_name> -port <netport>`

Record node1 information

Before you can shut down and retire node1, you need to record information about its cluster network, management, and FC ports as well as its NVRAM System ID. You need that information later in the procedure when you map node1 to node3 and reassign disks.

Steps

1. Enter the following command and capture its output:

   `network route show`

   The system displays output similar to the following example:

```
cluster::> network route show

Vserver         Destination    Gateway      Metric
 -------------- -------------- ------------ -------
 iscsi vserver  0.0.0.0/0      10.10.50.1   20
 node1          0.0.0.0/0      10.10.20.1   10
 ....
 node2          0.0.0.0/0      192.169.1.1 20
```

2. Enter the following command and capture its output:

   `vserver services name-service dns show`

   The system displays output similar to the following example:

```
 cluster::> vserver services name-service dns show
                                                                    Name
 Vserver         State     Domains                                  Servers
 -------------- --------- ----------------------------------
---------------
 node 1 2       enabled   alpha.beta.gamma.netapp.com
10.10.60.10,

10.10.60.20
 vs_base1       enabled   alpha.beta.gamma.netapp.com,
10.10.60.10,
                          beta.gamma.netapp.com,
10.10.60.20
 ...
 ...
 vs peer1        enabled  alpha.beta.gamma.netapp.com,
10.10.60.10,
                          gamma.netapp.com
10.10.60.20
```

3. Find the cluster network and node-management ports on node1 by entering the following command on either controller:

   `network interface show -curr-node <node1> -role cluster,intercluster,node-mgmt,cluster-mgmt`

   The system displays the cluster, intercluster, node-management, and cluster-management LIFs for the node in the cluster, as shown in the following example:

```
 cluster::> network interface show -curr-node <node1>
            -role cluster,intercluster,node-mgmt,cluster-mgmt

              Logical       Status     Network            Current
Current Is
  Vserver     Interface     Admin/Oper Address/Mask       Node     Port
Home
  ----------- ------------- ---------- ------------------ --------
------- ----
  vserver1
              cluster mgmt   up/up     192.168.x.xxx/24   node1    e0c
true
  node1
              intercluster   up/up     192.168.x.xxx/24   node1    e0e
true
              clus1          up/up     169.254.xx.xx/24   node1    e0a
true
              clus2          up/up     169.254.xx.xx/24   node1    e0b
true
              mgmt1          up/up     192.168.x.xxx/24   node1    e0c
true
 5 entries were displayed.
```

Your system might not have intercluster LIFs.

4. Capture the information in the output of the command in Step 3 to use in the section Map ports from node1 to node3.

   The output information is required to map the new controller ports to the old controller ports.

5. Enter the following command on node1:

   `network port show -node <node1> -type physical`

   The system displays the physical ports on the node as shown in the following example:

```
 sti8080mcc-htp-008::> network port show -node sti8080mcc-htp-008 -type
physical

 Node: sti8080mcc-htp-008

                                                                        Ignore
                                                  Speed(Mbps)  Health   Health
 Port   IPspace  Broadcast Domain Link MTU   Admin/Oper   Status   Status
 ----   -------  ---------------- ---- ----  -----------  -------
-------
 e0M    Default  Mgmt             up   1500  auto/1000    healthy  false
 e0a    Default  Default          up   9000  auto/10000   healthy  false
 e0b    Default  -                up   9000  auto/10000   healthy  false
 e0c    Default  -                down 9000  auto/-       -        false
 e0d    Default  -                down 9000  auto/-       -        false
 e0e    Cluster  Cluster          up   9000  auto/10000   healthy  false
 e0f    Default  -                up   9000  auto/10000   healthy  false
 e0g    Cluster  Cluster          up   9000  auto/10000   healthy  false
 e0h    Default  Default          up   9000  auto/10000   healthy  false
 9 entries were displayed.
```

6. Record the ports and their broadcast domains.

   The broadcast domains will need to be mapped to the new ports on the new controller later in the procedure.

7. Enter the following command on node1:

   `network fcp adapter show -node <node1>`

   The system displays the FC ports on the node, as shown in the following example:

```
 cluster::> fcp adapter show -node <node1>
                      Connection  Host
 Node         Adapter Established Port Address
 ------------ ------- ----------- ------------
 node1
               0a     ptp         11400
 node1
               0c     ptp         11700
 node1
               6a     loop        0
 node1
               6b     loop        0
 4 entries were displayed.
```

8. Record the ports.

   The output information is required to map the new FC ports on the new controller later in the procedure.

9. If you did not do so earlier, check whether there are interface groups or VLANs configured on node1 by entering the following commands:

   ```
   network port ifgrp show

   network port vlan show
   ```

   You will use the information in the section Map ports from node1 to node3.

10. Take one of the following actions:

If you…	Then…
Recorded the NVRAM System ID number in the section Prepare the nodes for the upgrade.	Go on to the next section, Retire node1.
Did not record the NVRAM System ID number in the section Prepare the nodes for the upgrade	Complete Step 11 and Step 12 and then continue to Retire node1.

