Troubleshooting Cisco Nexus 7000 Series Switches

BRKCRS-3144
Session Goal

- To provide you with a thorough understanding of the Cisco Nexus™ 7000 built-in troubleshooting tools and troubleshooting technique of nx-os operating system
- This session will provide summary of the troubleshooting tools and usage most of them through troubleshooting data collection for selected functional areas of the Nexus™ 7000 system but will not focus on deep dive troubleshooting data analysis
- Related sessions: BRKARC-3470
Agenda

- Before You Get Started
  - Traditional Versus NX-OS Troubleshooting Approach
    - Nexus 7000 Built-in Troubleshooting Tools
    - Nexus 7000 Module and Forwarding Engine Architecture Overview
  - Troubleshooting
    - CPU
    - Control-Plane – CoPP
    - Memory Utilization
    - vPC
    - Unicast Layer 2 and Layer 3 Forwarding and ARP
    - Multicast Layer 2 and Layer 3 Forwarding
    - Switch Fabric
    - ACL
    - QoS
Before You Get Started
Traditional Versus NX-OS Troubleshooting Approach

- Troubleshooting objective is to find out
  - What is broken or not functioning as expected
  - Why is it broken and is there a workaround
  - What triggered unexpected or broken behavior

- Successful and effective troubleshooting requires:
  - Proper and accurate problem description
  - Platform specific knowledge
    - Hardware architecture and capabilities
    - Data path through the system
    - Feature hardware dependency, supported features and their combination
    - Knowledge of available troubleshooting tools and their usage
  - Topology knowledge and data path through topology
  - Cisco hardware and vendor devices interaction knowledge
Before You Get Started
Traditional Versus NX-OS Troubleshooting Approach (Cont.)

Traditional Approach

- Problem triggered
- Problem detected
  - Cisco TAC engaged
- Problem identified
  - TAC initial data collection
  - TAC recreate additional data
  - Special Code additional data
- Problem not identified
  - TAC recreate additional data
  - Problem not identified
  - No sufficient data
  - Problem identified
- Problem resolved
  - Case closed
 NX-OS Approach

Problem triggered → Problem detected → *Data* *Collection*

*Data* *Collection* → Cisco TAC engaged → 90% cases problem identified → Problem identified

IF

Problem not identified no sufficient data → TAC recreate additional data

Problem resolved Case closed → Special Code additional data

Problem identified → Problem triggered

90% cases problem identified → Problem not identified no sufficient data

Case closed
Before You Get Started
Traditional Versus NX-OS Troubleshooting Approach (Cont.)

### Suggestions
- Identify **detection** and **trigger** time as accurately as possible to set ‘good’ start up point for collected data search and analysis
- Minimize **delta time** between trigger/detection time and **data collection** time
- Try to recall all activities before trigger/detection time
- Get proficient as much as possible with built-in **tool box**
- Get familiar with specific feature troubleshooting cli, feature show tech-support output for **on-the-fly** troubleshooting and analysis

### Remember
- Internal data logs have **limited size**, adjust them ahead of time for relevant features you have deployed
- Even max-ed log size may not prevent data wrap up
- Use configuration rollback or other configuration backup method while troubleshooting and making configuration changes
- Forensic data survives reload or switchover via ‘**Onboard logging**’, ‘**accounting-log**’, ‘**nvram**’
Agenda

- Before You Get Started
  - Traditional Versus NX-OS Troubleshooting Approach
  - Nexus 7000 Built-in Troubleshooting Tools
  - Nexus 7000 I/O Module and Forwarding Engine Architecture Overview

- Troubleshooting
  - CPU
  - Control-Plane – CoPP
  - Memory Utilization
  - vPC
  - Unicast Layer 2 and Layer 3 Forwarding and ARP
  - Multicast Layer 2 and Layer 3 Forwarding
  - Switch Fabric
  - ACL
  - QoS
Built-In Troubleshooting Tools
Make Troubleshooting Easier and more Effective — Almost Fun to Do 😊

- Powerful `show` cli output
  - Standard documented cli
    - Platform independent (PI) and dependent (PD) output
    - Hardware keyword indicates platform hardware specific output
  - Undocumented ‘engineering’ cli
    - Internal keyword
    - No XML or SNMP support

- Extensive feature and software component `event-history` logging
  - Permanent engineering debugs output of process Finite State Machine (FSM)
    - `show ip ospf internal event-history cli`
    - `show system internal pixm event-history msgs`

- Extensive system activity logging to dedicated `logflash` with filtering to display only ‘what I want to see’
  - `show logging logfile last X | start-time <time> | start-seqn X`
  - There is implicit ‘all’ if none of the filters is used
Built-In Troubleshooting Tools
Make Troubleshooting Easier and more Effective — Almost Fun to Do 😊

- Onboard logging, accounting log logging (config and exec)
  - Forensic data surviving reload and switchover
  - Hardware component events and manipulation activity
  - Use it to ‘recall’ all activity around ‘trigger and detection’ time

- GOLD system
  - Hardware components health monitoring

- Beacon feature
  - Hardware component LED FLASH locator
  - Useful with new installs, cabling and replacements

- Ping, Traceroute
- Span, Netflow, XML, EEM
- Build in Linux tools e.g. grep, egrep, last, less, sed, wc, sort, diff, redirect, exclude, include, pipe etc
Built-In Troubleshooting Tools
Make Troubleshooting Easier and more Effective — Almost Fun to Do 😊

- Traditional feature related debugs e.g.
  
  ```
  debug ip packet protocol igmp, debug ipv6 icmp, debug icmp
  ```

- NX-OS debugs with `debug-filter`, e.g.
  
  ```
  debug-filter ip packet direction inbound
  ```

- Easy to read asic counters and registers
  - Software copy not clear-on-read, must use clear cli to clear them
  - Comprehensive per module, asic, port, counter category filtering

- Embedded Logic Analyzer Module (ELAM capture)
  - Detailed frame internal header information
  - Built-in `wireshark` analyzer capturing mgmt interface and CPU traffic
  - May be combined with `acl-log` with no performance degradation, as traffic is still forwarded or dropped in hardware

**TIP #1:** Use `show cli syntax | grep -i <keyword_of_interest> grep –v <show>`

**TIP #2:** For configuration cli, you must go to `configuration mode` first
Agenda

- Before You Get Started
  - Traditional Versus NX-OS Troubleshooting Approach
  - Nexus 7000 Built-in Troubleshooting Tools
  - Nexus 7000 I/O Module and Forwarding Engine Architecture Overview

- Troubleshooting
  - CPU,
  - Control-Plane – CoPP
  - Memory Utilization
  - vPC
  - Unicast Layer 2 and Layer 3 Forwarding and ARP
  - Multicast Layer 2 and Layer 3 Forwarding
  - Switch Fabric
  - ACL
  - QoS
Nexus 7000 Module Architecture

- Main I/O Module functional blocks
  - Ethernet MAC
  - Replication Engine
  - Virtual Output Queue
  - Fabric State number 1 or 3
  - CTS – Security
  - Forwarding engine

- `show hardware internal dev-port-map`
  - Provides physical ‘data path’ through the I/O module
  - Maps front panel port number to all other ASIC/devices instances
  - Zero based numbering
  - Eliminates need for I/O Module block diagram during t’shooting

Note: All I/O Modules block diagrams and ‘show hardware internal dev-port-map’ cli output available via hidden slides for reference
## Nexus 7000 Module Architecture

New Hardware Support Features—LED Locator

- **LED locator (Beacon) New hardware feature**
  - Makes it easy to locate any hardware component for maintenance
  - Operated from ‘default’ vdc only
  - Once turned on, stays on till turned off

### LED Locator (Beacon)

- New hardware feature
- Makes it easy to locate any hardware component for maintenance
- Operated from ‘default’ vdc only
- Once turned on, stays on till turned off

### Nexus 7000 Module Architecture

#### New Hardware Support Features

- LED Locator

#### LED Locator Command Examples

**N7K-1# locator-led ?**

- chassis Blink chassis led
- fan Blink Fan led
- module Blink module led
- powersupply Blink powersupply led
- xbar Xbar

**N7K-1# show locator-led status**

<table>
<thead>
<tr>
<th>Component</th>
<th>Locator LED Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chassis</td>
<td>off</td>
</tr>
<tr>
<td>Module 1</td>
<td>ON</td>
</tr>
<tr>
<td>Module 2</td>
<td>off</td>
</tr>
<tr>
<td>Module 4</td>
<td>off</td>
</tr>
<tr>
<td>Module 5</td>
<td>off</td>
</tr>
<tr>
<td>Module 6</td>
<td>off</td>
</tr>
<tr>
<td>Module 7</td>
<td>Not powered up.</td>
</tr>
<tr>
<td>Module 10</td>
<td>off</td>
</tr>
<tr>
<td>Xbar 1</td>
<td>off</td>
</tr>
<tr>
<td>Xbar 2</td>
<td>ON</td>
</tr>
<tr>
<td>Xbar 3</td>
<td>off</td>
</tr>
</tbody>
</table>

### Module 1 Front Panel LED

- Module 1 Front Panel LED is flashing

---

© 2011 Cisco and/or its affiliates. All rights reserved. Cisco Public
### Nexus 7000 Module Architecture (Cont.)

**New Hardware Support Features — IDS**

N7K-1# `show hardware forwarding ip verify module 1`

<table>
<thead>
<tr>
<th>IPv4 and v6 IDS Checks</th>
<th>Status</th>
<th>Packets Failed</th>
</tr>
</thead>
<tbody>
<tr>
<td>address source broadcast</td>
<td>Enabled</td>
<td>0</td>
</tr>
<tr>
<td>address source multicast</td>
<td>Enabled</td>
<td>0</td>
</tr>
<tr>
<td>address destination zero</td>
<td>Enabled</td>
<td>0</td>
</tr>
<tr>
<td>address identical</td>
<td>Disabled</td>
<td></td>
</tr>
<tr>
<td>address reserved</td>
<td>Disabled</td>
<td></td>
</tr>
<tr>
<td>address class-e</td>
<td>Disabled</td>
<td></td>
</tr>
<tr>
<td>checksum</td>
<td>Enabled</td>
<td>0</td>
</tr>
<tr>
<td>protocol</td>
<td>Enabled</td>
<td>0</td>
</tr>
<tr>
<td>fragment</td>
<td>Disabled</td>
<td></td>
</tr>
<tr>
<td>length minimum</td>
<td>Enabled</td>
<td>0</td>
</tr>
<tr>
<td>length consistent</td>
<td>Enabled</td>
<td>0</td>
</tr>
<tr>
<td>length maximum max-frag</td>
<td>Enabled</td>
<td>0</td>
</tr>
<tr>
<td>length maximum udp</td>
<td>Disabled</td>
<td></td>
</tr>
<tr>
<td>length maximum max-tcp</td>
<td>Enabled</td>
<td>0</td>
</tr>
<tr>
<td>tcp flags</td>
<td>Disabled</td>
<td></td>
</tr>
<tr>
<td>tcp tiny-frag</td>
<td>Enabled</td>
<td>0</td>
</tr>
<tr>
<td>version</td>
<td>Enabled</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IPv6 IDS Checks</th>
<th>Status</th>
<th>Packets Failed</th>
</tr>
</thead>
<tbody>
<tr>
<td>length consistent</td>
<td>Enabled</td>
<td>0</td>
</tr>
<tr>
<td>length maximum max-frag</td>
<td>Enabled</td>
<td>0</td>
</tr>
<tr>
<td>length maximum udp</td>
<td>Disabled</td>
<td></td>
</tr>
<tr>
<td>length maximum max-tcp</td>
<td>Enabled</td>
<td>0</td>
</tr>
<tr>
<td>tcp tiny-frag</td>
<td>Enabled</td>
<td>0</td>
</tr>
</tbody>
</table>

- **Intrusion Detection System (IDS)** performed by forwarding engine hardware
- **Global system wide** hardware feature
- Some features require specific check to be turned off to function properly (**BFD** requires ‘address indentical’ check to be disabled)
- Some IDS checks are disabled by default
- Use `hardware ip verify cli` in default vDC to modify settings
- To clear IDS counters use cli `clear hardware forwarding ip verify all`
Nexus 7000 Module Architecture (Cont.)

New Hardware Support Features — Onboard Logging

- Onboard nvram logging
  - 32M flash per module
  - Makes it easy to track I/O module
    - Bootup time
    - Exceptions and asic component crashes, temperature
    - Insertion, removal and reset reasons
  - Can be accessed from any vdc
    - Becomes very powerful combined with accounting log
  - logging content configurable per module

N7K-1# show logging onboard module 7 boot-uptime
-----------------------------------------------
Module:  7
-----------------------------------------------

Sat Mar 26 09:15:32 2011: Card Uptime Record
-----------------------------------------------
Uptime: 27, 0 days 0 hour(s) 0 minute(s) 27 second(s)
Reset Reason: Module is powered down or power cycled (72)
Reset Reason SW: Unknown (0)
Reset Reason HW: System reset by active sup (by writing to PMFPGA regs) (100)
Card Mode..........................: Runtime
## Nexus 7000 Module Architecture (Cont.)

### New Hardware Support Features

```
N7K-1# show accounting log start-time 2010 Mar 26 09:00:00
Sat Mar 26 09:14:21 2011:type=update:id=172.20.2.212@pts/1:user=admin:cmd=clear logging onboard module 7
   (SUCCESS)
Sat Mar 26 09:14:31 2011:type=update:id=172.20.2.212@pts/1:user=admin:cmd=Powering down module 7
Sat Mar 26 09:14:31 2011:type=update:id=172.20.2.212@pts/1:user=admin:cmd=configure terminal ; poweroff module 7
   (SUCCESS)
Sat Mar 26 09:14:43 2011:type=update:id=172.20.2.212@pts/1:user=admin:cmd=Powering up module 7
Sat Mar 26 09:14:43 2011:type=update:id=172.20.2.212@pts/1:user=admin:cmd=configure terminal ; no poweroff
   module 7 (SUCCESS)
```

```
N7K-1# show logging onboard module 7 ?
   <CR>
   > Redirect it to a file
   >> Redirect it to a file in append mode
   boot-uptime Boot-uptime
   counter-stats Show OBFL counter statistics
   device-version Device-version
   endtime Show OBFL logs till end time mm/dd/yy-HH:MM:SS
   environmental-history Environmental-history
   error-stats Show OBFL error statistics
   exception-log Exception-log
   internal Show Logging Onboard Internal
   interrupt-stats Interrupt-stats
   obfl-history Obfl-history
   stack-trace Stack-trace
   starttime Show OBFL logs from start time mm/dd/yy-HH:MM:SS
   status Status
   | Pipe command output to filter
```

- Any module manipulation can be tracked down
- Allows per asic exception-log filtering
- `internal reset-reason` option provides 'upgrade history'
Agenda

- Before You Get Started
  - Traditional Versus NX-OS Troubleshooting Approach
  - Nexus 7000 Module and Forwarding Engine Architecture Overview
  - Built-in Troubleshooting Tools

- Troubleshooting
  - CPU
  - Control-Plane - CoPP
  - Memory Utilization
  - vPC
  - Unicast Layer 2 and Layer 3 Forwarding and ARP
  - Multicast Layer 2 and Layer 3 Forwarding
  - Switch Fabric
  - ACL
  - QoS
Troubleshooting
CPU

- High CPU utilization is **not automatically** problem indication!
  - NEXUS 7000 is dual core linux based system with robust preemptive scheduler (one functional unit for both rp and sp)
  - **Strict** control-plane and data-plane separation
  - Scheduler assures **fair access** to CPU for all processes
  - Lower level processes (drivers) run in FIFO or non-preemptive mode

- Common reasons for high CPU
  - Excessive CPU bound traffic, control-plane churn
  - Access-list processing, hardware programming
  - Misbehaving process

- Suggested troubleshooting to get started
  - `show hardware internal cpu-mac inband stat`
  - `show system internal processes cpu`
  - wireshark
Troubleshooting
CPU — Supervisor, General Health Check

N7k-3-VDC3# show system resources
Load average: 1 minute: 0.64   5 minutes: 1.08   15 minutes: 1.30
Processes : 3912 total, 2 running
CPU states : 4.5% user,  5.0% kernel,  90.5% idle
Memory usage: 4115232K total,  3434268K used,  680964K free

N7k-3-VDC3# show processes cpu history

1 2 111 1111211233 1 1 111 1 1 1 6 112 1 1 21132 1 111 123
919275058862141899918384800583739174756080779143297264026770

CPU% per second (last 60 seconds)
# = average CPU%

How many processes were scheduled to run in average per whole system in last 1, 5 and 15 minutes
How much of CPU cycles are used by user configured processes and kernel processes
Output IS calibrated for 2 cores

