Setup and configure NVIDIA SN2100 switches

ONTAP Systems Switches

NetApp
July 19, 2022

Table of Contents

Setup and configure NVIDIA SN2100 switches

Setup and configure the NVIDIA SN2100 switches

The NVIDIA SN2100 switch is a 10/25/40/100 Gb Ethernet switch running Cumulus Linux. The SN2100 switch serves Cluster and Storage applications in ONTAP 9.10.1P3 over different switch-pairs.

Cumulus Linux (CL) OS can be installed either when the switch is running Cumulus Linux or ONIE. For this release, the minimal version of Cumulus Linux supported is 4.4.2.

> The procedures here use Network Command Line Utility (NCLU) which is a command line interface that ensures Cumulus Linux is fully accessible to all. The net command is the wrapper utility you use to execute actions from a terminal.

> When using breakout cables for 10G and 25G, make sure that auto-negotiation is off and hard set the port speed on the switch. See Cabling and configuration considerations for further details.

Cabling and configuration considerations

Before configuring your NVIDIA SN2100 switch, review the following information:

1. Only optical connections are supported on SN2100 switches with X1151A NIC, X1146A NIC, or onboard 100GbE ports. For example:
   a. AFF A800 on ports e0a and e0b
   b. AFF A320 on ports e0g and e0h
2. When a QSA adapter is used to connect to the onboard Intel cluster ports on a platform, not all links come up.

   Example platforms are: FAS2750, AFF A300, and FAS8200 (all 10G) and AFF A250 (25G).

   To resolve this issue, do the following:

   a. For Intel 10G, manually set the swp1s0-3 link speed to 10000 and set auto-negotiation to offset
   b. For Chelsio 25G, manually set the swp2s0-3 link speed to 25000 and set auto-negotiation to off

   > Using 10G/25G QSA, use the non-breakout 40/100G ports. Do not insert the QSA adapter on ports that are configured for breakout.

3. Depending on the transceiver in the switchport, you might need to set the speed on the switchport to fixed speed. If using 10G and 25G breakout ports, make sure that auto-negotiation is off and hard set the port speed on the switch. For example:

```
cumulus@cumulus:mgmt:~$ net add int swp1s3 link autoneg on && net com
--- /etc/network/interfaces     2019-11-17 00:17:13.470687027 +0000
+++ /run/nclu/ifupdown2/interfaces.tmp  2019-11-24 00:09:19.435226258
+0000
@@ -37,21 +37,21 @@
     alias 10G Intra-Cluster Node
     link-autoneg off
     link-speed 10000
     mstpctl-bpduguard yes
     mstpctl-portadminedge yes
     mtu 9216

auto swp1s3
iface swp1s3
     alias 10G Intra-Cluster Node
-    link-autoneg off
+    link-autoneg on
     link-speed 10000
     mstpctl-bpduguard yes
     mstpctl-portadminedge yes
     mtu 9216

auto swp2s0
iface swp2s0
     alias 25G Intra-Cluster Node
     link-autoneg off
     link-speed 25000
```

Install Cumulus Linux in Cumulus mode

Cumulus Linux (CL) OS can be installed either when the switch is running Cumulus Linux or ONIE.

Before you begin

The following assumptions are made:

- You have intermediate-level Linux knowledge.
- You are familiar with basic text editing, UNIX file permissions, and process monitoring. A variety of text editors are pre-installed, including `vi` and `nano`.
- You must have access to a Linux or UNIX shell. If you are running Windows, use a Linux environment as your command line tool for interacting with Cumulus Linux.

Each time Cumulus Linux is installed, the entire file system structure is erased and rebuilt.

The default password for the cumulus user account is **cumulus**. The first time you log into Cumulus Linux, you must change this default password. Be sure to update any automation scripts before installing a new image. Cumulus Linux provides command line options to change the default password automatically during the installation process.

