Troubleshooting the Cisco Nexus 9000v

This chapter contains the following sections:

Common Issues For All Hypervisors

How to boot when VM falls into "loader >" prompt

Generally, the first time boot is successful. However, the system boot could fail and drops to the "loader >" prompt on the VGA console or serial console depending on how the VM is provisioned.

Example:

```
loader > dir
Setting listing for bootflash:
Number of devices detected by BIOS is 1
Number of devices detected by BIOS is 1
Number of devices detected by BIOS is 1
Going to print files for device bootflash:
 .rpmstore
 nxos.7.9.3.15.9.66. bin
Number of devices detected by BIOS is 1
Number of devices detected by BIOS is 1
Number of devices detected by BIOS is 1
Clearing listing for bootflash:

loader >
```

To continue the boot, enter the **boot nxos.7.0.3.I5.0.66.bin** command at the "loader >" prompt

How to prevent VM from dropping into "loader >" prompt

As soon as you set up your Cisco Nexus 9000v (following set up of POAP interface), you need to configure the boot image in your system to avoid dropping to the "loader >" prompt after reload/shut down.

Example:

```
config t

  boot nxos n9000-dk9.7.0.3.I2.0.454.bin

copy running starting
```

ESXi Hypervisor

How to use SATA controller to speed up Cisco Nexus 9000v booting process

Cisco Nexus 9000v uses the same hardware platform image boot on hypervisors. ESXi 5.5 and later versions support a SATA controller on an ESXi server that you can use to speed up Cisco Nexus 9000v boot time. To create a VM with a SATA controller, the regular ESXi VM creation steps are applicable except the following are required for a successful VM boot:

- The VMware vSphere Web Client is needed to access this support.
- Download the vmdk image into the ESXi server.

  Convert this monolith vmdk into a VMware native disk type using vmkfstools (command line tool available with the ESXI server)

  Example:
  ```
  vmkfstools -i nexus9000v-final.7.0.3.I5.0.66.vmdk nexus9000v-final.7.0.3.I5.0.66-esx.vmdk)
  ```
- Create a VM that is compatible with ESXi 5.5 (or later) and VM version 10.
- Add the SATA controller.
- Add the existing disk with the SATA controller selected.
- Continue the VM booting process from the ESXi VM creation instruction.

How to access the "loader >" prompt from the serial console

EFI BIOS defaults all input/output to the VM console. When a VM drops to "loader >" prompt, you must go to the vSphere client to access "loader >" to boot another image. You can change this behavior by adding an extra configuration in the ESXi VM editing mode.

You can use one of the following methods:

- In the vSphere client Configuration Parameters window, you can add one row in the configuration (Edit Settings > VM Options > Advanced > Edit Configuration).
- You can add **efi.serialconsole.enabled = "TRUE"** to the .vmx file once the VM is created.

How to connect to the switch on ESXi if the EFI serial console is not enabled

On ESXi when you are monitoring the VM console, you might see "Leaving grub land". After this, even though it appears that nothing is happening, the communication has transferred to the serial port you had configured.

```
Read length 646737920
Hd5 for size 646737920
    [Initrd, addr-Ox59236000, size=0x268c70000]

segment header
length: 4, vendor: 16 flags: 4, loadaddr: 2500000, image len: 600 memory length
: 600
Reading data for vendor seg . Length 1536

Image length: 651842048 bytes

image hash: d411d638 b48101f6 2e5e7fOb f0130b67
Leaving grub land
```

To connect to the switch you need to open a terminal and enter the **telnet** *<esxi host> <port number>* command.

```
rahushen@rtp-ads-15Ø->
rahushen@rtp-ads-15Ø->telnet fe-ucs-dt7 7ØØØ
Trying 1Ø.122.84.213...
Connected to fe-ucs-dt7.
Escape character is '^]'.

User Access Verification
switch login: admin
Password :
Cisco NX-OS Software
Copyright (c) 2ØØ2-2Ø15, Cisco Systems, Inc. All rights reserved.
Cisco Nexus 9000v software ("Cisco Nexus 9000v") and related documentation,
files or other reference materials ("Documentation") are
the proprietary property and confidential information of Cisco
Systems, Inc. ("Cisco") and are protected, without limitation,
pursuant to United States and International copyright and trademark
laws in the applicable jurisdiction which provide civil and criminal
penalties for copying or distribution without Cisco's authorization.

Any use or disclosure, in whole or in part, of the Cisco Nexus 9000v Software
or Documentation to any third party for any purposes is expressly
prohibited except as otherwise authorized by Cisco in writing.
The copyrights to certain works contained herein are owned by other
third parties and are used and distributed under license. Some parts
of this software may be covered under the GNU Public License or the
GNU Lesser General Public License. A copy of each such license is
available at
http://www.gnu.org/licenses/gpl.html and
http://www.gnu.org/Iicenses/lgpl.html
*************************************************************************
* Cisco Nexus 9000v is strictly limited to use for evaluation, demonstration    *
* and NX-OS education. Cisco Nexus 9000v is provided as-is and is not supported  *
* by Cisco's Technical Advisory Center. Any use or disclosure, in whole  *
* or in part of the Cisco Nexus 9000v Software or Documentation to any third     *
* party for any purposes is expressly prohibited except as otherwise    *
* authorized by Cisco in writing.                                       *
*************************************************************************

switch#
```