CPU utilization 60 seconds ago
# Troubleshooting

## CPU — Supervisor, General Health Check

N7K-3-VDC3# `show system internal processes cpu`

```
top - 14:01:06 up 21 days, 15:35, 4 users, load average: 0.77, 0.73, 1.07
Tasks: 3257 total, 1 running, 422 sleeping, 0 stopped, 2834 zombie
Cpu(s): 5.8%us, 6.0%sy, 0.1%ni, 84.1%id, 0.4%wa, 0.1%hi, 3.4%si, 0.0%st
Mem: 4115232k total, 3875988k used, 239244k free, 82400k buffers
Swap: 0k total, 0k used, 0k free, 1817776k cached
```

<table>
<thead>
<tr>
<th>PID</th>
<th>USER</th>
<th>PR</th>
<th>NI</th>
<th>VIRT</th>
<th>RES</th>
<th>SHR</th>
<th>S</th>
<th>%CPU</th>
<th>%MEM</th>
<th>TIME+</th>
<th>COMMAND</th>
</tr>
</thead>
<tbody>
<tr>
<td>22683</td>
<td>root</td>
<td>20</td>
<td>0</td>
<td>182m</td>
<td>63m</td>
<td>14m</td>
<td>S</td>
<td>93.7</td>
<td>1.6</td>
<td>636:17.84</td>
<td>netstack</td>
</tr>
<tr>
<td>29391</td>
<td>admin#03</td>
<td>20</td>
<td>0</td>
<td>5364</td>
<td>3312</td>
<td>1140</td>
<td>R</td>
<td>22.3</td>
<td>0.1</td>
<td>0:00.30</td>
<td>top</td>
</tr>
<tr>
<td>3892</td>
<td>root</td>
<td>20</td>
<td>0</td>
<td>164m</td>
<td>54m</td>
<td>23m</td>
<td>S</td>
<td>6.0</td>
<td>1.3</td>
<td>1095:00</td>
<td>netstack</td>
</tr>
<tr>
<td>4149</td>
<td>root</td>
<td>20</td>
<td>0</td>
<td>111m</td>
<td>41m</td>
<td>19m</td>
<td>S</td>
<td>4.5</td>
<td>1.0</td>
<td>994:43.26</td>
<td>stp</td>
</tr>
<tr>
<td>3175</td>
<td>root</td>
<td>20</td>
<td>0</td>
<td>78100</td>
<td>19m</td>
<td>17m</td>
<td>S</td>
<td>3.0</td>
<td>0.5</td>
<td>175:07.02</td>
<td>diagmgr</td>
</tr>
<tr>
<td>23028</td>
<td>root</td>
<td>20</td>
<td>0</td>
<td>101m</td>
<td>23m</td>
<td>9968</td>
<td>S</td>
<td>3.0</td>
<td>0.6</td>
<td>598:14.57</td>
<td>stp</td>
</tr>
<tr>
<td>3181</td>
<td>root</td>
<td>20</td>
<td>0</td>
<td>77684</td>
<td>4564</td>
<td>3352</td>
<td>S</td>
<td>1.5</td>
<td>0.1</td>
<td>0:30.35</td>
<td>securityd</td>
</tr>
<tr>
<td>3591</td>
<td>root</td>
<td>20</td>
<td>0</td>
<td>222m</td>
<td>13m</td>
<td>7132</td>
<td>S</td>
<td>1.5</td>
<td>0.3</td>
<td>0:09.61</td>
<td>igmp</td>
</tr>
<tr>
<td>4753</td>
<td>root</td>
<td>20</td>
<td>0</td>
<td>162m</td>
<td>45m</td>
<td>16m</td>
<td>S</td>
<td>1.5</td>
<td>1.1</td>
<td>34:59.22</td>
<td>netstack</td>
</tr>
<tr>
<td>1</td>
<td>root</td>
<td>20</td>
<td>0</td>
<td>1988</td>
<td>612</td>
<td>532</td>
<td>S</td>
<td>0.0</td>
<td>0.0</td>
<td>0:16.32</td>
<td>init</td>
</tr>
<tr>
<td>2</td>
<td>root</td>
<td>15</td>
<td>-5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>S</td>
<td>0.0</td>
<td>0.0</td>
<td>0:00.00</td>
<td>kthreadd</td>
</tr>
<tr>
<td>3</td>
<td>root</td>
<td>RT</td>
<td>-5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>S</td>
<td>0.0</td>
<td>0.0</td>
<td>0:00.22</td>
<td>migration/0</td>
</tr>
<tr>
<td>4</td>
<td>root</td>
<td>15</td>
<td>-5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>S</td>
<td>0.0</td>
<td>0.0</td>
<td>5:49.32</td>
<td>ksoftirqd/0</td>
</tr>
</tbody>
</table>

---

**Use X | no-more, where X is interval in seconds to get more snapshots**

**Equivalent of Linux TOP monitoring tool output showing system processes across all vDCs**

**Use it to cross check accuracy of ‘show system resources’ output**

**Output is NOT calibrated for 2 cores so it would be expected to see 2 processes using 100% CPU**

**Output show processes from all vDCs**
Troubleshooting
CPU — Supervisor, General Health Check

N7K-3-VDC3# show processes cpu | egrep "PID|--|ospf"

<table>
<thead>
<tr>
<th>PID</th>
<th>Runtime (ms)</th>
<th>Invoked</th>
<th>uSecs</th>
<th>1Sec</th>
<th>Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>9337</td>
<td>102</td>
<td>72</td>
<td>1418</td>
<td>0.0%</td>
<td>ospfv3</td>
</tr>
<tr>
<td>22916</td>
<td>118</td>
<td>62</td>
<td>1905</td>
<td>13.1%</td>
<td>ospf</td>
</tr>
</tbody>
</table>

N7K-3-VDC3# show system internal sysmgr service pid 22916
Service "__inst_001__ospf" ("ospf", 58):
  UUID = 0x41000119, PID = 22916, SAP = 320
  Restart count: 1
The service never crashed since the last reboot.
  Tag = 6467
  Plugin ID: 1

N7K-3-VDC3# show system internal sysmgr service name ospfv3 tag 8893
Service "__inst_001__ospfv3" ("ospfv3", 59):
  UUID = 0x4100011A, PID = 9337, SAP = 328
  Restart count: 2
The service never crashed since the last reboot.
  Tag = 8893
  Plugin ID: 1

Useful process level details

For testing purposes, process was manually restarted using ‘restart ospfv3 8893’ cli
# Troubleshooting

## CPU — Traffic Caused High CPU Utilization and Control-Plane Instability

- Typical datacenter traffic causing high cpu utilization
  - arp, nd (ipv6)
  - dhcp traffic
  - Glean traffic (no arp or nd)
  - Malicious traffic to 224.0.0.0/24 subnet
  - Fragments or malicious L2 mcast or ‘other’ traffic

- CoPP provides more granular targeted CPU protection whereas RLs work better with traffic categories where specifics (sip/dip) may not be known

- CPU protection via CoPP policers
- CPU protection via L2/L3 hardware rate-limiters (RL)
- CoPP and RL default settings may need tweaking based on network requirement specifics
  - Both are configured/enabled per M1 I/O Module
  - Total rp bound traffic allowed is sum across all M1 I/O Modules

- Going ‘over-protective’ is not a solution
- CoPP and RL tweaking must allow ‘reasonable’ protocol convergence and CPU protection at the same time
Troubleshooting
CPU — Traffic Caused High CPU Utilization and Control-Plane Instability

N7K-1-VDC2# show system resources
Load average:  1 minute: 2.92   5 minutes: 2.38   15 minutes: 2.27
Processes :  1267 total, 4 running
CPU states :  34.0% user,  42.5% kernel,  23.5% idle
Memory usage:  4115232K total,  3638780K used,  476452K free

N7K-1-VDC2# show processes cpu sort

<table>
<thead>
<tr>
<th>PID</th>
<th>Runtime (ms)</th>
<th>Invoked</th>
<th>uSecs</th>
<th>1Sec</th>
<th>Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>3981</td>
<td>127</td>
<td>276</td>
<td>462</td>
<td>43.2%</td>
<td>ospf</td>
</tr>
<tr>
<td>3841</td>
<td>267</td>
<td>78</td>
<td>3427</td>
<td>16.4%</td>
<td>netstack</td>
</tr>
<tr>
<td>2941</td>
<td>34146488</td>
<td>7377876</td>
<td>4628</td>
<td>0.9%</td>
<td>platform</td>
</tr>
<tr>
<td>3982</td>
<td>118</td>
<td>245</td>
<td>485</td>
<td>0.9%</td>
<td>ospfv3</td>
</tr>
</tbody>
</table>

192.251.19.22 on Vlan19 from INIT to DOWN, DEADTIME

2011 Mar 26 15:38:56.584 N7K-1-VDC2 %OSPF-5-NBRSTATE: ospf-6467 [3981] Process 6467, Nbr
192.251.19.22 on Vlan19 from DOWN to INIT, HELLORCVD

192.251.19.22 on Vlan19 from INIT to DOWN, DEADTIME

192.251.19.22 on Vlan19 from DOWN to INIT, HELLORCVD
Troubleshooting
CPU — Traffic Caused High CPU Utilization and Control-Plane Instability

```
N7K-1# show hardware internal cpu-mac inband stats | egrep " Rx|Tx|counters|Throttle|Tick|rate|total|good|XOFF p|XON p"
```

<table>
<thead>
<tr>
<th>RMON counters</th>
<th>Rx</th>
<th>Tx</th>
</tr>
</thead>
<tbody>
<tr>
<td>total packets</td>
<td>779905245</td>
<td>1421785114</td>
</tr>
<tr>
<td>good packets</td>
<td>779905245</td>
<td>1421650279</td>
</tr>
<tr>
<td>good octets (hi)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>good octets (low)</td>
<td>172303021767</td>
<td>192965708376</td>
</tr>
<tr>
<td>total octets (hi)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>total octets (low)</td>
<td>172302724342</td>
<td>192974265660</td>
</tr>
<tr>
<td>XON packets</td>
<td>0</td>
<td>67627</td>
</tr>
<tr>
<td>XOFF packets</td>
<td>0</td>
<td>67208</td>
</tr>
</tbody>
</table>

Interrupt counters
Error counters
Throttle statistics
Throttle interval ............ 2 * 100ms
Packet rate limit ............ 32000 pps
Tick counter ................. 12414130
Rx packet rate (current/max) 4993 / 20296 pps
Tx packet rate (current/max) 60 / 3474 pps

MAC counters
```
<table>
<thead>
<tr>
<th>MAC0 (R2D2)</th>
<th>MAC1 (CPU)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rx</td>
<td>Tx</td>
</tr>
<tr>
<td>total packets</td>
<td>779905246</td>
</tr>
<tr>
<td>total bytes</td>
<td>2470922140</td>
</tr>
<tr>
<td>XOFF packets</td>
<td>5447</td>
</tr>
<tr>
<td>XOFF packets</td>
<td>7590855</td>
</tr>
<tr>
<td>XON packets</td>
<td>0</td>
</tr>
</tbody>
</table>
```

Challenge is how to identify offending traffic type and its source.
Troubleshooting
CPU — Traffic Caused High CPU Utilization and Control-Plane Instability

N7K-1-VDC2# show system internal pktmgr interface vlan 64
Vlan64, ordinal: 117
SUP-traffic statistics: (sent/received)
   Packets: 3771848 / 40687558
   Bytes: 304360445 / 36018498390
   Instant packet rate: 0 pps / 4951 pps
Packet statistics:
   Tx: Unicast 3747450, Multicast 24381
       Broadcast 17
   Rx: Unicast 3751291, Multicast 36936251
       Broadcast 16

Use this cli first without specific interface to identify ‘offending’ one with highest rate
Alternatively use ‘show system internal pktmgr internal vdc inband’ which identifies vDC interfaces and number of packet sent to CPU

With high rate traffic, screen would get flooded with debug output. Be ready to type ‘undebug all’ or alternatively use ‘debug logfile <file>’. Logfile is file written to dedicated log flash where debug output is redirected to.

N7K-1-VDC2# debug-filter pktmgr vlan 64
N7K-1-VDC2# show debug-filter all
ddebug-filter pktmgr vlan 64

N7K-1-VDC2# debug pktmgr frame
2011 Mar 26 21:22:30.599670 netstack: In  Vlan 64 0x0800 992 7 0000.1301.1301 ->
   0100.5e00.0005 Vlan64

N7K-1-VDC2# show ip arp vlan 64 | i 0000.1301.1301
N7K-1-VDC2# show mac address-table address 0000.1301.1301 vlan 64
Legend:
   * - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
   age - seconds since last seen,+ - primary entry using vPC Peer-Link

<table>
<thead>
<tr>
<th>VLAN</th>
<th>MAC Address</th>
<th>Type</th>
<th>age</th>
<th>Secure</th>
<th>NTFY Ports/SWID.SSID.LID</th>
</tr>
</thead>
<tbody>
<tr>
<td>* 64</td>
<td>0000.1301.1301</td>
<td>dynamic</td>
<td>0</td>
<td>F</td>
<td>F</td>
</tr>
</tbody>
</table>

No ARP entry??
Offending host mac
Offending port
Troubleshooting
CPU — Traffic Caused High CPU Utilization and Control-Plane Instability

N7K-1-VDC2# debug-filter ip ospf interface vlan 64
N7K-1-VDC2# debug logfile offending_traffic
N7K-1-VDC2# debug ip ospf packets
N7K-1-VDC2# undebug all
N7K-1-VDC2#

Alternative method to capture offending traffic to file using nx-os debug-filter
Logfile is 4MB circular buffer

N7K-1-VDC2# show debug logfile offending_traffic
  rid:0.0.0.0 area:0.0.0.0 crc:0xfdd2 aut:0 aukid:0 from 192.253.64.254/Vlan64
2011 Mar 26 23:33:25.992780 ospf: 6467 [3981] Invalid src address 192.253.64.254, should not be seen on Vlan64
  rid:0.0.0.0 area:0.0.0.0 crc:0xfdd2 aut:0 aukid:0 from 192.253.64.254/Vlan64
2011 Mar 26 23:33:25.992966 ospf: 6467 [3981] Invalid src address 192.253.64.254, should not be seen on Vlan64

N7K-1-VDC2# copy log:offending_traffic bootflash:offending_traffic
Copy complete, now saving to disk (please wait)...  
N7K-1-VDC2# gzip bootflash:offending_traffic
N7K-1-VDC2# dir bootflash: | grep offending_traffic
 86968  Mar 26 23:48:33 2011  offending_traffic.gz

N7K-1-VDC2# copy bootflash:offending_traffic.gz scp://cisco@172.20.2.212/home/cisco vrf management
Password: offending_traffic.gz
100%  85KB
Copy complete, now saving to disk (please wait)...