The baud rate requirement must be set to 115200 on the serial console switch for NVIDIA SN2100 switch console access, as follows:

- 115200 baud
- 8 data bits
- 1 stop bit
- parity: none
- flow control: none

Steps

1. Log in to the switch. First time log in to the switch requires username/password of **cumulus**/**cumulus** with `sudo` privileges:

```
cumulus login: cumulus
Password: cumulus
You are required to change your password immediately (administrator
enforced)
Changing password for cumulus.
Current password: cumulus
New password: netapp1!
Retype new password: netapp1!
```

2. Check the Cumulus Linux version:

```
cumulus@cumulus:mgmt:~$ net show system
Hostname......... cumulus
Build............ Cumulus Linux 4.4.2
Uptime........... 0:08:20.860000
Model............ Mlnx X86
CPU.............. x86_64 Intel Atom C2558 2.40GHz
Memory........... 8GB
Disk............. 14.7GB
ASIC............. Mellanox Spectrum MT52132
Ports............ 16 x 100G-QSFP28
Part Number...... MSN2100-CB2FC
Serial Number.... MT2105T05177
Platform Name.... x86_64-mlnx_x86-r0
Product Name..... MSN2100
ONIE Version..... 2019.11-5.2.0020-115200
Base MAC Address. 04:3F:72:43:92:80
Manufacturer..... Mellanox
```

3. Configure the hostname, IP address, subnet mask, and default gateway. The new hostname only becomes effective after restarting the console/SSH session.

> A Cumulus Linux switch provides at least one dedicated Ethernet management port called `eth0`. This interface is specifically for out-of-band management use. By default, the management interface uses DHCPv4 for addressing.

> Do not use an underscore (_), apostrophe ('), or non-ASCII characters in the hostname.

```
cumulus@cumulus:mgmt:~$ net add hostname sw1
cumulus@cumulus:mgmt:~$ net add interface eth0 ip address
10.233.204.71/23
cumulus@cumulus:mgmt:~$ net add interface eth0 ip gateway 10.233.204.1
cumulus@cumulus:mgmt:~$ net pending
cumulus@cumulus:mgmt:~$ net commit
```

This command modifies both the `/etc/hostname` and `/etc/hosts` files.

4. Confirm that the hostname, IP address, subnet mask, and default gateway have been updated:

```
cumulus@sw1:mgmt:~$ hostname sw1
cumulus@sw1:mgmt:~$ ifconfig eth0
eth0: flags=4163<UP,BROADCAST,RUNNING,MULTICAST>  mtu 1500
inet 10.233.204.71  netmask 255.255.254.0  broadcast 10.233.205.255
inet6 fe80::bace:f6ff:fe19:1df6  prefixlen 64  scopeid 0x20<link>
ether b8:ce:f6:19:1d:f6  txqueuelen 1000  (Ethernet)
RX packets 75364  bytes 23013528 (21.9 MiB)
RX errors 0  dropped 7  overruns 0  frame 0
TX packets 4053  bytes 827280 (807.8 KiB)
TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0 device memory
0xdfc00000-dfc1ffff

cumulus@sw1::mgmt:~$ ip route show vrf mgmt
default via 10.233.204.1 dev eth0
unreachable default metric 4278198272
10.233.204.0/23 dev eth0 proto kernel scope link src 10.233.204.71
127.0.0.0/8 dev mgmt proto kernel scope link src 127.0.0.1
```

5. Configure the time zone using NTP interactive mode.
   a. On a terminal, run the following command:

      ```
      cumulus@sw1:~$ sudo dpkg-reconfigure tzdata
      ```

   b. Follow the on-screen menu options to select the geographic area and region.
   c. To set the time zone for all services and daemons, reboot the switch.
   d. Verify that the date and time on the switch are correct and update if necessary.
6. Install Cumulus Linux 4.4.2:

   ```
   cumulus@sw1:mgmt:~$ sudo onie-install -a -i
   http://10.60.132.97/x/eng/testbedN,svl/nic/files/cumulus-linux-4.4.2-
   mlx-amd64.bin
   ```

   The installer starts the download. Type **y** when prompted.
7. Reboot the NVIDIA SN2100 switch:

   ```
   cumulus@sw1:mgmt:~$ sudo reboot
   ```

8. The installation starts automatically, and the following GRUB screens appear. Do **not** make any selections:
   - Cumulus-Linux GNU/Linux
   - ONIE: Install OS