The vCenter or UCS server connectivity is lost as soon as Cisco Nexus 9000v is up

⚠️

Caution   When connecting a vNIC into a vSwitch or bridge, an incorrect network connection might result in losing the connectivity to your hypervisor server or vCenter on ESXi.

Cisco Nexus 9000v uses vNICs users entered from the KVM/QMEU command line or from a graphical representation on ESXi for networking, either externally or internally within a hypervisor server. The first NIC is always used as the Cisco Nexus 9000v management interface. The subsequent NICs are used as a data port, such as e1/1, e1/2, and up to e1/9.

Connect only the first NIC for the Cisco Nexus 9000v VM as the management interface to your lab LAN physical switch or vSwitch (VM Network) connecting directly to physical switch in the lab (or do not connect any data port vNIC to any physical switch conflicting with your server management connectivity).

Cisco Nexus 9000v data port is not passing traffic in ESXi server

To ensure a smooth operation, specific configuration settings on vSwitch must be enabled:

1. Ensure all instances of vSwitch connecting to Cisco Nexus 9000v be in "Promiscuous Mode" = "Accept", pointing to the UCS server. You can access this option through "Configuration > Properties > Edit" from the vSphere Client.

2. Ensure all instances of vSwitch pass through all VLANs. You can access this option through "Configuration > Properties > Edit" from the vSphere Client.

KVM or QEMU Hypervisor

Multicast on KVM or QEMU Hypervisor

The Cisco Nexus 9000v multicast feature is supported as broadcast. To get this feature work properly, the IGMP multicast snooping must be disabled in this environment on all bridge interfaces.

The following example shows how to disable vxlan_br1, vxlan_br2, vxlan_br3, and vxlan_br4 from the linux prompt.

```
echo 0 > /sys/devices/virtual/net/vxlan_br1/bridge/multicast_snooping

echo 0 > /sys/devices/virtual/net/vxlan_br2/bridge/multicast_snooping

echo 0 > /sys/devices/virtual/net/vxlan_br3/bridge/multicast_snooping

echo 0 > /sys/devices/virtual/net/vxlan_br4/bridge/multicast_snooping
```

VirtualBox

Networking on VirtualBox or Vagrant

To use the dataplane interfaces on VirtualBox or Vagrant, ensure the following:

- The interfaces must be in 'promiscuous' mode.

  In the VirtualBox network settings, select "Allow All" for the Promiscuous mode.

- Ensure all instances of Cisco Nexus 9000v in your topology have unique MAC addresses by using the **show interface mac** command.

VM Fails to Boot up on VirtualBox/Vagrant

Check the following:

- Ensure that enough resources, such as memory or vCPU, are available. Close all applications that consume a significant amount of memory in your PC or server. Check the available free memory.

- Go to the VirtualBox GUI and power down the corresponding VM created from the Vagrant software (long name with tag specified in Vagrant configuration file) or VM created manually from vmdk.

- Make sure that the "serial console" is correctly provisioned.

- Check block disk type and make ensure it is using the SATA controller.

- PowerOn the VM again. The VGA console should appear with the "loader >" prompt. Follow "How to Boot If VM Fails to loader > prompt" troubleshooting topic, and monitor the booting up process through the serial console.