11. Enter the following command on either controller:

    ```
    system node show -instance -node <node1>
    ```

    The system displays information about node1 as shown in the following example:

    ```
    cluster::> system node show -instance -node <node1>
                                Node: node1
                               Owner:
                            Location: GDl
                               Model: FAS6240
                       Serial Number: 700000484678
                           Asset Tag: -
                              Uptime: 20 days 00:07
                     NVRAM System ID: 1873757983
                           System ID: 1873757983
                              Vendor: NetApp
                              Health: true
                         Eligibility: true
    ```

12. Record the NVRAM System ID number to use in the section Install and boot node3.

Retire node1

To retire node1, you need to disable the HA pair with node2, shut node1 down properly, and remove it from the rack or chassis.

Steps

1. Verify the number of nodes in the cluster:

   `cluster show`

   The system displays the nodes in the cluster, as shown in the following example:

   ```
   cluster::> cluster show
   Node                  Health  Eligibility
   --------------------- ------- ------------
   node1                 true    true
   node2                 true    true
   2 entries were displayed.
   ```

2. Disable storage failover, as applicable:

If the cluster is…	Then…

A two-node cluster	a. Disable cluster high availability by entering the following command on either node:
	
	`cluster ha modify -configured false`
	
	a. Disable storage failover:
	
	`storage failover modify -node <node1> -enabled false`

A cluster with more than two nodes	Disable storage failover: `storage failover modify -node <node1> -enabled false`

3. Verify that storage failover was disabled:

   `storage failover show`

   The following example shows the output of the `storage failover show` command when storage failover has been disabled for a node:

```
 cluster::> storage failover show
                                 Takeover
 Node           Partner         Possible State Description
 -------------- -------------- --------
-------------------------------------
 node1          node2          false    Connected to node2, Takeover
                                        is not possible: Storage
failover is
                                        disabled

 node2          node1          false    Node owns partner's aggregates
as part
                                        of the nondisruptive controller
upgrade
                                        procedure. Takeover is not
possible:
                                        Storage failover is disabled
 2 entries were displayed.
```

4. Verify the data LIF status:

   `network interface show -role data -curr-node <node2> -home-node <node1>`

   Look in the **Status Admin/Oper** column to see if any LIFs are down. If any LIFs are down, consult the Troublehsoot section.

5. Take one of the following actions:

If the cluster is…	Then…
A two-node cluster	Go to Step 6.
A cluster with more than two nodes	Go to Step 8.

6. Access the advanced privilege level on either node:

   `set -privilege advanced`

7. Verify that the cluster HA has been disabled:

   `cluster ha show`

   The system displays the following message:

   ```
   High Availability Configured: false
   ```

   If cluster HA has not been disabled, repeat Step 2.

8. Check whether node1 currently holds epsilon:

   `cluster show`

   Because there is the possibility of a tie in a cluster that has an even number of nodes, one node has an extra fractional voting weight called epsilon. Refer to References to link to the *System Administration Reference* for more information.

   > If you have a four-node cluster, epsilon might be on a node in a different HA pair in the cluster.
   >
   > If you are upgrading a HA pair in a cluster with multiple HA pairs, you should move epsilon to the node of a HA pair not undergoing a controller upgrade. For example, if you are upgrading nodeA/nodeB in a cluster with the HA pair configuration nodeA/nodeB and nodeC/nodeD, you should move epsilon to nodeC or nodeD.

   The following example shows that node1 holds epsilon:

   ```
   cluster::*> cluster show

   Node                  Health  Eligibility  Epsilon
   --------------------- ------- ------------ ------------
   node1                 true    true         true
   node2                 true    true         false
   ```

9. If node1 holds epsilon, then mark epsilon `false` on the node so that it can be transferred to the node2:

   `cluster modify -node <node1> -epsilon false`

10. Transfer epsilon to node2 by marking epsilon `true` on node2:

    `cluster modify -node <node2> -epsilon true`

11. Verify that the change to node2 occurred:

    `cluster show`

    ```
    cluster::*> cluster show
    Node                  Health  Eligibility  Epsilon
    --------------------- ------- ------------ ------------
    node1                 true    true         false
    node2                 true    true         true
    ```

    The epsilon for node2 should now be true and the epsilon for node1 should be false.

12. Verify whether the setup is a two-node switchless cluster:

    `network options switchless-cluster show`

```
cluster::*> network options switchless-cluster show

Enable Switchless Cluster: false/true
```

The value of this command must match the physical state of the system.

13. Return to the admin level:

    `set -privilege admin`

14. Halt node1 from the node1 prompt:

    `system node halt -node <node1>`

    > **Attention**: If node1 is in same chassis as node2, do not power off the chassis by using the power switch or by pulling the power cable. If you do so, node2, which is serving data, will go down.

15. When the system prompts you to confirm that you want to halt the system, enter `y`.

    The node stops at the boot environment prompt.

16. When node1 displays the boot environment prompt, remove it from the chassis or the rack.

    You can decommission node1 after the upgrade is completed. See Decommission the old system.