- CUMULUS-INSTALL
- Cumulus-Linux GNU/Linux

9. Repeat steps 1 to 4 to log in.
10. Verify that the Cumulus Linux version is 4.4.2:

```
cumulus@sw1:mgmt:~$ net show version
NCLU_VERSION=1.0-cl4.4.2u0
DISTRIB_ID="Cumulus Linux"
DISTRIB_RELEASE=4.4.2
DISTRIB_DESCRIPTION="Cumulus Linux 4.4.2"
```

11. Create a new user and add this user to the `sudo` group. This user only becomes effective after the console/SSH session is restarted:

```
cumulus@sw1:mgmt:~$ sudo adduser --ingroup netedit admin
[sudo] password for cumulus:
Adding user `admin' ...
Adding new user `admin' (1001) with group `netedit' ...
Creating home directory `/home/admin' ...
Copying files from `/etc/skel' ...
New password:
Retype new password:
passwd: password updated successfully
Changing the user information for admin
Enter the new value, or press ENTER for the default
Full Name []:
Room Number []:
Work Phone []:
Home Phone []:
Other []:
Is the information correct? [Y/n] y

cumulus@sw1:mgmt:~$ sudo adduser admin sudo
[sudo] password for cumulus:
Adding user `admin' to group `sudo' ...
Adding user admin to group sudo
Done.
cumulus@sw1:mgmt:~$ exit
logout
Connection to 10.233.204.71 closed.

[admin@cycrh6svl01 ~]$ ssh admin@10.233.204.71
admin@10.233.204.71's password:
Linux sw1 4.19.0-cl-1-amd64 #1 SMP Cumulus 4.19.206-1+cl4.4.1u1 (2021-
09-09) x86_64
Welcome to NVIDIA Cumulus (R) Linux (R)

For support and online technical documentation, visit
http://www.cumulusnetworks.com/support

The registered trademark Linux (R) is used pursuant to a sublicense from
LMI, the exclusive licensee of Linus Torvalds, owner of the mark on a
world-wide basis.
admin@sw1:mgmt:~$
```

Install Cumulus Linux in ONIE mode

Cumulus Linux (CL) OS can be installed either when the switch is running Cumulus Linux

or ONIE.

Before you begin

You can install the Cumulus Linux using Open Network Install Environment (ONIE) that allows for automatic discovery of a network installer image. This facilitates the system model of securing switches with an operating system choice, such as Cumulus Linux. The easiest way to install Cumulus Linux with ONIE is with local HTTP discovery.

If your host is IPv6-enabled, make sure it is running a web server. If your host is IPv4-enabled, make sure it is running DHCP in addition to a web server.

This procedure demonstrates how to upgrade Cumulus Linux after the admin has booted in ONIE.

Steps

1. Download the Cumulus Linux installation file to the root directory of the web server. Rename this file `onie-installer`.
2. Connect your host to the management Ethernet port of the switch using an Ethernet cable.
3. Power on the switch. The switch downloads the ONIE image installer and boots. After the installation completes, the Cumulus Linux login prompt appears in the terminal window.

   Each time Cumulus Linux is installed, the entire file system structure is erased and rebuilt.

4. Reboot the SN2100 switch:

```
cumulus@cumulus:mgmt:~$ sudo reboot
```

5. Hit the **Esc** key at the GNU GRUB screen to interrupt the normal boot process, select **ONIE** and press Enter.
6. On the next screen displayed, select **ONIE: Install OS**.
7. The ONIE installer discovery process runs searching for the automatic installation. Press Enter to temporarily stop the process.
8. When the discovery process has stopped:

```
ONIE:/ # onie-stop
discover: installer mode detected.
Stopping: discover...start-stop-daemon: warning: killing process 427:
No such process done.
```