L2FWDER Troubleshooting

Overview

L2fwder is a centralized forwarding component in Cisco Nexus 9000v which performs the following:

- Rx and Tx packets from or to the vmnics

- L2 switching orbridging

  - MAC learning

    - Dynamic MAC learned in packet path

    - Static MACs learned from L2FM via MTS notifications

      - VMACS

      - GW-MAC

- Switching

  - Maintains an array of potential bridge domains

    - Each Bridge domain keeps track of interfaces

      - In forwarding state

      - In Blocked state as an STP state

    - Switching of packets based on the destination MAC in bridge domain based MAC tables

      - Unicast traffic

      - BUM traffic

  - VXLAN Decapsulation

- Punting packets for Layer 3 processing to kstack and netstack

- VXLAN Decap

  - NVE peer-learning by punting the first packet to kstack/netstack for NVE processing.

  - Learning of remote MACs against the remote VTEP interface.

  - Punting ARP packets in case of Layer 3-gateway to kstack/netstack for ARP to learn the remote host routes.

- VXLAN Encap

  - Performed by netstack and packet manger. (Similar to process in hardware, Nexus 9000 platform, for sup-generated packets.)

- VXLAN BGP EVPN

  - In Cisco Nexus 9000v, MAC routes are produced by L2FWDER into L2RIB directly by replacing L2FM, while HMM continues to produce the MAC IP routes into L2RIB similarly as it occurs in Cisco Nexus 9000v.

Commands for L2FWDER

Common Commands	**debug l2fwder ?**	
	err	Control and data path errors.
	fdb	Events over fdb.
	ha	Events from sysmgr.
	ipc	Events over ipc.
	packet	Packet forwarding information.
	pkttrace	Packet trace.
	vxlan	VXLAN plugin.
Clear Commands	**clear mac address-table datapath dynamic**	
	clear mac address-table datapath static	

Troubleshooting RX/TX Path

- Rx-Path

  The logs to monitor for successful pickup from vmnics and sending it to kstack/netstack.

  ```
  l2fwder_get_data_with_wrr(515):Packet received over Driver type 0

  l2fwder_input(67):In   0x0800 78   0 5254.005b.cf97 -> 5254.004c.4e42 Eth1/4

  l2fwder_ethernet_output(196):Driver TUN

  l2fwder_action_send_to_stack(865):l2fwder_action_send_to_stack: tx to ifindex 0 iod 8

  l2fwder_ethernet_output(304):l2fwder_ethernet_output: driver_type[2] pktQ count[1]
  ```

- Tx-Path

  The logs to monitor for successful pickup from tuntap and sending it to kstack/netstack.

  ```
  l2fwder_get_data_with_wrr(515):Packet received over Driver type 2

  l2fwder_ethernet_output(199):Driver ETH

  l2fwder_ethernet_output(251):Out  0x0800 78   0 5254.004c.4e42 -> 5254.005b.cf97 Eth1/4

  l2fwder_ethernet_output(304):l2fwder_ethernet_output: driver_type[0] pktQ count[1]
  ```

- Known Unicast MAC forwarding

  ```
  l2fwder_action_process(934):l2fwder_action_process: process action 1
  ```

```
l2fwder_action_tx_unicast(796):l2fwder_action_tx_unicast: tx to ifindex 1a000600 iod 8
 h_type 0

l2fwder_ethernet_output(199):Driver ETH
```

- MAC database (FDB) lookup related logs for a success lookup (Other than BUM traffic)

```
l2fwder_get_mac_lookup_fwd_info(857):Lookup Result is * 0xPo200(1) ret is 1
l2fwder_get_mac_lookup_fwd_info(897):action ucast
```

- MAC database (FDB) lookup for BUM traffic

Troubleshooting MAC Learning

- Command to check the MAC database in L2FWDER:

```
switch# show system internal l2fwder mac

Legend:

        * - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC

        age - seconds since last seen,+ - primary entry using vPC Peer-Link,

        (T) - True, (F) - False, C - ControlPlane MAC

    VLAN     MAC Address     Type      age     Secure NTFY Ports

---------+----------------+--------+---------+------+----+------------------

G   100    5254.004c.4e42   static   -         F     F    sup-eth1(R)

G   200    5254.004c.4e42   static   -         F     F    sup-eth1(R)

*   200    5254.00c5.9daf   dynamic  00:07:45  F     F       Po200
```

- Event history command to check for static MAC learning:

```
Event:E_DEBUG, length:73, at 930108 usecs after Wed Sep 14 04:13:14 2016

    [117] [23935]: Learning SUCCESS for static 1 mac 52:54:00:c5:9d:af bd 200
```

- Debug log check for dynamic MAC learning:

```
l2fwder_fdb_insert_entry(231):FDB insert for MAC 52:54:00:c5:9d:af bd 200 total entries
 1
```