Zip it to smaller size and ship it to your server for further analysis
Agenda

- Before You Get Started
  - Traditional Versus NX-OS Troubleshooting Approach
  - Nexus 7000 Module and Forwarding Engine Architecture Overview
  - Built-in Troubleshooting Tools

- Troubleshooting
  - CPU
  - Control-Plane - CoPP
  - Memory Utilization
  - vPC
  - Unicast Layer 2 and Layer 3 Forwarding and ARP
  - Multicast Layer 2 and Layer 3 Forwarding
  - Switch Fabric
  - ACL
  - QoS
Troubleshooting
Control-Plane — CoPP and RL CPU Protection, Default CoPP modification

N7K-1# show policy-map interface control-plane module 2 | egrep "service-policy|critical|ospf|police cir 39600|malicious"

```
service-policy input: copp-system-policy
   class-map copp-system-class-critical (match-any)
      match access-grp name copp-system-acl-ospf
      match access-grp name copp-system-acl-ospf6
      police cir 39600 kbps , bc 250 ms
```

N7K-1# show class-map type control-plane copp-system-class-critical | egrep class|ospf

```
class-map type control-plane match-any copp-system-class-critical
   match access-grp name copp-system-acl-ospf
   match access-grp name copp-system-acl-ospf6
```

N7K-1# show ip access-lists copp-system-acl-ospf

```
IP access list copp-system-acl-ospf
   10 permit ospf any any
```

- **What is the solution?**
  - Modify `copp-system-acl-ospf` to permit only specific IPs or subnets
  - Create `copp-system-acl-224malicious` access-list
  - Add `copp-system-class-malicious` class with zero policer
  - Same approach can be used for any offending 224.0.0.0/24 traffic

- Keep in mind CoPP is applied for all vdc's but can only be modified from default vDC
Troubleshooting
Control-Plane — CoPP and RL CPU Protection

N7K-1# show ip access-lists copp-system-acl-ospf
IP access list copp-system-acl-ospf
  10 permit ospf any any
  20 permit ip 40.9.0.0/16 224.0.0.5/32
  30 permit ip 40.9.0.0/16 224.0.0.6/32
  40 permit ip 192.251.0.0/16 224.0.0.5/32
  50 permit ip 192.251.0.0/16 224.0.0.6/32
  60 permit ip 172.6.66.0/24 224.0.0.5/32
  70 permit ip 172.6.66.0/24 224.0.0.6/32
  80 permit ip 12.0.0.0/8 224.0.0.5/32
  90 permit ip 12.0.0.0/8 224.0.0.6/32

Add flapping neighbor subnet and other subnets where ospf should be protected

N7K-1# show ip access-lists copp-system-acl-224malicious
IP access list copp-system-acl-224malicious
  10 permit ip any 224.0.0.0/24

Create new access-list

N7K-1# show policy-map interface control-plane module 2 | egrep "service-policy|critical|ospf|police cir 39600|malicious|police cir 1"
service-policy input: copp-system-policy
class-map copp-system-class-critical (match-any)
  match access-grp name copp-system-acl-ospf
  match access-grp name copp-system-acl-ospf6
  police cir 39600 kbps , bc 250 ms
class-map copp-system-class-malicious (match-any)
  match access-grp name copp-system-acl-224malicious
  police cir 1 bps , bc 200 ms

Add new class right before last class-default
Zero rate policer to block all malicious traffic
Troubleshooting
Control-Plane — CoPP and RL CPU Protection

N7K-1# show policy-map interface control-plane module 2 class copp-system-class-malicious

control Plane
  service-policy input: copp-system-policy
    class-map copp-system-class-malicious (match-any)
      match access-grp name copp-system-acl-224malicious
      police cir 1 bps , bc 200 ms
    module 2:
      conformed 0 bytes; action: drop
      violated 1799505072 bytes; action: drop

N7K-1# show policy-map interface control-plane module 1 class copp-system-class-malicious

control Plane
  service-policy input: copp-system-policy
    class-map copp-system-class-malicious (match-any)
      match access-grp name copp-system-acl-224malicious
      police cir 1 bps , bc 200 ms
    module 1:
      conformed 0 bytes; action: drop
      violated 0 bytes; action: drop

Depending on how routing is done in vpc configuration, same CoPP tweaking may be required on both vpc peers.
**Troubleshooting**

Control-Plane — CoPP and RL CPU Protection, RL programming check

---

**N7K-1# show hardware internal forwarding l3 asic rate-limiter layer-3-control detail module 1**

- **Dev-id:** 0
- **Rate-limiter configuration:** layer-3 control
  - **Enabled:** 0
  - **Packets/sec:** 10000
  - **Packet burst:** 100 [burst period of 1 msec]

**Layer-3 Control Rate-limiter is disabled to allow CoPP to deal with 224.0.0.0/24 malicious traffic. CoPP provides better scalability in this protection.**

**N7K-1# show hardware internal forwarding l2 asic rate-limiter layer-3-glean detail module 1**

- **Device:** 1
- **Rate-Limiter configuration:** layer-3 glean
  - **Enabled:** 1
  - **Packets/sec:** 2500

**Match fields:**
- **Cap1 bit:** 0
- **Cap2 bit:** 0
- **DI select:** 1
  - **DI:** 0x401
- **Flood bit:** 0

**Replaced result fields:**
- **Cap1 bit:** 0
- **Cap2 bit:** 0
  - **DI:** 0x7fff

**N7K-1# show system internal pixm info ltl 0x401**

- **0x0401 is in SUP In-band LTL range**

**Bit bucket LTL index (any frame sent to not configured LTL index will be dropped)**

**N7K-1# show system internal pixm info ltl 0x7fff**

- **0x7fff is not configured**

---
Agenda

- **Before You Get Started**
  - Traditional Versus NX-OS Troubleshooting Approach
  - Nexus 7000 Module and Forwarding Engine Architecture Overview
  - Built-in Troubleshooting Tools

- **Troubleshooting**
  - CPU
  - Control-Plane
  - Memory Utilization
  - vPC
  - Unicast Layer 2 and Layer 3 Forwarding and ARP
  - Multicast Layer 2 and Layer 3 Forwarding
  - Switch Fabric
  - ACL
  - QoS
Troubleshooting
Memory Utilization — System Memory

- NX-OS built-in memory monitoring
  - From 4.2(4), the default memory alert thresholds are 85% Minor, 90% Severe, 95% Critical
  - System memory issues affect all vDCs

```
N7K-1# show logging logfile | grep -b 5 -i memory | grep "Mar 22"
2011 Mar 22 15:40:13 N7K-1 %BGP-5-MEMALERT: bgp-1 [3439] BGP memory status changed from OK to Minor Alert
2011 Mar 22 15:40:13 N7K-1 %PLATFORM-2-MEMORY_ALERT: Memory Status Alert : MINOR. Usage 85% of Available Memory
```

```
N7K-1# show system internal memory-status
MemStatus: Minor Alert
```

```
N7K-1# show system internal memory-alerts-log
MINOR ALERT INFO
Tue Mar 22 15:40:13 PDT 2011
***** /proc/memory_events *****
Alert MINOR Reached at 1300833613.000287556
***** /proc/meminfo *****

MemTotal: 4115232 kB
MemFree: 318452 kB
Buffers: 81524 kB
Cached: 1726848 kB
```

```
Neighbor V AS MsgRcvd MsgSent TblVer InQ OutQ Up/Down State/PfxRcd
9.9.9.9 4 64 0 0 0 0 0 00:53:51 Idle (NoMem)
200.18.0.2 4 1203 8 7 32 0 0 00:01:34
```

```
N7k-3(config)# system memory-thresholds minor 85 severe 90 critical 95
```

While alert is present, BGP may be idle with ‘NoMem’ error. Use bgp process level disable-memory-alert-check cli or per neighbor low-memory exempt cli System memory upgrade to 8G may be required.

System memory alert threshold can be modified as required.
## Troubleshooting
### Memory Utilization — System Memory General Health Check

```bash
N7k-3-VDC3# show system internal kernel meminfo

<table>
<thead>
<tr>
<th>Memory Type</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MemTotal</td>
<td>4115232 kB</td>
</tr>
<tr>
<td>MemFree</td>
<td>263684 kB</td>
</tr>
<tr>
<td>Buffers</td>
<td>82400 kB</td>
</tr>
<tr>
<td>Cached</td>
<td>1817788 kB</td>
</tr>
<tr>
<td>ShmFS</td>
<td>1533324 kB</td>
</tr>
<tr>
<td>Allowed</td>
<td>1028808 Pages</td>
</tr>
<tr>
<td>Free</td>
<td>65921 Pages</td>
</tr>
<tr>
<td>Available</td>
<td>164026 Pages</td>
</tr>
<tr>
<td>SwapCached</td>
<td>0 kB</td>
</tr>
<tr>
<td>Active</td>
<td>2080320 kB</td>
</tr>
<tr>
<td>Inactive</td>
<td>1433752 kB</td>
</tr>
<tr>
<td>HighTotal</td>
<td>3338960 kB</td>
</tr>
<tr>
<td>HighFree</td>
<td>4092 kB</td>
</tr>
<tr>
<td>LowTotal</td>
<td>776272 kB</td>
</tr>
<tr>
<td>LowFree</td>
<td>259592 kB</td>
</tr>
<tr>
<td>SwapTotal</td>
<td>0 kB</td>
</tr>
<tr>
<td>SwapFree</td>
<td>0 kB</td>
</tr>
<tr>
<td>Dirty</td>
<td>0 kB</td>
</tr>
<tr>
<td>Writeback</td>
<td>0 kB</td>
</tr>
<tr>
<td>AnonPages</td>
<td>1613748 kB</td>
</tr>
<tr>
<td>Mapped</td>
<td>456088 kB</td>
</tr>
<tr>
<td>Slab</td>
<td>142884 kB</td>
</tr>
<tr>
<td>SReclaimable</td>
<td>25556 kB</td>
</tr>
<tr>
<td>SUnreclaim</td>
<td>117328 kB</td>
</tr>
<tr>
<td>PageTables</td>
<td>32756 kB</td>
</tr>
<tr>
<td>NFS_Unstable</td>
<td>0 kB</td>
</tr>
<tr>
<td>Bounce</td>
<td>0 kB</td>
</tr>
<tr>
<td>WritebackTmp</td>
<td>0 kB</td>
</tr>
<tr>
<td>CommitLimit</td>
<td>2057616 kB</td>
</tr>
<tr>
<td>Committed_AS</td>
<td>4837280 kB</td>
</tr>
<tr>
<td>VmallocTotal</td>
<td>188408 kB</td>
</tr>
<tr>
<td>VmallocUsed</td>
<td>161092 kB</td>
</tr>
<tr>
<td>VmallocChunk</td>
<td>27272 kB</td>
</tr>
<tr>
<td>HugePages_Total</td>
<td>0</td>
</tr>
<tr>
<td>HugePages_Free</td>
<td>0</td>
</tr>
<tr>
<td>HugePages_Rsvd</td>
<td>0</td>
</tr>
<tr>
<td>HugePages_Surp</td>
<td>0</td>
</tr>
<tr>
<td>Hugepagesize</td>
<td>2048 kB</td>
</tr>
<tr>
<td>DirectMap4k</td>
<td>2048 kB</td>
</tr>
<tr>
<td>DirectMap2M</td>
<td>841728 kB</td>
</tr>
</tbody>
</table>
```

- **MemTotal**: Total amount of memory in the system (4GB in N7K Sup1)
- **Cached**: Memory used by page cache (tmp fs mounts and data cached from bootflash)
- **Available**: Amount of free memory in pages (takes into account the space that could be made available in page cache and free lists)
- **Mapped**: Memory mapped into page tables (data being used by non-kernel processes)
- **Slab**: Rough indication of kernel memory consumption

**NOTE**: 1 Page = 4kB
Troubleshooting
Memory Utilization — System Memory General Health Check

N7K-1-VDC2# show system resources
Load average:  1 minute: 0.11   5 minutes: 0.09   15 minutes: 0.14
Processes : 1241 total, 2 running
CPU states : 2.0% user, 3.4% kernel, 94.6% idle
Memory usage: 4115232K total, 3606556K used, 508676K free

N7K-1-VDC2# show processes memory | egrep "PID|--|ospf|bgp"

<table>
<thead>
<tr>
<th>PID</th>
<th>MemAlloc</th>
<th>MemLimit</th>
<th>MemUsed</th>
<th>StackBase/Ptr</th>
<th>Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>3981</td>
<td>43761664</td>
<td>446487641</td>
<td>247361536</td>
<td>bff070c0/bff06b80</td>
<td>ospf</td>
</tr>
<tr>
<td>3982</td>
<td>9428992</td>
<td>446266867</td>
<td>230895616</td>
<td>bff070c0/bff06b80</td>
<td>ospfv3</td>
</tr>
<tr>
<td>3986</td>
<td>18247680</td>
<td>2411763200</td>
<td>271065088</td>
<td>bfe7a850/bfe7a760</td>
<td>bgp</td>
</tr>
</tbody>
</table>

N7K-1-VDC2# show system internal processes memory | egrep "PID|ospf|bgp"

<table>
<thead>
<tr>
<th>PID</th>
<th>TTY</th>
<th>STAT</th>
<th>TIME</th>
<th>MAJFLT</th>
<th>TRS</th>
<th>RSS</th>
<th>VSZ</th>
<th>%MEM</th>
<th>COMMAND</th>
</tr>
</thead>
<tbody>
<tr>
<td>3981</td>
<td>?</td>
<td>Ssl</td>
<td>11:52:06</td>
<td>0</td>
<td>690</td>
<td>64840</td>
<td>176028</td>
<td>1.5</td>
<td>/isan/bin/routing-sw/ospf -t 6467</td>
</tr>
<tr>
<td>4392</td>
<td>?</td>
<td>Ssl</td>
<td>02:15:41</td>
<td>0</td>
<td>690</td>
<td>63136</td>
<td>157424</td>
<td>1.5</td>
<td>/isan/bin/routing-sw/ospf -t 6467</td>
</tr>
<tr>
<td>4396</td>
<td>?</td>
<td>Ssl</td>
<td>00:35:01</td>
<td>0</td>
<td>1460</td>
<td>40856</td>
<td>180744</td>
<td>0.9</td>
<td>/isan/bin/routing-sw/bgp -t 1204</td>
</tr>
<tr>
<td>3986</td>
<td>?</td>
<td>Ssl</td>
<td>00:37:57</td>
<td>0</td>
<td>1460</td>
<td>39944</td>
<td>199176</td>
<td>0.9</td>
<td>/isan/bin/routing-sw/bgp -t 1203</td>
</tr>
<tr>
<td>3982</td>
<td>?</td>
<td>Ssl</td>
<td>01:16:17</td>
<td>0</td>
<td>728</td>
<td>22448</td>
<td>159948</td>
<td>0.5</td>
<td>/isan/bin/routing-sw/ospfv3 -t 8893</td>
</tr>
<tr>
<td>4393</td>
<td>?</td>
<td>Ssl</td>
<td>01:14:42</td>
<td>0</td>
<td>728</td>
<td>21436</td>
<td>141808</td>
<td>0.5</td>
<td>/isan/bin/routing-sw/ospfv3 -t 8893</td>
</tr>
<tr>
<td>3431</td>
<td>?</td>
<td>Ssl</td>
<td>01:09:00</td>
<td>0</td>
<td>728</td>
<td>15356</td>
<td>173136</td>
<td>0.3</td>
<td>/isan/bin/routing-sw/ospfv3 -t 1</td>
</tr>
<tr>
<td>3430</td>
<td>?</td>
<td>Ssl</td>
<td>01:08:23</td>
<td>0</td>
<td>690</td>
<td>15144</td>
<td>142376</td>
<td>0.3</td>
<td>/isan/bin/routing-sw/ospf -t 1</td>
</tr>
<tr>
<td>4811</td>
<td>?</td>
<td>Ssl</td>
<td>01:08:52</td>
<td>0</td>
<td>690</td>
<td>14832</td>
<td>123944</td>
<td>0.3</td>
<td>/isan/bin/routing-sw/ospf -t 1</td>
</tr>
<tr>
<td>3436</td>
<td>?</td>
<td>Ssl</td>
<td>01:07:37</td>
<td>0</td>
<td>690</td>
<td>14416</td>
<td>141872</td>
<td>0.3</td>
<td>/isan/bin/routing-sw/ospf -t 6467</td>
</tr>
</tbody>
</table>

Output taken from any vDC shows processes from all configured vDCs
Troubleshooting
Memory Utilization — System Memory Per Process Utilization

N7K-1-VDC2# show system internal sysmogr service pid 3986
Service "bgp" ("bgp", 80):
  UUID = 0x11B, PID = 3986, SAP = 2351
  Restart count: 1
  The service never crashed since the last reboot.
  Tag = 1203
  Plugin ID: 1

N7K-1-VDC2# show system internal kernel memory uuid 0x11B
MEMORY TYPE    TOTAL  RSS  PSS  SHARED  PRIVATE
bgp            TEXT   1464 1224 1224 1204    20
               DATA    24   16   16     0    16
Anonymous      HEAP  8328 8308 8308     0  8308
ld-2.8.so      TEXT   104  100  100  100     0
               RO_DATA  4   4   4     0    4
               DATA    4   4   4     0    4
libc-2.8.so    TEXT  1252  440  440  440     0
               RO_DATA  8   8   8     0    8
              DATA    4   4   4     0    4
Anonymous      MALLOC/MMAP  9488 8368 8368     0  8368
libdl-2.8.so   TEXT   8    8   8     8    0
               RO_DATA  4   4   4     0    4
              DATA    4   4   4     0    4
libpthread-2.8.so TEXT   80   68   68   68     0
               RO_DATA  4   4   4     0    4
[snip]
Troubleshooting
Memory Utilization — System Memory Per Process Utilization

N7K-1# show system internal pktmgr internal mem-stats detail | grep -b 13 -a 3
TCP_MEM_client_t

| TYPE  | NAME                      | ALLOCS       |            |            |            |            |            |
|-------|---------------------------|--------------|------------|------------|------------|------------|
| 2     | TCP_MEM_inpcb             | 18           | 66         | 3240       | 11880      |
| 3     | TCP_MEM_socket            | 18           | 66         | 11160      | 40920      |
| 4     | TCP_MEM_getsockaddr       | 0            | 1          | 0          | 40         |
| 5     | TCP_MEM_tcp_msg_t         | 17           | 17         | 14892      | 14892      |
| 6     | TCP_MEM_tseg_qent         | 0            | 1          | 0          | 28         |
| 7     | TCP_MEM_tcpcb             | 3            | 51         | 732        | 12444      |
| 9     | TCP_MEM_sockaddr_in_dcos  | 0            | 1          | 0          | 24         |
| 10    | TCP_MEM_syncache          | 0            | 33         | 0          | 4620       |
| 11    | TCP_MEM_syncache_head     | 1            | 1          | 12296      | 12296      |
| 12    | TCP_MEM_client_t          | 4153         | 4154       | 71099360   | 71116480   |