9. If the DHCP service is running on your network, verify that the IP address, subnet mask, and the default gateway are correctly assigned:

```
ONIE:/ # ifconfig eth0
eth0   Link encap:Ethernet  HWaddr B8:CE:F6:19:1D:F6
       inet addr:10.233.204.71  Bcast:10.233.205.255  Mask:255.255.254.0
       inet6 addr: fe80::bace:f6ff:fe19:1df6/64 Scope:Link
       UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
       RX packets:21344 errors:0 dropped:2135 overruns:0 frame:0
       TX packets:3500 errors:0 dropped:0 overruns:0 carrier:0
       collisions:0 txqueuelen:1000
       RX bytes:6119398 (5.8 MiB)  TX bytes:472975 (461.8 KiB)
       Memory:dfc00000-dfc1ffff

ONIE:/ # route
Kernel IP routing table
Destination     Gateway         Genmask         Flags Metric Ref    Use
Iface

default         10.233.204.1    0.0.0.0         UG    0      0        0
eth0
10.233.204.0    *               255.255.254.0   U     0      0        0
eth0
```

10. If the IP addressing scheme is manually defined, do the following:

```
ONIE:/ # ifconfig eth0 10.233.204.71 netmask 255.255.254.0
ONIE:/ # route add default gw 10.233.204.1
```

11. Repeat step 9 to verify that the static information is correctly entered.
12. Install Cumulus Linux:

```
ONIE:/ # route

  Kernel IP routing table

  ONIE:/ # onie-nos-install
http://10.60.132.97/x/eng/testbedN,svl/nic/files/cumulus-linux-4.4.2-
mlx-amd64.bin

  Stopping: discover... done.
  Info: Attempting
http://10.60.132.97/x/eng/testbedN,svl/nic/files/cumulus-linux-4.4.2-
mlx-amd64.bin ...
  Connecting to 10.60.132.97 (10.60.132.97:80)
  installer            100% |*|   552M  0:00:00 ETA
  ...
  ...
```

13. Once the installation has completed, log in to the switch:

```
cumulus login: cumulus
Password: cumulus
You are required to change your password immediately (administrator
enforced)
Changing password for cumulus.
Current password: cumulus
New password: netapp1!
Retype new password: netapp1!
```

14. Verify the Cumulus Linux version:

```
cumulus@cumulus:mgmt:~$ net show version
NCLU_VERSION=1.0-cl4.4.2u4
DISTRIB_ID="Cumulus Linux"
DISTRIB_RELEASE=4.4.2
DISTRIB_DESCRIPTION="Cumulus Linux 4.4.2”
```

Install the RCF script

Before installing the RCF script, ensure that the following are available on the switch:

- Cumulus Linux 4.4.2 is installed.
- IP address, subnet mask, and default gateway defined via DHCP or manually configured.

See Cabling and configuration considerations for caveats and further details.

Current RCF script versions

There are two RCF scripts available for Clustering and Storage applications. The procedure for each is the same.

- Clustering: **MSN2100-RCF-v1.8-Cluster**
- Storage: **MSN2100-RCF-v1.8-Storage**

The following example procedure shows how to download and apply the RCF script for Clusters.

Steps

1. Display the available interfaces on the SN2100 switch:

```
cumulus@cumulus:mgmt:~$ net show interface all

State  Name    Spd  MTU    Mode          LLDP                 Summary
-----  -----   ---  -----  -----------   ------------------
--------------
...
...
ADMDN  swp1    N/A  9216   NotConfigured
ADMDN  swp2    N/A  9216   NotConfigured
ADMDN  swp3    N/A  9216   NotConfigured
ADMDN  swp4    N/A  9216   NotConfigured
ADMDN  swp5    N/A  9216   NotConfigured
ADMDN  swp6    N/A  9216   NotConfigured
ADMDN  swp7    N/A  9216   NotConfigure
ADMDN  swp8    N/A  9216   NotConfigured
ADMDN  swp9    N/A  9216   NotConfigured
ADMDN  swp10   N/A  9216   NotConfigured
ADMDN  swp11   N/A  9216   NotConfigured
ADMDN  swp12   N/A  9216   NotConfigured
ADMDN  swp13   N/A  9216   NotConfigured
ADMDN  swp14   N/A  9216   NotConfigured
ADMDN  swp15   N/A  9216   NotConfigured
ADMDN  swp16   N/A  9216   NotConfigured
```