Troubleshooting Packet Drops in l2fwder/pktmgr/netstack for layer 2/Layer 3 Traffic

- L2FWDER Global Counters:

```
switch(config)# show l2fwder statistics
```

```
Decap stats:

                          RX     DROP

              DCE_CORE     0       0

       2 dot1q decap       0       0

       Sub-interface       0       0

          Switchport  140940       0

           Undefined  210758       0

               Stack  635671       0

       1 dot1q decap       0       0

               VXLAN       0       0

        PORT_CHANNEL  105986       0


Encap stats:

                          TX     DROP

              DCE_CORE     0       0

       2 dot1q decap       0       0

       Sub-interface       0       0

          Switchport  482493       0

           Undefined  211186       0

               Stack       0       0

       1 dot1q decap       0       0

               VXLAN       0       0

        PORT_CHANNEL       0       0


Switching stats:

             Unicast     860

               Flood   29372

           Multicast       0

                Punt   29615

                Drop       0

   LTL Packet Count       0
```

```
Punt stats:

      Packets punted 351004


SMM stats:

MAC                Eth-type   Hit-count

======================================

   0180.c200.0014   0x0000          0

   0180.c200.0015   0x0000          0

   0100.0cdf.dfdf   0x0000          0

   ffff.ffff.ffff   0x0806      29078

   0180.c200.0041   0x22f4          0

   0100.0ccc.cccc   0x0000      13963

   0180.c200.0002   0x0000          0

   0180.c200.0003   0x0000          0

   0180.c200.000e   0x0000          0

   0180.c200.0000   0x0000       1652

   0100.0ccc.cccd   0x0000      97087

   0001.0203.0405   0x0000       1604

   0000.0000.0000   0x0000          0


  Dropped         31

  Consumed    115690

  No Action    29070

  lookup fail  206781


RMM stats:

  Dropped          0

  Consumed    205699

  Rate Limit Dropped      0


VACL stats:

sw-bd         VACL           Hit-count
```

```
=======================================

  Dropped          0

  Consumed         0

  Copy+Fwd         0

  No Action        0


Port-Channel stats:

 VSL Drop Packets        0


MAC Learning Disabled stats:

 Packets recieved on Peer-Link:MAC Learning Disabled     313


Action Flood Stats:

 Port-Channel Split-Horizon Packets     48

 VSL Drop Packets                       0


    Forwarding state of ports in bridge domains


switch# show system internal l2fwder bd

Following is the BD State:-

BD_ID  State  Enh_Fwd  Mode

-----  -----  -------  -----

   1      1       0      0

List of all IODs: 9

List of BLK IODs: 8

----------------------------

BD_ID  State  Enh_Fwd  Mode

-----  -----  -------  -----

  100      0       0      0

List of all IODs: 5 7 16
```

```
        List of BLK IODs: 16
```

Troubleshooting VXLAN BGP EVPN

In the Cisco Nexus 9000v, L2FWDER is the emulated data plane and is responsible for the MAC learning of the connected hosts through source MAC learning.

Note For more information about BGP EVPN, see the *Cisco 9000 Series NX-OS VXLAN Configuration Guide*.

The example in this section considers the following two VTEP end points:

- Leaf0 (VTEP 1) which has hosts with MAC addresses 2222.3333.4444, 000c.2980.d40a in VLAN 1001 and 1002 respectively.

- Leaf1(VTEP 2) which has hosts with MAC addresses 000c.29b9.1375, 000c.29b9.1375 in VLAN 1001 and 1002 respectively.

The following examples shows the MAC and MAC IP route exchange between the two VTEP end points:

- **Local MAC and MAC IP routes in Leaf0**

  - Command to view the source MAC learning:

    ```
    leaf0# show sys int l2fwder mac | inc dynamic
    *  1002    000c.2980.d40a   dynamic   01:13:40   F      F      Eth1/2
    *  1001    2222.3333.4444   dynamic   00:58:38   F      F      Eth1/2
    ```

  - L2FWDER produces the learnt end host MACs as MAC routes in the L2RIB table. The command to display the learnt MAC routes in L2RIB:

    ```
    leaf0# show l2route mac all | inc Local

    Flags -(Rmac):Router MAC (Stt):Static (L):Local (R):Remote (V):vPC link
    1001        2222.3333.4444 Local  L,              0          Eth1/2
    1002        000c.2980.d40a Local  L,              0          Eth1/2
    ```