Total bytes: 71141680 (69474k)

Symptoms indicate memory leak in TCP_MEM_client and match CSCto12912
## Troubleshooting

Memory Utilization — Shared Memory, Estimated Utilization

N7k-3-VDC3# `show routing ip multicast memory estimate groups 200 sources-per-group 16 oifs-per-entry 16`

<table>
<thead>
<tr>
<th>Shared memory estimates:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Current max</td>
<td>8 MB; 204 groups</td>
</tr>
<tr>
<td></td>
<td>16 sources-per-group</td>
</tr>
<tr>
<td></td>
<td>16 oifs-per-entry</td>
</tr>
<tr>
<td>In-use</td>
<td>4 MB; 1 groups</td>
</tr>
<tr>
<td></td>
<td>1 sources-per-group (average)</td>
</tr>
<tr>
<td></td>
<td>0 oifs-per-entry (average)</td>
</tr>
<tr>
<td>Configured max</td>
<td>8 MB; 204 groups</td>
</tr>
<tr>
<td></td>
<td>16 sources-per-group</td>
</tr>
<tr>
<td></td>
<td>16 oifs-per-entry</td>
</tr>
<tr>
<td>Estimate</td>
<td>8 MB; 200 groups</td>
</tr>
<tr>
<td></td>
<td>16 sources-per-group</td>
</tr>
<tr>
<td></td>
<td>16 oifs-per-entry</td>
</tr>
</tbody>
</table>

N7k-3-VDC3# `show routing ip unicast memory estimate routes 180000 next-hops 4`

<table>
<thead>
<tr>
<th>Shared memory estimates:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Current max</td>
<td>8 MB; 6868 routes with 16 nhs</td>
</tr>
<tr>
<td></td>
<td>in-use 1 MB; 143 routes with 2 nhs (average)</td>
</tr>
<tr>
<td>Configured max</td>
<td>8 MB; 6868 routes with 16 nhs</td>
</tr>
<tr>
<td>Estimate</td>
<td>69 MB; 180000 routes with 4 nhs</td>
</tr>
</tbody>
</table>

Useful cli to predict mrib shared memory utilization based on number of multicast groups, sources and output interfaces (oifs)

Useful cli to predict urib shared memory utilization based on number of unicast prefixes and next-hops
## Troubleshooting
### Memory Utilization — Shared Memory Allocation Failure

2010 Jun 12 15:05:13 N7K-1-VDC2%MRIB-3-MALLOC_FAILED: mrib [6971] sm_malloc() failed for mrib_notify_buffer

2010 Jun 12 15:05:23 N7K-1-VDC2 %MRIB-4-SYSLOG_SL_MSG_WARNING:MRIB-3-MALLOC_FAILED: message repeated 3835 times in last 60 sec

```
N7K-1-VDC2# show resource

<table>
<thead>
<tr>
<th>Resource</th>
<th>Min</th>
<th>Max</th>
<th>Used</th>
<th>Unused</th>
<th>Avail</th>
</tr>
</thead>
<tbody>
<tr>
<td>vlan</td>
<td>16</td>
<td>4094</td>
<td>603</td>
<td>0</td>
<td>3491</td>
</tr>
<tr>
<td>monitor-session</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>monitor-session-erspan-dst</td>
<td>0</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>vrf</td>
<td>16</td>
<td>200</td>
<td>2</td>
<td>14</td>
<td>198</td>
</tr>
<tr>
<td>port-channel</td>
<td>0</td>
<td>768</td>
<td>2</td>
<td>0</td>
<td>759</td>
</tr>
<tr>
<td>u4route-mem</td>
<td>8</td>
<td>8</td>
<td>1</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>u6route-mem</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>m4route-mem</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>m6route-mem</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>
```

N7K-1(config-vdc)# limit-resource m4route-mem minimum 24 maximum 24

```
N7K-1-VDC2# show resource | egrep "Resource|---|m4route-mem"

<table>
<thead>
<tr>
<th>Resource</th>
<th>Min</th>
<th>Max</th>
<th>Used</th>
<th>Unused</th>
<th>Avail</th>
</tr>
</thead>
<tbody>
<tr>
<td>m4route-mem</td>
<td>24</td>
<td>24</td>
<td>4</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>
```

Message indicates that there was lack of shared memory for multicast rib and default setting adjustment was required.

Minimum and maximum shared memory allocation must be equal.

Switchover, vDC reload or system reload is required to get new shared memory allocation into effect.
Agenda

- Before You Get Started
  - Traditional Versus NX-OS Troubleshooting Approach
  - Nexus 7000 Module and Forwarding Engine Architecture Overview
  - Built-in Troubleshooting Tools

- Troubleshooting
  - CPU
  - Control-Plane
  - Memory Utilization
  - vPC
    - Unicast Layer 2 and Layer 3 Forwarding and ARP
    - Multicast Layer 2 and Layer 3 Forwarding
  - Switch Fabric
  - ACL
  - QoS
Troubleshooting vPC

- **vPC characteristic**
  - Dual control-plane
  - Eliminates STP blocking ports
  - FHRP active/active mode
  - **Loop-avoidance logic** (drop packet received on vPC peer link (PL) and direct to another vPC port-channel, vsl bit set on ingress and checked on egress)
  - **Cisco Fabric Services (CFS)** protocol is used to synchronize configuration and state machines between vpc peers (igmp, pim etc)

- **vPC does not support**
  - L3 adjacencies between vpc peers and 3rd device behind vpc port-channel connected to L2 switch
  - non-default pim/ospf/hsrp timers
  - PIM-DM, SSM. PIM bi-dir
  - pim spt-threshold infinity

In case your network has any of ‘not supported’, eliminate it before you spend any time troubleshooting your network issue.
Troubleshooting vPC

- **Generic vPC recommendations**
  - PL 10G ports (only) in dedicated mode
  - Dedicated L3 vPC peer keel-alive (PKL) link
  - peer-gateway to accommodate non RFC compliant hosts connected to L2 switch
  - peer-gateway exclude <vlan-list> in case vPC PL resides on F1 I/O module
  - peer-switch for faster stp convergence (both peers appear to be roots for rest of L2 topology)

- **Routing vPC recommendations**
  - Dedicated L3 link between vPC peers or
  - Dedicated L2 link between vPC peers with p2p svi interfaces or
  - Dedicated vlan carried on vPC PL and not extended to vPC connected L2 switch with p2p svi interfaces
  - ip pim pre-build-spt for faster multicast failover
**Troubleshooting**

vPC — General Health Check

**PeerB# show vpc brief**

Legend:

(*) - local vPC is down, forwarding via vPC peer-link

<table>
<thead>
<tr>
<th>vPC domain id</th>
<th>: 64</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peer status</td>
<td>peer adjacency formed ok</td>
</tr>
<tr>
<td>vPC keep-alive status</td>
<td>peer is alive</td>
</tr>
<tr>
<td>Configuration consistency status</td>
<td>success</td>
</tr>
<tr>
<td>Type-2 consistency status</td>
<td>failed</td>
</tr>
<tr>
<td>Type-2 consistency reason</td>
<td>SVI type-2 configuration incompatible</td>
</tr>
<tr>
<td>vPC role</td>
<td>primary</td>
</tr>
<tr>
<td>Number of vPCs configured</td>
<td>2</td>
</tr>
<tr>
<td>Peer Gateway</td>
<td>Disabled</td>
</tr>
<tr>
<td>Peer gateway excluded VLANs</td>
<td>-</td>
</tr>
<tr>
<td>Dual-active excluded VLANs</td>
<td>-</td>
</tr>
</tbody>
</table>

**vPC Peer-link status**

<table>
<thead>
<tr>
<th>id</th>
<th>Port</th>
<th>Status</th>
<th>Active vlans</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Po664</td>
<td>up</td>
<td>1,19,31-35,2000,4092-4093</td>
</tr>
</tbody>
</table>

**vPC status**

<table>
<thead>
<tr>
<th>id</th>
<th>Port</th>
<th>Status</th>
<th>Consistency</th>
<th>Reason</th>
<th>Active vlans</th>
</tr>
</thead>
<tbody>
<tr>
<td>667</td>
<td>Po667(SU)</td>
<td>Eth</td>
<td>LACP</td>
<td>Eth1/10(P)</td>
<td>1,19,31-35,2000,4092-4093</td>
</tr>
<tr>
<td>4093</td>
<td>Po4093</td>
<td>up</td>
<td>success</td>
<td>success</td>
<td>4093</td>
</tr>
</tbody>
</table>

Type-2 inconsistency indicates that one vPC peer has SVI configured and in up/up state and the other does not have it.

**Troubleshooting vPC — General Health Check**

**PeerB# show system internal ethpm info interface e1/1 |i rate delay(1), bw(10000000), rate-mode(dedicated)**

**PeerB# show port-channel summary interface port-channel 667 | grep 667**

**PeerB# show port-channel summary interface port-channel 667 | grep 667**
### Troubleshooting

**vPC — General Health Check**

PeerA# `show vpc consistency-parameters global`

Legend:

**Type 1:** vPC will be suspended in case of mismatch

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Local Value</th>
<th>Peer Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>STP Mode</td>
<td>1</td>
<td>Rapid-PVST</td>
<td>Rapid-PVST</td>
</tr>
<tr>
<td>STP Disabled</td>
<td>1</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>STP MST Region Name</td>
<td>1</td>
<td>&quot;&quot;</td>
<td>&quot;&quot;</td>
</tr>
<tr>
<td>STP MST Region Revision</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>STP MST Region Instance to VLAN Mapping</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STP Loopguard</td>
<td>1</td>
<td>Disabled</td>
<td>Disabled</td>
</tr>
<tr>
<td>STP Bridge Assurance</td>
<td>1</td>
<td>Enabled</td>
<td>Enabled</td>
</tr>
<tr>
<td>STP Port Type, Edge BPDUGuard</td>
<td>1</td>
<td>Normal, Disabled,</td>
<td>Normal, Disabled,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Disabled</td>
<td>Disabled</td>
</tr>
<tr>
<td>STP MST Simulate PVST</td>
<td>1</td>
<td>Enabled</td>
<td>Enabled</td>
</tr>
<tr>
<td>VTP domain</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VTP version</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>VTP mode</td>
<td>2</td>
<td>Server</td>
<td>Server</td>
</tr>
<tr>
<td>VTP password</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VTP pruning status</td>
<td>2</td>
<td>Disabled</td>
<td>Disabled</td>
</tr>
<tr>
<td>Interface-vlan admin up</td>
<td>2</td>
<td>19,31-35, 2000,4092-4093</td>
<td>19,31-35,4092-4093</td>
</tr>
<tr>
<td>Interface-vlan routing capability</td>
<td>2</td>
<td>1,19,31-35,2000,4092-4</td>
<td>1,19,31-35,4092-4</td>
</tr>
<tr>
<td>Allowed VLANs</td>
<td>-</td>
<td>1,19,31-35, 2000,4092-4</td>
<td>1,19,31-35, 2000,4092-4</td>
</tr>
<tr>
<td>Local suspended VLANs</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Note: Both vPC peers will be in active (primary) state if both PL and PKL fail and stay active if only PL is recovered. In case only PL fails, secondary vPC peer suspends all of its vPCs.

Note that interface vlan2000 is missing!
### Troubleshooting

**vPC — General Health Check**

**PeerA#** show vpc brief

Legend:

- (*) - local vPC is down, forwarding via vPC peer-link

<table>
<thead>
<tr>
<th>vPC domain id</th>
<th>: 64</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peer status</td>
<td>: peer adjacency formed ok</td>
</tr>
<tr>
<td>vPC keep-alive status</td>
<td>: peer is alive</td>
</tr>
<tr>
<td>Configuration consistency status</td>
<td>: failed</td>
</tr>
<tr>
<td>Configuration consistency reason</td>
<td>: vPC type-1 configuration incompatible - STP Mode inconsistent</td>
</tr>
<tr>
<td>Type-2 consistency status</td>
<td>: success</td>
</tr>
<tr>
<td>vPC role</td>
<td>: secondary</td>
</tr>
<tr>
<td>Number of vPCs configured</td>
<td>: 2</td>
</tr>
<tr>
<td>Peer Gateway</td>
<td>: Disabled</td>
</tr>
<tr>
<td>Peer gateway excluded VLANs</td>
<td>: -</td>
</tr>
<tr>
<td>Dual-active excluded VLANs</td>
<td>: -</td>
</tr>
</tbody>
</table>

**vPC Peer-link status**

<table>
<thead>
<tr>
<th>id</th>
<th>Port</th>
<th>Status</th>
<th>Active vlans</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Po664</td>
<td>up</td>
<td>-</td>
</tr>
</tbody>
</table>

**vPC status**

<table>
<thead>
<tr>
<th>id</th>
<th>Port</th>
<th>Status</th>
<th>Consistency</th>
<th>Reason</th>
<th>Active vlans</th>
</tr>
</thead>
<tbody>
<tr>
<td>667</td>
<td>Po667</td>
<td>up</td>
<td>failed</td>
<td>Global compat check failed</td>
<td>-</td>
</tr>
<tr>
<td>4093</td>
<td>Po4093</td>
<td>up</td>
<td>failed</td>
<td>Global compat check failed</td>
<td>-</td>
</tr>
</tbody>
</table>

*STP incompatibility was introduced and vpc was suspended*

*Vlan2000 SVI issue was fixed*

*PeerA# show system internal ethpm info interface e1/9 |i rate delay(1), bw(10000000), rate-mode(dedicated)*

*PeerA# show port-channel summary interface port-channel 667 | grep 667*

*PeerA# show port-channel summary interface port-channel 667 | grep 667*

*667 Po667(SU) Eth LACP Eth1/2(P)*
Troubleshooting vPC — Why Does Routing not Work Without Peer-Gateway? EC Hash Check

**R14# show ip ospf neighbor vlan 19**

<table>
<thead>
<tr>
<th>Neighbor ID</th>
<th>Pri</th>
<th>State</th>
<th>Dead Time</th>
<th>Address</th>
<th>Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>200.17.0.1</td>
<td>1</td>
<td>FULL/DR</td>
<td>00:00:03</td>
<td>192.251.19.11</td>
<td>Vlan19</td>
</tr>
<tr>
<td>200.18.0.1</td>
<td>1</td>
<td>FULL/DROTHER</td>
<td>00:00:03</td>
<td>192.251.19.22</td>
<td>Vlan19</td>
</tr>
</tbody>
</table>

**R14# show ip route 40.64.64.40 | i Vlan19**

Last update from 192.251.19.22 on Vlan19, 04:42:03 ago

* 192.251.19.22, from 200.18.0.1, 04:42:03 ago, via Vlan19
  192.251.19.11, from 200.17.0.1, 04:42:03 ago, via Vlan19

Peer-gateway is NOT configured, ospf is stable!