2. Copy the RCF script to the switch:

```
cumulus@cumulus:mgmt:~$ pwd
/home/cumulus
cumulus@cumulus:mgmt: /tmp$ scp
ssologin@10.60.132.97:/x/eng/testbedN,svl/nic/files/MSN2100-RCF-v1.8-
Cluster
ssologin@10.60.132.97's password:
MSN2100-RCF-v1.8-Cluster                  100% 8607   111.2KB/s
00:00
```

3. Apply the RCF script **MSN2100-RCF-v1.8-Cluster**:

```
cumulus@cumulus:mgmt:/tmp$ sudo python3 MSN2100-RCF-v1.8-Cluster
[sudo] password for cumulus:
...
```

The RCF script completes the following steps:

a. Updates the banner MOTD
b. Disables the apt-get for OS updates
c. Defines breakout and non-breakout interfaces
d. Configures interfaces and SNMP
e. Disables CDP
f. Changes the LLDP configuration
g. Adds a RoCE configuration
h. Modifies the RoCE configuration for HA and Cluster RDMA
i. Reboots the switch

4. Verify the configuration after the reboot:

```
cumulus@cumulus:mgmt:~$ net show interface all

State  Name      Spd   MTU    Mode       LLDP               Summary
-----  --------- ----  -----  ---------- ------------------ --------
...
...
DN     swp1s0    N/A   9216   Trunk/L2                      Master:
bridge(UP)
DN     swp1s1    N/A   9216   Trunk/L2                      Master:
bridge(UP)
DN     swp1s2    N/A   9216   Trunk/L2                      Master:
bridge(UP)
DN     swp1s3    N/A   9216   Trunk/L2                      Master:
bridge(UP)
```

```
DN     swp2s0    N/A   9216   Trunk/L2                    Master:
bridge(UP)
DN     swp2s1    N/A   9216   Trunk/L2                    Master:
bridge(UP)
DN     swp2s2    N/A   9216   Trunk/L2                    Master:
bridge(UP)
DN     swp2s3    N/A   9216   Trunk/L2                    Master:
bridge(UP)
UP     swp3      100G  9216   Trunk/L2                    Master:
bridge(UP)
UP     swp4      100G  9216   Trunk/L2                    Master:
bridge(UP)
DN     swp5      N/A   9216   Trunk/L2                    Master:
bridge(UP)
DN     swp6      N/A   9216   Trunk/L2                    Master:
bridge(UP)
DN     swp7      N/A   9216   Trunk/L2                    Master:
bridge(UP)
DN     swp8      N/A   9216   Trunk/L2                    Master:
bridge(UP)
DN     swp9      N/A   9216   Trunk/L2                    Master:
bridge(UP)
DN     swp10     N/A   9216   Trunk/L2                    Master:
bridge(UP)
DN     swp11     N/A   9216   Trunk/L2                    Master:
bridge(UP)
DN     swp12     N/A   9216   Trunk/L2                    Master:
bridge(UP)
DN     swp13     N/A   9216   Trunk/L2                    Master:
bridge(UP)
DN     swp14     N/A   9216   Trunk/L2                    Master:
bridge(UP)
UP     swp15     N/A   9216   BondMember                  Master:
bond_15_16(UP)
UP     swp16     N/A   9216   BondMember                  Master:
bond_15_16(UP)
...
...

cumulus@cumulus:mgmt:~$ net show roce config
RoCE mode.......... lossless
Congestion Control:
  Enabled SPs.... 0 2 5
  Mode........... ECN
  Min Threshold.. 150 KB
  Max Threshold.. 1500 KB
```

```
PFC:
  Status......... enabled
  Enabled SPs.... 2 5
  Interfaces......... swp10-16,swp1s0-3,swp2s0-3,swp3-9

DSCP                     802.1p  switch-priority
-----------------------  ------  ---------------
0 1 2 3 4 5 6 7               0                0
8 9 10 11 12 13 14 15         1                1
16 17 18 19 20 21 22 23       2                2
24 25 26 27 28 29 30 31       3                3
32 33 34 35 36 37 38 39       4                4
40 41 42 43 44 45 46 47       5                5
48 49 50 51 52 53 54 55       6                6
56 57 58 59 60 61 62 63       7                7

switch-priority  TC  ETS
---------------  --  --------
0 1 3 4 6 7       0  DWRR 28%
2                 2  DWRR 28%
5                 5  DWRR 43%
```