  - While L2FWDER is responsible for producing the mac routes, the MAC IP route information is produced by Host Mobility Manager(HMM) in L2RIB. The command to display the MAC IP route information in L2RIB is:

    ```
    switch# sh l2route mac-ip all | inc Local
    Flags -(Rmac):Router MAC (Stt):Static (L):Local (R):Remote (V):vPC link
    1001        2222.3333.4444 HMM    --              0          5.1.1.1        Local
    1002        000c.2980.d40a HMM    --              0          5.2.1.1        Local
    ```

  - The MAC IP route information is produced by the Host Mobility Manager (HMM) in L2RIB. The command to display the MAC IP route information is:

    ```
    leaf0# show l2route mac-ip all | inc Local

    Flags -(Rmac):Router MAC (Stt):Static (L):Local (R):Remote (V):vPC link
    1001        2222.3333.4444 HMM    --              0          5.1.1.1        Local
    1002        000c.2980.d40a HMM    --              0          5.2.1.1        Local
    ```

  - The command to display the BGP learnt local MAC and MAC IP routes per VNI is:

```
leaft1# show bgp l2vpn evpn vni-id 5001
BGP routing table information for VRF default, address family L2VPN EVPN

BGP table version is 79, local router ID is 6.1.1.1

Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, *-valid, >-best

Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist,
I-injected
Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath, & - backup
*>l[2]:[0]:[0]:[48]:[2222.3333.4444]:[0]:[0.0.0.0]/216
                      6.1.1.1
              100        32768 i
*>l[2]:[0]:[0]:[48]:[2222.3333.4444]:[32]:[5.1.1.1]/272
                      6.1.1.1
          100        32768 i
```

- **Remote MAC and MAC IP routes in Leaf1**

  - In the remote VTEP, the MAC and the MAC IP route information flows through BGP into the L2RIB, and finally L2FWDER receives the end host MAC reachability information.

```
leaft1# show bgp l2vpn evpn vni-id 5001
BGP routing table information for VRF default, address family L2VPN EVPN
BGP table version is 53, local router ID is 6.2.2.2
Status: s-suppressed, x-deleted, S-stale, d-dampened, h-history, *-valid, >-best
Path type: i-internal, e-external, c-confed, l-local, a-aggregate, r-redist, I-i
njected
Origin codes: i - IGP, e - EGP, ? - incomplete, | - multipath, & - backup

   Network          Next Hop           Metric     LocPrf        Weight      Path
*>i[2]:[0]:[0]:[48]:[2222.3333.4444]:[0]:[0.0.0.0]/216
                      6.1.1.1
              100             0 i
*>i[2]:[0]:[0]:[48]:[2222.3333.4444]:[32]:[5.1.1.1]/272
                      6.1.1.1
              100             0 i

leaf1# show l2route mac all | inc BGP
1001      2222.3333.4444 BGP     SplRcv        0            6.1.1.1
1002      000c.2980.d40a BGP     SplRcv        0            6.1.1.1

eaf1# show l2route mac-ip all | inc BGP
1001      2222.3333.4444 BGP     --            0            5.1.1.1       6.1.1.1
1002      000c.2980.d40a BGP     --            0            5.2.1.1       6.1.1.1

leaf1# show system internal l2fwder mac  | inc nve-peer
* 1002    000c.2980.d40a    static   -         F      F   (0x47000001) nve-peer1
6.1.1.1
* 1001    2222.3333.4444    static   -         F      F   (0x47000001) nve-peer1
6.1.1.1
```

Troubleshooting VXLAN Encap/Decap

The following is in addition to the normal datapath debugging described in other sections:

NVE manager commands to check the provisioning and learning of NVE peers.	**show nve vni**
	show nve peers all
	show ip overlay-traffic

Commands

Counter gauging commands.	**show l2fwder statistics**
	show system internal pktmgr stats
	show ip traffic
Debug commands to capture packet in datapath.	**debug l2fwder** [**packet** | **pktrace** | **error**]
	debug pktmgr [**frame** | **pkt-errors** | **data** | **tunnel**]
	debug ip packet
	tcpdump **Note** (Debug on the vmnic.)

Collecting VM Logs

The Cisco Nexus 9000v uses all code from the physical hardware platform. Therefore, all logging and core files collected from the hardware platform apply to the Cisco Nexus 9000v system. If any issues arise, we recommend that you take a snapshot of the VM or make a copy of the .vmdk or .qcow2 file for further analysis.