*via 40.9.3.1, Vlan4093, [1/0], 05:49:10, static

**R14# remote command switch test ether-channel load interface po141 ip 14.14.14.15 40.64.64.40**

Computed RBH: 0x5
Would select Te3/3 of Po141

PeerA# show interface port-channel1667 | i unicast | diff

< 3 unicast packets 32 multicast packets 0 broadcast packets
< 4 unicast packets 55 multicast packets 0 broadcast packets

---

> 1011 unicast packets 117 multicast packets 0 broadcast packets
> 13 unicast packets 201 multicast packets 0 broadcast packets

PeerB# show int po667 | i uni

1 unicast packets 9 multicast packets 0 broadcast packets
3 unicast packets 46 multicast packets 0 broadcast packets

Sent 1k pings from R14 to R15
Troubleshooting
vPC — Why Does Routing not Work Without Peer-Gateway? PL Traffic
Elam Capture

Use *show hardware internal dev-port-map* to find port to PB1/PB2 mapping
48-port: PB1->Metro0, PB2->Metro1, 32-port PB1->Metro2-3, PB2->Metro0-1

PeerB# attach mod 1
Attaching to module 1 ...
To exit type 'exit', to abort type '$.'
module-1# elam asic eureka instance 1
destination-ip 40.64.64.40 rbi-corelate
module-1(eureka-elam)# trigger rbus rbi pb2 ip if cap2 1
module-1(eureka-elam)# start
module-1(eureka-elam)# status
Instance: 1
EU-DBUS: Triggered
trigger dbus dbi ingress ipv4 if source-ipv4-address 14.14.14.14 destination-ipv4-address 40.64.64.40 rbi-corelate
EU-RBUS: Triggered
trigger rbus rbi pb1 ip if cap2 1

**DBUS** – data bus header prepended to packet by ingress port asic
**RBUS** – result bus header created by forwarding engine and executed by rewrite asic
**PB1/PB2** – forwarding engine packet buffer
**CAP2** – rbus field to synchronize dbus and rbus for ELAM (rbi-corelate)
**DBI** – dbus input interface
**RBI** – rbus input interface

R15 destination IP
R14 source IP
ELAM confirms PeerA is sending traffic destined to R15 across PL to PeerB
Troubleshooting
vPC — Why Does Routing not Work Without Peer-Gateway? PL Elam Capture Result Analysis

```
module-1(eureka-elam)# show dbus | egrep "seq|vlan|source|dest|l3_p|mac"
seq = 0x0e
vlan = 19
source_flood = 0x0
source_index = 0x00a2a
l3_packet_type = 0x0, (0:Ethernet, 1:IPX, 2-4: IEEE 802.3)
l3_protocol = 0x0 (0:IPv4, 6:IPv6)
l3_protocol_type = 0x01, (1:ICMP, 2:IGMP, 4:IP, 6:TCP, 17:UDP)
destination_flood = 0x0
destination_index = 0x00400
dmac = 00.22.19.19.19.19
smac = 00.13.5f.1f.46.c0
ip_source = 014.014.014.014
ip_destination = 040.064.064.040
```

```
module-1(eureka-elam)# show rbus | egrep "seq|ccc|cap2|flood|rbh|vlan|index|rit"
seq = 0x0e
ccc = 0x4
cap2 = 0x1
flood = 0x0
dest_index = 0x00a2e
vlan = 4093
rbh = 0x4
fabric_priority = 0x0
data(rit/dmac/recir) = 00.24.98.e9.11.43
data(rit/smac/recir) = 00.24.98.e9.11.42
```

- Sequence number
- Ingress vlan
- LTL index of ingress port
- Protocol

- Sequence number
- Rewrite instruction
- LTL index of egress port
- Egress vlan
- Egress Port-channel hash

Dest_index belongs to Po4093 which confirm PeerB is trying to send traffic out to R15 but vPC loop-avoidance logic is dropping it
vPC loop-avoidance logic asic drops on 48-port 1G M1 I/O Module

PeerB# show system internal pixm info ltl 0x00a2e

<table>
<thead>
<tr>
<th>PC_TYPE</th>
<th>PORT</th>
<th>LTL</th>
<th>RES_ID</th>
<th>LTL_FLAG</th>
<th>CB_FLAG</th>
<th>MEMB_CNT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>Po4093</td>
<td>0x0a2e</td>
<td>0x16000ffc</td>
<td>0x00000000</td>
<td>0x00000002</td>
<td>2</td>
</tr>
<tr>
<td>Member</td>
<td>rbh</td>
<td>rbh_cnt</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eth3/34</td>
<td>0x000000f0</td>
<td>0x04</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eth3/32</td>
<td>0x000000f0</td>
<td>0x04</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

module-3# show hardware internal dev-port-map | egrep "32|34|FP"

<table>
<thead>
<tr>
<th>FP port</th>
<th>PHYS</th>
<th>Secur</th>
<th>MAC_0</th>
<th>RWR_0</th>
<th>L2LKP</th>
<th>L3LKP</th>
<th>QUEUE</th>
<th>SWICHF</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>3</td>
<td>7</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>34</td>
<td>4</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

PeerB# show hardware internal statistics module 3 device mac errors port 32 | egrep -b 9 aric

<table>
<thead>
<tr>
<th>Device: R2D2</th>
<th>Role: MAC</th>
<th>Mod: 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Last cleared @ Mon Mar 28 21:46:42 2011</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Device Statistics Category :: ERROR</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<p>| Instance: 2 |</p>
<table>
<thead>
<tr>
<th>ID Name</th>
<th>Value</th>
<th>Ports</th>
</tr>
</thead>
<tbody>
<tr>
<td>-- -- -----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>4422 mstat_rdrop</td>
<td>0000000000001791</td>
<td>32 -</td>
</tr>
<tr>
<td>28688 aric_no_port_select_error</td>
<td>0000000000001037</td>
<td>25-36</td>
</tr>
</tbody>
</table>
Troubleshooting vPC — Why Does Routing not Work with Peer-Gateway? OSPF Check

PeerA# show ip ospf neighbor vlan19 | grep -a 2 Neighbor

<table>
<thead>
<tr>
<th>Neighbor ID</th>
<th>Pri State</th>
<th>Up Time</th>
<th>Address</th>
<th>Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>200.18.0.1</td>
<td>FULL/BDR</td>
<td>1d05h</td>
<td>192.251.19.22</td>
<td>Vlan19</td>
</tr>
</tbody>
</table>

PeerB# show ip ospf neighbor vlan 19 | grep -a 2 Neighbor

<table>
<thead>
<tr>
<th>Neighbor ID</th>
<th>Pri State</th>
<th>Up Time</th>
<th>Address</th>
<th>Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>200.17.0.1</td>
<td>FULL/DR</td>
<td>1d05h</td>
<td>192.251.19.11</td>
<td>Vlan19</td>
</tr>
</tbody>
</table>

peer-gateway IS configured, ospf is unable to come up

OSPF multicast packet are ok but unicast communication does not work due to ttl=1 and G-bit set which forces routing

R14#show ip ospf neighbor vlan 19

<table>
<thead>
<tr>
<th>Neighbor ID</th>
<th>Pri State</th>
<th>Dead Time</th>
<th>Address</th>
<th>Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>200.17.0.1</td>
<td>EXCHANGE/DR</td>
<td>00:00:03</td>
<td>192.251.19.11</td>
<td>Vlan19</td>
</tr>
<tr>
<td>200.18.0.1</td>
<td>EXCHANGE/BDR</td>
<td>00:00:03</td>
<td>192.251.19.22</td>
<td>Vlan19</td>
</tr>
</tbody>
</table>

R14#
Mar 29 16:53:55.691: %OSPF-5-ADJCHG: Process 6467, Nbr 200.18.0.1 on Vlan19 from EXCHANGE to DOWN, Neighbor Down: Too many retransmissions
Mar 29 16:54:01.111: %OSPF-5-ADJCHG: Process 6467, Nbr 200.17.0.1 on Vlan19 from EXCHANGE to DOWN, Neighbor Down: Too many retransmissions
Mar 29 16:54:55.692: %OSPF-5-ADJCHG: Process 6467, Nbr 200.18.0.1 on Vlan19 from DOWN to DOWN, Neighbor Down: Ignore timer expired
Mar 29 16:55:01.112: %OSPF-5-ADJCHG: Process 6467, Nbr 200.17.0.1 on Vlan19 from DOWN to DOWN, Neighbor Down: Ignore timer expired
vPC — Why Does Routing not Work with Peer-Gateway? Wireshark Capture

PeerA# ethanalyzer local interface inband decode-internal capture-filter "proto 89 and host 192.251.19.14 and host 192.251.19.22" limit-captured-frames 1 detail > bootflash:ospf_neighbor.txt

Capturing on inband
1 packet captured

PeerA# ethanalyzer local interface inband capture-filter "proto 89 and host 192.251.19.14 and host 192.251.19.22" limit-captured-frames 1 write bootflash:ospf_wrong_neighbor.pcap

Capturing on inband
1 packet captured
Program exited with status 0.

PeerB# ethanalyzer local interface inband decode-internal capture-filter "proto 89 and host 192.251.19.14 and host 192.251.19.11" limit-captured-frames 1 detail > bootflash:ospf_neighbor.txt

Capturing on inband
1 packet captured

PeerB# ethanalyzer local interface inband capture-filter "proto 89 and host 192.251.19.14 and host 192.251.19.11" limit-captured-frames 1 write bootflash:ospf_wrong_neighbor.pcap

Capturing on inband
1 packet captured
Program exited with status 0.
Troubleshooting
vPC — Why Does Routing not Work?

PeerA# show file bootflash:ospf_neighbor.txt | egrep -i "sour|dest|add|Time to live:" | exclu

Global

NXOS SOURCE INDEX: 2626
NXOS DEST INDEX: 1024
   .... 0 ........ .... = IG bit: Individual address (unicast)
Source: 00:13:5f:1f:46:c0 (00:13:5f:1f:46:c0)
   Address: 00:13:5f:1f:46:c0 (00:13:5f:1f:46:c0)
   .... 0 ........ .... = IG bit: Individual address (unicast)
Time to live: 1
Destination: 192.251.19.22 (192.251.19.22)

PeerB show file bootflash:ospf_neighbor.txt | egrep -i "sour|dest|add|Time to live:" | exclu

Global

NXOS SOURCE INDEX: 2604
NXOS DEST INDEX: 1024
Destination: 00:11:19:19:19:19 (00:11:19:19:19:19)
   Address: 00:11:19:19:19:19 (00:11:19:19:19:19)
   .... 0 ........ .... = IG bit: Individual address (unicast)
Source: 00:13:5f:1f:46:c0 (00:13:5f:1f:46:c0)
   Address: 00:13:5f:1f:46:c0 (00:13:5f:1f:46:c0)
   .... 0 ........ .... = IG bit: Individual address (unicast)
Time to live: 1
Destination: 192.251.19.11 (192.251.19.11)
In case the issue you have encountered is complicated and you can’t figure it out, collect **show tech-support** output asap!
Agenda

- Before You Get Started
  - Traditional Versus NX-OS Troubleshooting Approach
  - Nexus 7000 Module and Forwarding Engine Architecture Overview
  - Built-in Troubleshooting Tools

- Troubleshooting
  - CPU
  - Control-Plane
  - Memory Utilization
  - vPC
  - Unicast Layer 2 and Layer 3 Forwarding and ARP
  - Multicast Layer 2 and Layer 3 Forwarding
  - Switch Fabric
  - ACL
  - QoS
### Unicast L2 and L3 Forwarding, ARP

#### L2 — Mac Addresses, Software Entry

**Legend:**
- * - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
- age - seconds since last seen, + - primary entry using vPC Peer-Link

<table>
<thead>
<tr>
<th>VLAN</th>
<th>MAC Address</th>
<th>Type</th>
<th>age</th>
<th>Secure</th>
<th>NTFY</th>
<th>Ports/SWID.SSID.LID</th>
</tr>
</thead>
<tbody>
<tr>
<td>G 32</td>
<td>0000.0c07.ac20</td>
<td>static</td>
<td>-</td>
<td>F</td>
<td>F</td>
<td>sup-eth1(R)</td>
</tr>
<tr>
<td>G 32</td>
<td>0011.3232.3232</td>
<td>static</td>
<td>-</td>
<td>F</td>
<td>F</td>
<td>sup-eth1(R)</td>
</tr>
<tr>
<td>* 32</td>
<td>0022.3232.3232</td>
<td>static</td>
<td>-</td>
<td>F</td>
<td>F</td>
<td>vPC Peer-Link</td>
</tr>
<tr>
<td>* 32</td>
<td>0000.98b9.4868</td>
<td>dynamic</td>
<td>60</td>
<td>F</td>
<td>F</td>
<td>Po667</td>
</tr>
<tr>
<td>* 32</td>
<td>0013.5f1f.46c0</td>
<td>dynamic</td>
<td>120</td>
<td>F</td>
<td>F</td>
<td>Po667</td>
</tr>
</tbody>
</table>

**vPC topology from previous slides is used**

**peer-gateway is NOT configured and therefore only Vlan32 SVI mac and HSRP mac are flagged by G-bit**

---

**N7K-3-PeerB# show mac address-table vlan 32**

<table>
<thead>
<tr>
<th>VLAN</th>
<th>MAC Address</th>
<th>Type</th>
<th>age</th>
<th>Secure</th>
<th>NTFY</th>
<th>Ports/SWID.SSID.LID</th>
</tr>
</thead>
<tbody>
<tr>
<td>G 32</td>
<td>0000.0c07.ac20</td>
<td>static</td>
<td>-</td>
<td>F</td>
<td>F</td>
<td>vPC Peer-Link(R)</td>
</tr>
<tr>
<td>* 32</td>
<td>0011.3232.3232</td>
<td>static</td>
<td>-</td>
<td>F</td>
<td>F</td>
<td>vPC Peer-Link</td>
</tr>
<tr>
<td>G 32</td>
<td>0022.3232.3232</td>
<td>static</td>
<td>-</td>
<td>F</td>
<td>F</td>
<td>sup-eth1(R)</td>
</tr>
<tr>
<td>* 32</td>
<td>0000.98b9.4868</td>
<td>dynamic</td>
<td>0</td>
<td>F</td>
<td>F</td>
<td>Po667</td>
</tr>
<tr>
<td>* 32</td>
<td>0013.5f1f.46c0</td>
<td>dynamic</td>
<td>60</td>
<td>F</td>
<td>F</td>
<td>Po667</td>
</tr>
</tbody>
</table>

---

**N7K-1-PeerA# show mac address-table vlan 32 | egrep "G|Vlan|--"**

<table>
<thead>
<tr>
<th>VLAN</th>
<th>MAC Address</th>
<th>Type</th>
<th>age</th>
<th>Secure</th>
<th>NTFY</th>
<th>Ports/SWID.SSID.LID</th>
</tr>
</thead>
<tbody>
<tr>
<td>G 32</td>
<td>0000.0c07.ac20</td>
<td>static</td>
<td>-</td>
<td>F</td>
<td>F</td>
<td>sup-eth1(R)</td>
</tr>
<tr>
<td>G 32</td>
<td>0011.3232.3232</td>
<td>static</td>
<td>-</td>
<td>F</td>
<td>F</td>
<td>sup-eth1(R)</td>
</tr>
<tr>
<td>G 32</td>
<td>0022.3232.3232</td>
<td>static</td>
<td>-</td>
<td>F</td>
<td>F</td>
<td>vPC Peer-Link(R)</td>
</tr>
<tr>
<td>* 32</td>
<td>0000.98b9.4868</td>
<td>dynamic</td>
<td>420</td>
<td>F</td>
<td>F</td>
<td>Po667</td>
</tr>
<tr>
<td>* 32</td>
<td>0013.5f1f.46c0</td>
<td>dynamic</td>
<td>480</td>
<td>F</td>
<td>F</td>
<td>Po667</td>
</tr>
</tbody>
</table>

**peer-gateway IS configured and therefore PeerB Vlan32 svi mac is also flagged by G-bit**
### Unicast L2 and L3 Forwarding, ARP

**L2 — Mac Addresses, Hardware Entry**

<table>
<thead>
<tr>
<th>FE</th>
<th>Valid</th>
<th>PI</th>
<th>BD</th>
<th>MAC</th>
<th>Index</th>
<th>Stat</th>
<th>SW</th>
<th>Mod</th>
<th>Age</th>
<th>Tmr</th>
<th>GM</th>
<th>Sec</th>
<th>TR</th>
<th>NT</th>
<th>RM</th>
<th>RMA</th>
<th>Cap</th>
<th>Fld</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>958</td>
<td>0000.98aa.8ac9</td>
<td>0x00a42</td>
<td>0</td>
<td>0x003</td>
<td>215</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>958</td>
<td>0022.3232.3232</td>
<td>0x00a40</td>
<td>1</td>
<td>0x000</td>
<td>42</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>958</td>
<td>0000.0c07.ac20</td>
<td>0x00400</td>
<td>1</td>
<td>0x000</td>
<td>56</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>958</td>
<td>0100.0ff.fffe</td>
<td>0x00400</td>
<td>1</td>
<td>0x000</td>
<td>169</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>958</td>
<td>0011.3232.3232</td>
<td>0x00400</td>
<td>1</td>
<td>0x000</td>
<td>41</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:** equivalent output can be obtained from `rp` using ‘show hardware mac address-table 1 vlan 32’ cli (1 = module#)

**N7K-1-PeerA#** show system internal pixm info vlan-bd-db | b "VDC: 2"

**BD info for VDC: 2**

<table>
<thead>
<tr>
<th>VLAN</th>
<th>BD</th>
<th>BD LTL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>17</td>
<td>0x8011</td>
</tr>
</tbody>
</table>