5. Verify information for the transceiver in the interface. `net show interface pluggables`

```
cumulus@cumulus:mgmt:~$ net show interface pluggables
Interface  Identifier     Vendor Name  Vendor PN        Vendor SN
Vendor Rev
---------  -------------  -----------  ---------------  --------------
----------
swp3       0x11 (QSFP28)  Amphenol     112-00574        APF20379253516
B0
swp4       0x11 (QSFP28)  AVAGO        332-00440        AF1815GU05Z
A0
swp15      0x11 (QSFP28)  Amphenol     112-00573        APF21109348001
B0
swp16      0x11 (QSFP28)  Amphenol     112-00573        APF21109347895
B0
```

6. Verify that the nodes each have a connection to each switch: `net show lldp`

```
cumulus@cumulus:mgmt:~$ net show lldp

LocalPort  Speed  Mode        RemoteHost              RemotePort
---------  -----  ----------  ----------------------  -----------
swp3       100G   Trunk/L2    sw1                     e3a
swp4       100G   Trunk/L2    sw2                     e3b
swp15      100G   BondMember  sw13                    swp15
swp16      100G   BondMember  sw14                    swp16
```

7. Verify the health of cluster ports on the cluster.
   a. Verify that e0d ports are up and healthy across all nodes in the cluster: `network port show -role cluster`

```
cluster1::*> network port show -role cluster

Node: node1

Ignore
                                                  Speed(Mbps) Health
Health
Port      IPspace      Broadcast Domain Link MTU  Admin/Oper  Status
Status
--------- ------------ ---------------- ---- ---- -----------
-------- ------
e3a       Cluster      Cluster          up   9000  auto/10000 healthy
false
e3b       Cluster      Cluster          up   9000  auto/10000 healthy
false

Node: node2

Ignore
                                                  Speed(Mbps) Health
Health
Port      IPspace      Broadcast Domain Link MTU  Admin/Oper  Status
Status
--------- ------------ ---------------- ---- ---- -----------
-------- ------
e3a       Cluster      Cluster          up   9000  auto/10000 healthy
false
e3b       Cluster      Cluster          up   9000  auto/10000 healthy
false
```

   b. Verify the switch health from the cluster (this might not show switch sw2, since LIFs are not homed on

e0d).

```
cluster1::*> network device-discovery show -protocol lldp
Node/       Local  Discovered
Protocol    Port   Device (LLDP: ChassisID)  Interface Platform
----------- ------ ------------------------- --------- ----------
node1/lldp
            e3a    sw1 (b8:ce:f6:19:1a:7e)   swp3      -
            e3b    sw2 (b8:ce:f6:19:1b:96)   swp3      -

node2/lldp
            e3a    sw1 (b8:ce:f6:19:1a:7e)   swp4      -
            e3b    sw2 (b8:ce:f6:19:1b:96)   swp4      -


cluster1::*> system switch ethernet show -is-monitoring-enabled
-operational true
Switch                      Type               Address          Model
--------------------------- ------------------ ---------------- -----
sw1                         cluster-network    10.233.205.90
MSN2100-CB2RC
     Serial Number: MNXXXXXXGD
      Is Monitored: true
            Reason: None
  Software Version: Cumulus Linux version 4.4.2 running on Mellanox
                    Technologies Ltd. MSN2100
    Version Source: LLDP

sw2                         cluster-network    10.233.205.91
MSN2100-CB2RC
     Serial Number: MNCXXXXXXGS
      Is Monitored: true
            Reason: None
  Software Version: Cumulus Linux version 4.4.2 running on Mellanox
                    Technologies Ltd. MSN2100
    Version Source: LLDP
```

Cable NS224 shelves as switch-attached storage

If you have a system in which the NS224 drive shelves need to be cabled as switch-attached storage (not direct-attached storage), use the information provided here.

- Cable NS224 drive shelves through storage switches:

  Information for cabling switch-attached NS224 drive shelves

- Install your storage switches:

  AFF and FAS Switch Documentation

- Confirm supported hardware, such as storage switches and cables, for your platform model:

  NetApp Hardware Universe