[snip]

| 32   | 958 | 0x83be       |
| 33   | 959 | 0x83bf       |

**module-1#** show hardware internal statistics device l2lu errors

```
<table>
<thead>
<tr>
<th>Device:Eureka</th>
<th>Role:L2</th>
<th>Mod: 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Last cleared @ Fri Feb 25 21:30:09 2011</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Device Statistics Category :: ERROR |

<table>
<thead>
<tr>
<th>Instance: 0</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>Value</th>
<th>Ports</th>
</tr>
</thead>
<tbody>
<tr>
<td>185</td>
<td>Non-flood packets sent with drop-index</td>
<td>00000000000000039</td>
<td>1-32 I1</td>
</tr>
</tbody>
</table>
### Unicast L2 and L3 Forwarding, ARP

**L2 — Mac Asics Statistics**

```
N7K-1-PeerA# slot 1 show hardware internal statistics device mac pktflow port 2 | grep -v `^$

<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Device: R2D2                                           Role: MAC       Mod: 1</td>
</tr>
<tr>
<td>Last cleared @ Wed Apr 13 08:32:20 2011</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Instance: 0</th>
<th>ID</th>
<th>Name</th>
<th>Value</th>
<th>Ports</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4096</td>
<td>mstat_rx_pkts</td>
<td>00000000000656926</td>
<td>2,4,6,8 -</td>
</tr>
<tr>
<td></td>
<td>4128</td>
<td>mstat_rx_pkts_65_127</td>
<td>0000000000436692</td>
<td>2,4,6,8 -</td>
</tr>
</tbody>
</table>

[snip]

<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Device: Ashburton                                        Role: MAC       Mod: 1</td>
</tr>
<tr>
<td>Last cleared @ Wed Apr 13 08:32:20 2011</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Instance: 0</th>
<th>ID</th>
<th>Name</th>
<th>Value</th>
<th>Ports</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>ashburton_ingress_port0_total_pkt_count</td>
<td>0000000003708053</td>
<td>2 -</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>ashburton_ingress_port0_dollq_pkt_count</td>
<td>0000000000594413</td>
<td>2 -</td>
</tr>
</tbody>
</table>

[snip]

<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Device: Naxos                                            Role: MAC SECURITY  Mod: 1</td>
</tr>
<tr>
<td>Last cleared @ Wed Apr 13 08:32:20 2011</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Instance: 0</th>
<th>ID</th>
<th>Name</th>
<th>Value</th>
<th>Ports</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>11</td>
<td>sys_egress_octets</td>
<td>0000054061058144</td>
<td>2 -</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>sys_egress_unicast_frames</td>
<td>0000000000574369</td>
<td>2 -</td>
</tr>
</tbody>
</table>

[snip]

<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>---------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>---------------------------------------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>

**mstat** — Mac level counters  
**sys** — Fabric-side counters  
**phy** — Network-side counters  
**ingress** — Ingress from n7k perspective  
**egress** — Egress from n7k perspective
Unicast L2 and L3 Forwarding, ARP
L2 — Additional Useful CLI

N7K-1-PeerA# show system internal pixm info interface port-channel 664 vdc 2

<table>
<thead>
<tr>
<th>PC_TYPE</th>
<th>PORT</th>
<th>LTL</th>
<th>RES_ID</th>
<th>LTL_FLAG</th>
<th>CB_FLAG</th>
<th>MEMB_CNT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>Po664</td>
<td>0x0a40</td>
<td>0x16000297</td>
<td>0x00000000</td>
<td>0x00000002</td>
<td>1</td>
</tr>
</tbody>
</table>

Member rbh rbh_cnt
Eth1/9 0x000000ff 0x08

VLAN| BD| CBL |BD-St & CBL Direction:
--------------------------
|    |   |    |

[snip ...]
32 | 0x3be | FORWARDING | INCLUDE_IF_IN_BD | BOTH
33 | 0x3bf | FORWARDING | INCLUDE_IF_IN_BD | BOTH
[snip]

N7K-1-PeerA# show system internal l2fm info summary
Distribution: TRUE
Local Macdb Recovery: Done Remote Macdb Recovery: Done
[snip]
Global BPDU Recpt Enable: TRUE
Global L3 SVI Enable: TRUE
Get Alw Lrn Peer After Issu: FALSE
[snip]
Default Aging Time: 1800 seconds
peer_gwmac_special_sup_di: FALSE
peer_gwmac_peer_link: TRUE
Global Flush Underway: FALSE
Number of VLANs: 606, Reason: TMR_DEFAULT
Gateway Mac: 0023.ac64.46c2

N7K-1-PeerA# show port-channel rbh-distribution int po667
ChanId Member port RBH values Num of buckets
-------- ------------- --------------------------
- 667 Eth1/2 0,1,2,3,4,5,6,7 8

N7K-1-PeerA# show system internal l2fm info move_db
Vlan From Intf To Intf FE Bitmap
----------------------------------------------
No entries in move db

N7K-1-PeerA(config)# logging level l2fm 7
### Unicast L2 and L3 Forwarding, ARP

**L2 — Spanning-Tree**

**N7K-1-PeerA#**

```
show spanning-tree vlan 32 | grep -v "^$"
```

**VLAN0032**

Spanning tree enabled protocol rstp

<table>
<thead>
<tr>
<th>Root ID</th>
<th>Priority</th>
<th>Address</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>32</td>
<td>0023.04ee.be40</td>
</tr>
</tbody>
</table>

This bridge is the root

<table>
<thead>
<tr>
<th>Hello Time</th>
<th>Max Age</th>
<th>Forward Delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 sec</td>
<td>20 sec</td>
<td>15 sec</td>
</tr>
</tbody>
</table>

**Bridge ID**

<table>
<thead>
<tr>
<th>Priority</th>
<th>Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>0023.04ee.be40</td>
</tr>
</tbody>
</table>

Hello Time 2 sec Max Age 20 sec Forward Delay 15 sec

<table>
<thead>
<tr>
<th>Interface</th>
<th>Role</th>
<th>Sts</th>
<th>Cost</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Po664</td>
<td>Desg</td>
<td>FWD</td>
<td>128.4759</td>
<td>(vPC peer-link) Network P2p</td>
</tr>
<tr>
<td>Po667</td>
<td>Desg</td>
<td>FWD</td>
<td>128.4762</td>
<td>(vPC) P2p</td>
</tr>
</tbody>
</table>

**peer-switch** is configured on both PeerA and PeerB so they both appear as roots for rest of the L2 topology

- **N7K-1-PeerA#**

```
show spanning-tree internal event-history tree 32 interface port-channel 664 | grep -v "^$"
```

VDC02 VLAN0032 <port-channel1664>

0 Transition at 145271 usecs after Sat Apr 2 16:04:38 2011 State: BLK Role: Desg Age: 0 Inc: no [STP_PORT_STATE_CHANGE]

**N7K-3-PeerB#**

```
show spanning-tree vlan 32 | grep -v "^$"
```

**VLAN0032**

Spanning tree enabled protocol rstp

<table>
<thead>
<tr>
<th>Root ID</th>
<th>Priority</th>
<th>Address</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>32</td>
<td>0023.04ee.be40</td>
</tr>
</tbody>
</table>

This bridge is the root

<table>
<thead>
<tr>
<th>Hello Time</th>
<th>Max Age</th>
<th>Forward Delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 sec</td>
<td>20 sec</td>
<td>15 sec</td>
</tr>
</tbody>
</table>

**Bridge ID**

<table>
<thead>
<tr>
<th>Priority</th>
<th>Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>0023.04ee.be40</td>
</tr>
</tbody>
</table>

Hello Time 2 sec Max Age 20 sec Forward Delay 15 sec

<table>
<thead>
<tr>
<th>Interface</th>
<th>Role</th>
<th>Sts</th>
<th>Cost</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Po664</td>
<td>Desg</td>
<td>FWD</td>
<td>128.4759</td>
<td>(vPC peer-link) Network P2p</td>
</tr>
<tr>
<td>Po667</td>
<td>Desg</td>
<td>FWD</td>
<td>128.4762</td>
<td>(vPC) P2p</td>
</tr>
</tbody>
</table>
Unic和平L2 and L3 Forwarding, ARP
L3 — Software Entry, ECMP

N7K-1-VDC3# show ip route 172.31.31.0 | grep -b 2 via
172.31.31.0/24, ubest/mbest: 2/0
  *via 12.1.1.1, Po1, [110/80], 00:12:09, ospf-6467, intra
  *via 12.111.111.1, Po111, [110/80], 00:12:09, ospf-6467, intra

N7K-1-VDC3# show ip arp | grep "12.1.1.1|12.111.111.1"
12.1.1.1    00:01:36    0023.ac64.46c2    port-channel1
12.111.111.1 00:14:58  0023.ac64.46c2  port-channel111

N7K-1-VDC3# show ip adjacency 12.1.1.1 | grep -b 3 12.1
IP Adjacency Table for VRF default
Total number of entries: 1
Address     MAC Address     Pref Source   Interface
12.1.1.1    0023.ac64.46c2  50 arp      port-channel1

N7K-1-VDC3# show ip adjacency 12.111.111.1
Flags: # - Adjacencies Throttled for Glean
IP Adjacency Table for VRF default
Total number of entries: 1
Address     MAC Address     Pref Source   Interface
12.111.111.1 0023.ac64.46c2  50 arp  port-channel111

N7K-1-VDC3# show forwarding ip route 172.31.31.0/24 module 1
IPv4 routes for table default/base
<table>
<thead>
<tr>
<th>Prefix</th>
<th>Next-hop</th>
<th>Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>172.31.31.0/24</td>
<td>12.1.1.1</td>
<td>port-channel11</td>
</tr>
<tr>
<td></td>
<td>12.111.111.1</td>
<td>port-channel111</td>
</tr>
</tbody>
</table>

N7K-1-VDC3# show routing hash 9.9.9.9
172.31.31.250
Load-share parameters used for software forwarding:
load-share mode: address source-destination port source-destination
Universal-id seed: 0x6bd14cf5
Hash for VRF "default" resulting hash: 0x01 path '>'

172.31.31.250
Load-share parameters used for software forwarding:
load-share mode: address source-destination port source-destination
Universal-id seed: 0x6bd14cf5
Hash for VRF "default" resulting hash: 0x00 path '>'

Ingress module
Unicast L2 and L3 Forwarding, ARP
L3 — Hardware Entry, ECMP

N7K-1-VDC3# show system internal forwarding ip route 172.31.31.0/24 detail module 1
RPF Flags legend:
 S - Directly attached route (S_Star)
 V - RPF valid
 M - SMAC IP check enabled
 G - SGT valid
 E - RPF External table valid
172.31.31.0/24 , port-channel1
Dev: 1 , Idx: 0xf1f6 , RPF Flags: V , DGT: 0 , VPN: 33
RPF_intf_5: port-channel1 (0x4018 )
AdjIdx: 0x43032, LIFB: 0 , LIF: port-channel1 (0x4018 ), DI: 0xa46
DMAC: 0023.ac64.46c2 SMAC: 0023.ac64.46c3
AdjIdx: 0x43033, LIFB: 0 , LIF: port-channel1111 (0x40ba ), DI: 0xa57
DMAC: 0023.ac64.46c2 SMAC: 0023.ac64.46c3

N7K-1-VDC3# show system internal forwarding adjacency entry 0x43032 detail module 1
Device: 1   Index: 0x43032   DMAC: 0023.ac64.46c2 SMAC: 0023.ac64.46c3
LIF: 0x4018 (port-channel1) DI: 0xa46 ccc: 4   L2_FWD: NO  RDT: YES
packets: 356523bytes: 534784500zone enforce: 0

N7K-1-VDC3# show system internal forwarding adjacency entry 0x43033 detail module 1
Device: 1   Index: 0x43033   DMAC: 0023.ac64.46c2 SMAC: 0023.ac64.46c3
LIF: 0x40ba (port-channel111) DI: 0xa57 ccc: 4   L2_FWD: NO  RDT: YES
packets: 0 bytes: 0 zone enforce: 0

N7K-1-VDC3# show interface port-channel 1 | grep "output rate" | grep -v input
30 seconds output rate 960124944 bits/sec, 80025 packets/sec

N7K-1-VDC3# show interface port-channel 111 | grep "output rate" | grep -v input
30 seconds output rate 3840056544 bits/sec, 320040 packets/sec

Ingress traffic has 5 streams 80kpps each, hardware performs 4:1 load-sharing across 2 ECMP path
Unicast L2 and L3 Forwarding, ARP
L3 — Additional Useful CLI

N7K-1-VDC3# show ip route summary
IP Route Table for VRF "default"
Total number of routes: 20135
Total number of paths: 40195
Best paths per protocol:      Backup paths per protocol:
    am             : 3            ospf-6467      : 10
    local          : 8
    direct         : 8
    broadcast      : 12
    eigrp-6467     : 4
    ospf-6467      : 40146
    bgp-1204       : 4

Number of routes per mask-length:
   /31: 1     /32: 55

N7K-1-VDC3# show hardware internal forwarding table utilization module 1 | grep -v -i key
Module 1 usage:
Route Type Used %Used Free %Free Total
(Log/Phys) (Log/Phys) (Log/Phys) (Log/Phys)
IPv4 Unicast: 40373/40373 54 33355/33355 45 73728/73728
L2VPN Peer: 0/0 0 0/0 0 0/0
MPLS: 0/0 0 0/0 0 0/0
IPv4 Multicast: 22/44 0 16362/32724 99 16384/32768
L2VPN IPv4 Mcast: 0/0 0 0/0 0 0/0
IPv6 Unicast: 83/166 1 8109/16218 98 8192/16384
L2VPN IPv6 Mcast: 0/0 0 0/0 0 0/0
IPv6 Multicast: 15/60 0 2033/8132 99 2048/8192

Officially documented cli is show hardware capacity forwarding
Unicast L2 and L3 Forwarding, ARP
L3 — Adjacency Manager (AM) Installed Route

**Before arp for host 172.32.32.32 is resolved, 172.32.32.32/32 subnet route points to Po1**

```
N7K-1-VDC2# show ip arp 172.32.32.32 | egrep "ARP|Address|Vlan"
```

<table>
<thead>
<tr>
<th>IP ARP Table</th>
<th>Address</th>
<th>Age</th>
<th>MAC Address</th>
<th>Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>N7K-1-VDC2#</td>
<td>172.32.32.32</td>
<td>00:00:07</td>
<td>0000.98aa.8ac9</td>
<td>Vlan32</td>
</tr>
</tbody>
</table>

```
N7K-1-VDC2# show ip route 172.32.32.0/24 | egrep "attached|direct"
```

```
172.32.32.0/24, ubest/mbest: 1/0, attached
 *via 172.32.32.11, Vlan32, [0/0], 00:14:30, direct
```

```
N7K-1-VDC2# show ip route 172.32.32.0/25 | egrep "ubest|static"
```

```
172.32.32.0/25, ubest/mbest: 1/0
 *via 12.1.1.2, Po1, [1/0], 00:02:22, static
```

```
N7K-1-VDC2# show system internal forwarding ip route 172.32.32.32/32 detail module 1
```

```
[snip]

172.32.32.0/25 , port-channel1
Dev: 1 , Idx: 0x18802 , RPF Flags: V , DGT: 0 , VPN: 17
RPF_Intf_5: port-channel1 (0x400e )
AdjIdx: 0x43013, LIFB: 0 , LIF: port-channel1 (0x400e ), DI: 0xa36
DMAC: 0023.ac64.46c3 SMAC: 0023.ac64.46c2

CSCti79838 will provide cli to set AM installed route AD

```
N7K-1-VDC2# show ip arp 172.32.32.32 | egrep "ARP|Address|Vlan"
```

<table>
<thead>
<tr>
<th>IP ARP Table</th>
<th>Address</th>
<th>Age</th>
<th>MAC Address</th>
<th>Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>172.32.32.32</td>
<td>00:00:07</td>
<td>0000.98aa.8ac9</td>
<td>Vlan32</td>
<td></td>
</tr>
</tbody>
</table>

```
N7K-1-VDC2# show system internal forwarding ip route 172.32.32.32/32 det module 1
```

```
[snip]

172.32.32.32/32 , Vlan32
Dev: 1 , Idx: 0x120c7 , RPF Flags: VS , DGT: 0 , VPN: 17
RPF_Intf_5: Vlan32 (0x3be )
AdjIdx: 0x43013, LIFB: 0 , LIF: Vlan32 (0x3be ), DI: 0x0
DMAC: 0000.98aa.8ac9 SMAC: 0011.3232.3232

After arp for host 172.32.32.32 is resolved, AM installs route pointing to directly connected subnet

```
N7K-1-VDC2# show ip route 172.32.32.32/32 | egrep "attached|am"
```

```
172.32.32.32/32, ubest/mbest: 1/0, attached
 *via 172.32.32.32, Vlan32, [2/0], 00:01:47, am
```
Unicast L2 and L3 Forwarding, ARP

L3 — ARP, Glean Throttling (5.1.2 Onwards)

N7K-3-PeerB# show ip arp vlan 32 | grep -v ^$
Flags: * - Advertisements learnt on non-active FHRP router
+ - Advertisements synced via CFSoE
# - Advertisements Throttled for Glean
D - Static Advertisements attached to down interface

IP ARP Table
Total number of entries: 4
<table>
<thead>
<tr>
<th>Address</th>
<th>Age</th>
<th>MAC Address</th>
<th>Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>172.32.32.11</td>
<td>00:07:01</td>
<td>0011.3232.3232</td>
<td>Vlan32</td>
</tr>
<tr>
<td>172.32.32.14</td>
<td>00:06:35</td>
<td>0013.5f1f.46c0</td>
<td>Vlan32</td>
</tr>
<tr>
<td>172.32.32.150</td>
<td>00:01:26</td>
<td>INCOMPLETE</td>
<td>Vlan32</td>
</tr>
<tr>
<td>172.32.32.151</td>
<td>00:01:26</td>
<td>INCOMPLETE</td>
<td>Vlan32</td>
</tr>
</tbody>
</table>

N7K-3-PeerB# show run all | grep "hardware ip"
hardware ip glean throttle
hardware ip glean throttle maximum 1000
hardware ip glean throttle timeout 300
hardware ip glean throttle syslog 500

1000 – 1k throttled adjacencies
300 – how long adjacency would be installed
500 – traffic rate threshold to syslog a message

N7K-3-PeerB# show ip adjacency 172.32.32.150 detail | b default | grep -v ^$
IP Adjacency Table for VRF default
Total number of entries: 1
| Address :          | 172.32.32.150 |
| MacAddr :          | 0000.0000.0000 |
| Preference :       | 255           |
| Source :           | arp           |
| Interface :        | Vlan32        |
| Physical Interface :| Vlan32        |
| Packet Count :     | 62027         |
| Byte Count :       | 5954592       |
| Best :             | Yes           |
| Throttled :        | Yes           |

N7K-3-PeerB# show ip adjacency 172.32.32.14 detail | b default | grep -v "^$"
IP Adjacency Table for VRF default
Total number of entries: 1
| Address :          | 172.32.32.14  |
| MacAddr :          | 0013.5f1f.46c0 |
| Preference :       | 50            |
| Source :           | arp           |
| Interface :        | Vlan32        |
| Physical Interface :| port-channel1667 |
| Packet Count :     | 0             |
| Byte Count :       | 0             |
| Best :             | Yes           |
| Throttled :        | No            |
Unicast L2 and L3 Forwarding, ARP
L3 — Forwarding Engine Error Statistics

N7K-1-PeerA# show hardware internal statistics module 1 device L3lu errors port 2

Hardware statistics on module 01:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Device: Lamira</td>
<td>Role: L3</td>
<td>Mod: 1</td>
</tr>
<tr>
<td>Last cleared @ Fri Feb 25 21:30:09 2011</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Device Statistics Category :: ERROR</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Instance: 0</th>
<th>ID</th>
<th>Name</th>
<th>Value</th>
<th>Ports</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>75</td>
<td>RP IPv4 L3 filtering Pkt drop</td>
<td>0000000000000002</td>
<td>1-32 I1</td>
</tr>
<tr>
<td></td>
<td>76</td>
<td>RP IPv6 L3 filtering Pkt drop</td>
<td>0000000000000001</td>
<td>1-32 I1</td>
</tr>
<tr>
<td></td>
<td>93</td>
<td>CL1 Same IF check Fail Pkt count</td>
<td>0000000038480964</td>
<td>1-32 I1</td>
</tr>
<tr>
<td></td>
<td>188</td>
<td>PL OFE Global aggr drop pkt ctr</td>
<td>00000018316176</td>
<td>1-32 I1</td>
</tr>
<tr>
<td></td>
<td>189</td>
<td>PL OFE Global aggr drop byte ctr</td>
<td>000027451514923</td>
<td>1-32 I1</td>
</tr>
<tr>
<td></td>
<td>198</td>
<td>PL OFE Total police drop pkt ctr</td>
<td>00000018316176</td>
<td>1-32 I1</td>
</tr>
<tr>
<td></td>
<td>199</td>
<td>PL OFE Total police drop byte ctr</td>
<td>000027451514923</td>
<td>1-32 I1</td>
</tr>
<tr>
<td></td>
<td>207</td>
<td>PL OFE TTL expire pkt ctr</td>
<td>0000000000037961</td>
<td>1-32 I1</td>
</tr>
<tr>
<td></td>
<td>259</td>
<td>L3 Fib Miss Pkt ctr</td>
<td>00000588018615</td>
<td>1-32 I1</td>
</tr>
<tr>
<td></td>
<td>260</td>
<td>L3 IPv4 Option Pkt ctr</td>
<td>000000000000357</td>
<td>1-32 I1</td>
</tr>
<tr>
<td></td>
<td>261</td>
<td>L3 IPv6 Option Pkt ctr</td>
<td>000000000046652</td>
<td>1-32 I1</td>
</tr>
<tr>
<td></td>
<td>262</td>
<td>L3 Non-Rpf Drop Pkt ctr</td>
<td>0000000240773</td>
<td>1-32 I1</td>
</tr>
<tr>
<td></td>
<td>305</td>
<td>NF L3 ACL deny pkt ctr</td>
<td>000006154091492</td>
<td>1-32 I1</td>
</tr>
<tr>
<td></td>
<td>449</td>
<td>Exception cause: ICMP UNREACH (Unicast)</td>
<td>0000000538866</td>
<td>1-32 I1</td>
</tr>
<tr>
<td></td>
<td>454</td>
<td>Exception cause: L3 BRIDGE DROP (Unicast)</td>
<td>00000073007752</td>
<td>1-32 I1</td>
</tr>
<tr>
<td></td>
<td>455</td>
<td>Exception cause: DROP (Unicast)</td>
<td>0000000000000003</td>
<td>1-32 I1</td>
</tr>
<tr>
<td></td>
<td>461</td>
<td>Exception cause: OPTIONS (Multicast)</td>
<td>0000000047009</td>
<td>1-32 I1</td>
</tr>
<tr>
<td></td>
<td>463</td>
<td>Exception cause: TWO MCAST RPF (Multicast)</td>
<td>000000000000016</td>
<td>1-32 I1</td>
</tr>
<tr>
<td></td>
<td>464</td>
<td>Exception cause: L3 BRIDGE DROP (Multicast)</td>
<td>0000001080488</td>
<td>1-32 I1</td>
</tr>
</tbody>
</table>

CoPP dropped traffic
No route traffic drops
Acl dropped packets, when acl-log is configured packets hits also access-list-log rate-limiter
Packets packets received across vpc PL from mcast vpc forwarder
Unicast L2 and L3 Forwarding, ARP
L3 — show tech-support Data Collection

N7K-1-VDC2-CS1# show tech-support forwarding L3 unicast | grep "\`show "
`show forwarding route summary vrf all` 
`show forwarding route max-display-count 100000 vrf all` 
`show forwarding vrf all adjacency` 
`show forwarding ipv6 route summary vrf all` 
`show forwarding ipv6 route max-display-count 100000 vrf all` 
`show forwarding vrf all ipv6 adjacency` 
`show forwarding trace` 
`show forwarding internal errors` 
`show forwarding internal error counts` 
`show forwarding internal unicast counts vdc all` 
`show forwarding internal message counts`

N7K-1-VDC2-CS1# show tech-support netstack | grep "\`show" | grep tech-support
`show tech-support arp` 
`show tech-support adjmgrp` 
`show tech-support icmpv6` 
`show tech-support ip` 
`show tech-support ipv6` 
`show tech-support pktmgr` 
`show tech-support sockets`

N7K-1-VDC2-CS1# show tech-support netstack | grep "\`show " | wc -l
212

N7K-1-VDC2-CS1# show tech-support arp | grep "\`show "
`show running-config arp` 
`show ip arp internal event-history cli` 
`show ip arp vrf all` 
`show ip arp static vrf all` 
`show ip arp summary vrf all` 
`show ip arp tunnel-statistics`
Agenda

- Before You Get Started
  - Traditional Versus NX-OS Troubleshooting Approach
  - Nexus 7000 Module and Forwarding Engine Architecture Overview
  - Built-in Troubleshooting Tools

- Troubleshooting
  - CPU
  - Control-Plane
  - Memory Utilization
  - vPC
  - Unicast Layer 2 and Layer 3 Forwarding and ARP
  - Multicast Layer 2 and Layer 3 Forwarding
  - Switch Fabric
  - ACL
  - QoS
Multicast L2 and L3 Forwarding
Multicast Replication

- **L2 multicast replication**
  - Copy of original packet for each output fabric-channel or switchport
  - Performed by xbar (stage number 2) and port asics
  - Driven by Multicast indexes (MI) at fabric level and LTL indexes at port level
  - Multicast Distribution or MD copy is created by ingress replication asic

- **L3 egress (only) multicast replication**
  - Copy of original packet for each layer 3 interface (OIF)
  - Performed by replication asic aka ‘rewrite’ or ‘RWR_0’
  - Multicast Expansion Table (MET) in replication engines contains OIFs
  - Nexus 7000 system supports egress Layer 3 replication

- **I/O Module RWR_0 METs may have different content (asymmetric)**
  - Conserves replication asic and forwarding engine bandwidth (forwarding engine must provide lookup result for each individual packet copy)
  - If OIF is SVI which L2 Vlan spans across multiple I/O modules, each I/O module creates copy of original packet even no receivers are present
Multicast L2 and L3 Forwarding
Multicast Replication, MET Table

Across modules, MET block size and contents can be asymmetric.

On single module, MET block size must be identical but contents can be asymmetric.
Multicast L2 and L3 Forwarding
Packet Flow with L3 Ingress, vPC Egress
L2/L3 — Platform Independent, vPC Specific Check

N7K-1-PeerA# show vpc role | grep -v mac|^$
vPC Role status
----------------------------------------------------
 vPC role                        : secondary
 Dual Active Detection Status    : 0
 vPC system-priority             : 32667
 vPC local role-priority         : 128

N7K-1-PeerA# show ip igmp internal vpc | grep -v role|emul|peer
IGMP vPC operational state UP
IGMP ES operational state DOWN
IGMP is registered with vPC library
IGMP is registered with MCEC_TL/CFS
IGMP vPC Operating Version: 2
IGMP vPC Domain ID: 64
IGMP vPC Peer-link Exclude feature enabled

N7K-1-PeerA# show ip pim internal vpc | grep -v role|not
PIM vPC operational state UP
PIM emulated-switch operational state DOWN
PIM is registered with vPC manager
PIM is registered with MCEC_TL/CFS
VPC peer link is up on port-channel664
PIM vPC Operating Version: 2
PIM vPC Domain ID: 64

N7K-1-PeerA# show ip pim internal vpc rpf | grep -v ^$|vPC
Source: 172.23.25.65
 Pref/Metric: 110/63
 Source role: secondary
 Forwarding state: Win (forwarding)

Important: Win state is per s,g

N7K-3-PeerB# show vpc role | grep -v mac|^$
vPC Role status
----------------------------------------------------
vPC role                        : primary
 Dual Active Detection Status    : 0
 vPC system-priority             : 32667
 vPC local role-priority         : 64

N7K-3-PeerB# show ip igmp internal vpc | grep -v role|emul|peer
IGMP vPC operational state UP
IGMP ES operational state DOWN
IGMP is registered with vPC library
IGMP is registered with MCEC_TL/CFS
IGMP vPC Operating Version: 2
IGMP vPC Domain ID: 64
IGMP vPC Peer-link Exclude feature enabled

N7K-3-PeerB# show ip pim internal vpc | grep -v role|not
PIM vPC operational state UP
PIM emulated-switch operational state DOWN
PIM is registered with vPC manager
PIM is registered with MCEC_TL/CFS
VPC peer link is up on port-channel664
PIM vPC Operating Version: 2
PIM vPC Domain ID: 64

N7K-3-PeerB# show ip pim internal vpc rpf | grep -v ^$|vPC
Source: 172.23.25.65
 Pref/Metric: 110/83
 Source role: primary
 Forwarding state: Lose (not forwarding)

Important: Worse metric and therefore B is Loser
Multicast L2 and L3 Forwarding
Packet Flow with L3 Ingress, vPC Egress
L2 — Platform Independent

Both peers have igmp state synchronized via CFS regardless to which of them igmp joined arrived to based on port-channel hashing.

PeerA is a vPC forwarder but PeerB has the same (*,g) entry.
Multicast L2 and L3 Forwarding
Packet Flow with L3 Ingress, vPC Egress
L2 — Platform Dependent, Platform PI Index Check

```
N7K-1-PeerA# show ip igmp snooping groups 239.28.28.64 vlan 32 detail
IGMP Snooping group membership for vlan 32
Group addr: 239.28.28.64
Group ver: v2 [old-host-timer: not running]
Last reporter: 172.32.32.250
IGMPv1/v2 memb ports:
    port-channel1667 [1 GQ missed]
  vPC grp peer-link flag: include
M2RIB  vPC grp peer-link flag: include

N7K-1-PeerA# show ip igmp snooping mrouter vlan 32
Type: S - Static, D - Dynamic, V - vPC Peer Link
I - Internal, F - Fabricpath core port
U - User Configured
Vlan  Router-port Type       Uptime       Expires
32    Po664     SV   11:25:21     never
32    Vlan32    ID   11:25:10    00:04:28

N7K-1-PeerA# show forwarding distribution ip igmp snooping vlan 32 group 239.28.28.64
Vlan: 32, Group: 239.28.28.64, Source: 0.0.0.0
   Outgoing Interface List Index: 11
   Reference Count: 1
   Platform Index: 0x7ffa
   Number of Outgoing Interfaces: 2
   port-channel1664
   port-channel1667

N7K-1-PeerA# show forwarding distribution multicast outgoing-interface-list L2 11 | grep –v ^$
   Outgoing Interface List Index: 11
   Reference Count: 1
   Platform Index: 0x7ffa
   Number of Outgoing Interfaces: 2
   port-channel1664
   port-channel1667

N7K-1-PeerA# show system internal ip igmp snooping vlan 32 group 239.28.28.64 module 1 | grep –v ^$
   Vlan   Group   Source   Epoch   RID   DTL   hwptr
   32 239.28.28.64   1   11   0x7ffa   0x13fff
```

Platform index (PI) chosen by IGMP process is used to distribute multicast through system and must be same at all levels.

MFDM PI index matches IGMP PI index.

Ingress I/O Module (from igmp perspective) PI index matches IGMP and MFDM index.
Multicast L2 and L3 Forwarding
Packet Flow with L3 Ingress, vPC Egress
L2 — Platform Dependent, OIF LTL and Fabric MI Index Check

```
N7K-1-PeerA# show system internal pixm info ltl 0x7ffa det | grep -v ^$  
MCAST LTLs allocated for VDC:2
============================================
LTL IFIDX  LTL_FLAG CB_FLAG MI[0]  
0x7ffa 0x0000000b 0x0002 0x001
Member info
---------------------------------  
IFIDX  LTL
Po667 0x0a47
Po664 0x0a45
LTL CB information
============================================
LTL IFIDX  VDC ESPAN LTL_FLAG CB_FLAG LTL_TYPE
0x7ffa 0x0000000b 02 0x00000000 0x00 0x0002 MCAST_GROUP
0x00000000 02 0x00000000 0x00 0x0002 MCAST_GROUP
VQI/MI: 0x001 0x001 0x001 0x001 0x001 0x001 0x001 0x001
index MI  v4_fpoe v5_fpoe (fpoe cnt:1)
00000 0x0001 0x0006 0x0000
GENERIC OPER Member Count:0

N7K-1-PeerA# show system internal xbar static-mc
---------------------------------------------------------------
<table>
<thead>
<tr>
<th>Multicast Index</th>
<th>slot-mask</th>
<th>Slots in the group(1-based)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0001</td>
<td>0x000001</td>
<td>1</td>
</tr>
<tr>
<td>0002</td>
<td>0x000002</td>
<td>2</td>
</tr>
<tr>
<td>0003</td>
<td>0x000004</td>
<td>3</td>
</tr>
<tr>
<td>0004</td>
<td>0x000008</td>
<td>4</td>
</tr>
</tbody>
</table>
```

Group 239.28.28.64 within Vlan 32 is distributed only into module 1, within Vlan4093 to module 1 and 4

```
N7K-1-PeerA# show forwarding distribution ip igmp snooping vlan 4093 group 239.28.28.64
Vlan: 4093, Group: 239.28.28.64, Source: 0.0.0.0
Outgoing Interface List Index: 12
Reference Count: 1
Platform Index: 0x7ff8
Number of Outgoing Interfaces: 2
port-channel1664
port-channel4093
```

```
N7K-1-PeerA# show system internal ip igmp snooping vlan 4093 group 239.28.28.64 module 4
Vlan  Group  Source  Epoch
RID  DTL  hwptr
4093 239.28.28.64 1
 0x7ff8 0x18eff
```

```
N7K-1-PeerA# show system internal pixm info ltl 0x7ff8 detail | egrep Po|VQI
Po664 0x0a45
Po4093 0x0a49
VQI/MI: 0x00b 0x00b 0x00b 0x00b 0x00b 0x00b 0x00b 0x00b
```

```
N7K-1-PeerA# show system internal pixm info ltl 0a0a47 | grep -a 2 "Member rbh"
Member rbh rbh_cnt
Eth1/2 0x000000ff 0x08

MI index in HEX
```

```
N7K-1-PeerA# show system internal pixm info ltl 0x7ffa detail | egrep Po|VQI
Po664 0x0a45
Po4093 0x0a49
VQI/MI: 0x00b 0x00b 0x00b 0x00b 0x00b 0x00b 0x00b 0x00b
```

```
N7K-1-PeerA# show system internal pixm info ltl 0x7ffa detail | egrep Po|VQI
Po664 0x0a45
Po4093 0x0a49
VQI/MI: 0x00b 0x00b 0x00b 0x00b 0x00b 0x00b 0x00b 0x00b
```

```
N7K-1-PeerA# show system internal pixm info ltl 0x7ffa detail | egrep Po|VQI
Po664 0x0a45
Po4093 0x0a49
VQI/MI: 0x00b 0x00b 0x00b 0x00b 0x00b 0x00b 0x00b 0x00b
```

```
N7K-1-PeerA# show system internal pixm info ltl 0x7ffa detail | egrep Po|VQI
Po664 0x0a45
Po4093 0x0a49
VQI/MI: 0x00b 0x00b 0x00b 0x00b 0x00b 0x00b 0x00b 0x00b
```
Multicast L2 and L3 Forwarding
Packet Flow with L3 Ingress, vPC Egress
L2 — Platform Dependent, Egress I/O Module PI and MI Index Mapping

```
module-4# show hardware internal rewrite_engine ltl read instance_bitmap 0xf start 0x7ff8 end 0x7ff8 rbh 0xff
--- inst ltl rhh e1 e0 k fpoe mi/vqi ps ports ---
[snip]
1 0x7ff8 0 0 0 1 0x0000b (mi 0x00b) 0x1
1 0x7ff8 1 0 0 1 0x0000b (mi 0x00b) 0x1
[snip]
1 0x7ff8 4 0 0 1 0x0000b (mi 0x00b) 0x2
1 0x7ff8 5 0 0 1 0x0000b (mi 0x00b) 0x2
[snip]
module-4# show hardware internal qengine asic 0 em mc -mapping first 0x7ff8 last 0x7ff8
DI | OC DF | Metros | Ports
--------- +--------- +------------------------
0x00007ff8 | d0 | m0 | 25, 48
```

```
module-1# show hardware internal rewrite_engine ltl read instance_bitmap 0xf start 0x7ffa end 0x7ffa rbh 0xff
--- inst ltl rhh e1 e0 k fpoe mi/vqi ps ports ---
0 0x7ffa 0 0 0 1 0x00001 (mi 0x001) 0x2000 | 2, 4, 6, 8
0 0x7ffa 1 0 0 1 0x00001 (mi 0x001) 0x2000 | 2, 4, 6, 8
0 0x7ffa 2 0 0 1 0x00001 (mi 0x001) 0x2000 | 2, 4, 6, 8
0 0x7ffa 3 0 0 1 0x00001 (mi 0x001) 0x2000 | 2, 4, 6, 8
[snip]
2 0x7ffa 3 0 0 1 0x00001 (mi 0x001) 0x0 | no port selects
3 0x7ffa 0 0 0 1 0x00001 (mi 0x001) 0x2000 | 9, 11, 13, 15
3 0x7ffa 1 0 0 1 0x00001 (mi 0x001) 0x2000 | 9, 11, 13, 15
3 0x7ffa 2 0 0 1 0x00001 (mi 0x001) 0x2000 | 9, 11, 13, 15
3 0x7ffa 3 0 0 1 0x00001 (mi 0x001) 0x2000 | 9, 11, 13, 15
module-1# show hardware internal qengine asic 0 em mc -mapping first 0x7ffa last 0x7ffa
DI | OC DF | Metros | Ports
--------- +--------- +------------------------
0x00007ffa | d0 | m0 | 2, 4, 6, 8, 10, 12, 14, 16
```

DI – Platform Index choosen by IGMP Snooping (changes with s,g expiration)
OC DF – VoQ Asic link to Rewrite Engine Asic (0-3)
Metro – RWR_0 Asic instance
Ports – Physical ports connected to RWR_0 Asic

Output for qengine asic 1 would show ports 'd0 | m3 | 1,3,5,7,9,11,13,15'
Multicast L2 and L3 Forwarding
L2 — Data Collection

```
N7K-1-VDC2# show tech-support forwarding L2 multicast | grep "`show "
`show system internal l2mcast info statistics`
`show system internal ip igmp snooping summary`
`show system internal ip igmp snooping`
`show system internal ip igmp snooping pending-hwinstall`
[nlip]
`show system internal m2fib rid pending`
`show system internal mfdm info statistics`
[nlip]

N7K-1-VDC2# show tech-support ip igmp brief | grep "`show "
`show running-config igmp`
`show version internal build-identifier`
`show logging logfile | grep -i igmp`
`show ip igmp internal`
`show ip igmp route vrf all`
`show ip igmp interface vrf all`
`show system internal sysmgr service name igmp`
`show system internal feature-mgr feature state | include igmp`
`show ip igmp internal mem-stats all`
`show ip igmp internal vpc`

N7K-1-VDC2# show tech-support ip igmp snooping | grep "`show "
`show running-config igmp`
`show version internal build-identifier`
`show logging logfile | grep -i igmp`
`show ip igmp snooping groups detail`
`show ip igmp snooping querier`
`show ip igmp snooping mrouter detail`
`show ip igmp snooping`
[snip]
`show ip igmp snooping otv groups detail`
```
Multicast L2 and L3 Forwarding
Packet Flow with L3 Ingress, vPC Egress
L3 — Platform Independent, mcast Routing Table

N7K-1-PeerA# show ip mroute 239.28.28.64 source-tree
IP Multicast Routing Table for VRF "default"
(172.23.25.65/32, 239.28.28.64/32), uptime: 00:02:14, pim ip mrib msdp
Incoming interface: port-channel11, RPF nbr: 12.1.1.2
Outgoing interface list: (count: 3)
  Vlan4093, uptime: 00:02:14, mrib
  Vlan32, uptime: 00:02:14, mrib
  port-channel166, uptime: 00:02:14, pim

N7K-1-PeerA# show ip mroute 239.28.28.64 summary software-forwarded | b Source
Source          packets      bytes           aps     pps         bit-rate     oifs
(*)G            3            3966            1322    0         0.000   bps  2
  sw-pkts: 3
172.23.25.65    10131630     13434541376    1325    2005      21276112.000 bps 3
  sw-pkts: 1

N7K-1-PeerA# show ip mroute 239.28.28.64 summary rpf-failed | grep -v ^$  
IP Multicast Routing Table for VRF "default"
Total number of routes: 3
Total number of (*,G) routes: 1
Total number of (S,G) routes: 1
Total number of (*,G-prefix) routes: 1
Group count: 1, rough average sources per group: 1.0
Group: 239.28.28.64/32, Source count: 1
Source          packets      bytes           aps     pps         bit-rate     oifs
(*)G            3            3966            1322    0         0.000   bps  2
  RPF Failed Packets: 0
  RPF Failed Bytes: 0
172.23.25.65    10853176     14391311372    1325    2004      21258608.800 bps 3

N7K-1-PeerB# show ip pim interface vlan 32
PIM Interface Status for VRF "default"
Vlan32, Interface status: protocol-up/link-up/admin-up
  IP address: 172.32.32.11, IP subnet: 172.32.32.0/24
  PIM DR: 172.32.32.22, DR's priority: 1

PeerB forwarding state is ‘Lose not forwarding’ for the group (metric based)
Enhanced PIM assert mechanism over CFS protocol (PIM assert handshake) is used to avoid periodic duplicates sent to receivers behind vPC (show ip pim statistics and show ip pim internal event-history assert-receive)
Multicast L2 and L3 Forwarding
Packet Flow with L3 Ingress, vPC Egress
L3 — Platform Dependent, s,g per OIF Counters, OIF Index Check

N7K-1-PeerA# show forwarding multicast route group 239.28.28.64 source 172.23.25.65 module 1
(172.23.25.65/32, 239.28.28.64/32), RPF Interface: port-channel11, flags:
   Received Packets: 0 Bytes: 0
   Number of Outgoing Interfaces: 3
   Outgoing Interface List Index: 12
      Vlan32 Outgoing Packets:292302 Bytes:391684680
      Vlan4093 Outgoing Packets:146151 Bytes:195842340
      port-channel66 Outgoing Packets:0 Bytes:0

N7K-1-PeerA# show forwarding multicast route group 239.28.28.64 source 172.23.25.65 module 4
(172.23.25.65/32, 239.28.28.64/32), RPF Interface: port-channel11, flags:
   Received Packets: 149727 Bytes: 200634180
   Number of Outgoing Interfaces: 3
   Outgoing Interface List Index: 12
      Vlan32 Outgoing Packets:0 Bytes:0
      Vlan4093 Outgoing Packets:149729 Bytes:200636860
      port-channel66 Outgoing Packets:149736 Bytes:200646240

N7K-1-PeerA# show forwarding multicast outgoing-interface-list module 1 12
   Outgoing Interface List Index: 12
   Reference Count: 1
      Vlan32
      Vlan4093
      port-channel66

N7K-1-PeerA# show forwarding multicast outgoing-interface-list module 4 12
   Outgoing Interface List Index: 12
   Reference Count: 1
      Vlan32
      Vlan4093
      port-channel66

N7K-1-PeerA# show interface vlan4093 | grep rate | grep -v in
   60 seconds output rate 42909698 bits/sec, 4084 packets/sec
N7K-1-PeerA# show interface vlan32 | grep rate | grep -v in
   60 seconds output rate 42879827 bits/sec, 4038 packets/sec
N7K-1-PeerA# show interface po66 | grep rate | grep -v in
   30 seconds output rate 21467896 bits/sec, 2043 packets/sec
Multicast L2 and L3 Forwarding
Packet Flow with L3 Ingress, vPC Egress
L3 — Platform Dependent, (s,g) FIB Programming, Ingress Module

N7K-1-PeerA# show system internal forwarding multicast route group 239.28.28.64 source 172.23.25.65 module 4 detail

Hardware Multicast FIB Entries:
Flags Legend:
* - s_star_priority
S - sg_entry
D - Non-RPF Drop
B - Bi-dir route  W - Wildcard route

(172.23.25.65/32, 239.28.28.64/32), Flags: *S
- Lamira: 1, HWIndex: 0x2202, VPN: 17
- RPF Interface: port-channel11, LIF: 0x40d9, PD oiflist index: 0x5
- ML3 Adj Idx: 0xa022, MD: 0x2006, MET0: 0x2007, MET1: 0x2007, MTU Idx: 0x1
- Metro Instance: 0
  - Dev: 1 Index: 0xa038  Type: MDT  elif: 0xc0008
    - dest idx: 0x7ff0  recirc-dti: 0xe20000
  - Dev: 1 Index: 0xa034  Type: OIF  elif: 0x840de
    - dest idx: 0xa42  smac: 0023.ac64.46c2

- Metro Instance: 1
  - Dev: 1 Index: 0xa038  Type: MDT  elif: 0xc0008
    - dest idx: 0x7ff0  recirc-dti: 0xe20000
  - Dev: 1 Index: 0x6101  Type: OIF  elif: 0x80101
    - dest idx: 0x0  smac: 0023.ac64.46c2

Module 4 is ingress I/O module from multicast flow perspective (ingress interface is Po1) but also egress I/O Module as OIFs Vlan4093/Po4093 and Po66 are on it

HWIndex – Pointer to MFIB table
ML3 Adj Idx – Adjacency table pointer returned by FIB lookup of incoming mcast packet
MD – Distribution index to get mcast packet to fabric
MET0 – pointer to MET for tunnel interface if present
MET1 – pointer to MET location with OIF information

Use slot X show hardware internal dev-port-map to determine physical port-to-metro mapping
Metro Instance 0 – OIF1: Po66 - Eth4/11-12
Metro Instance 1 – OIF2: Vlan4093/Po4093 – Eth4/37-38
Multicast L2 and L3 Forwarding
Packet Flow with L3 Ingress, vPC Egress
L3 — Platform Dependent, (s,g) FIB Programming, Egress Module

N7K-1-PeerA# show system internal forwarding multicast route group 239.28.28.64 source 172.23.25.65 module 1 detail | b
(172.23.25.65/32, 239.28.28.64/32), Flags: *S
  Lamira: 1, HWIndex: 0x2202, VPN: 17
  RPF Interface: port-channel11, LIF: 0x40d9, PD oiflist index: 0x5
  ML3 Adj Idx: 0xa022, MD: 0x2007, MET0: 0x2008, MET1: 0x2008, MTU Idx: 0x1
  Metro Instance: 0
    Dev: 1 Index: 0xa038 Type: MDT elif: 0xc0008
dest idx: 0x7ff0 recirc-dti: 0xe20000
    Dev: 1 Index: 0x60d9 Type: OIF elif: 0x800d9 Vlan32
dest idx: 0x0 smac: 0011.3232.3232

Metro Instance: 1
Dev: 1 Index: 0xa038 Type: MDT elif: 0xc0008
dest idx: 0x7ff0 recirc-dti: 0xe20000

Metro Instance: 2
Dev: 1 Index: 0xa038 Type: MDT elif: 0xc0008
dest idx: 0x7ff0 recirc-dti: 0xe20000

Metro Instance: 3
Dev: 1 Index: 0xa038 Type: MDT elif: 0xc0008
dest idx: 0x7ff0 recirc-dti: 0xe20000
Dev: 1 Index: 0x60d9 Type: OIF elif: 0x800d9 Vlan32
dest idx: 0x0 smac: 0011.3232.3232
Dev: 1 Index: 0x6101 Type: OIF elif: 0x80101 Vlan4093
dest idx: 0x0 smac: 0023.ac64.46c2

N7K-1-PeerA# show system internal forwarding adjacency entry 0x60d9 module 1 detail
Device: 1 Index: 0x60d9 DMAC: 0000.0000.0000 SMAC: 0011.3232.3232
LIF: 0x800d9 (Vlan32) DI: 0x0 ccc: 4 L2_FWD: NO RDT: NO packets: 12848bytes: 17216320zone enforce: 0

Module 1 is only egress module from multicast flow perspective
ML3 Adj Idx is same for all modules
MET indexes do not need to be same for all modules
Empty MET tables in Metro 1 and 2 (no receivers, it saves replication and lookup resources
Index – OIF specific pointer to Adj table

DI – Dest index is zero as this information comes from L3 asic indicating L2 asic index will be used instead
Multicast L2 and L3 Forwarding
Packet Flow with L3 Ingress, vPC Egress
L3 — Platform Dependent, Replication Engine Counters

N7K-1-PeerA# show hardware internal statistics module 1 device rewrite pktflow asic-all | egrep Dev|Inst|Multicast [snip]
<table>
<thead>
<tr>
<th>Device:Metropolis</th>
<th>Role:REWR</th>
<th>Mod: 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>97 Multicast L3 MET replication pkt cnt</td>
<td>0000000000000500</td>
<td>2,4,6,8,10,12,14,16 I1</td>
</tr>
<tr>
<td>98 Multicast L3 PR replication pkt cnt</td>
<td>0000000000000500</td>
<td>2,4,6,8,10,12,14,16 I1</td>
</tr>
</tbody>
</table>

N7K-1-PeerA# show hardware internal statistics module 4 device rewrite pktflow asic-all | egrep Dev|Inst|Multicast [snip]
<table>
<thead>
<tr>
<th>Device:Metropolis</th>
<th>Role:REWR</th>
<th>Mod: 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>96 Multicast L2 MET replication pkt cnt</td>
<td>0000000000000500</td>
<td>1,3,5,7,9,11,13,15</td>
</tr>
<tr>
<td>97 Multicast L3 MET replication pkt cnt</td>
<td>00000000000001000</td>
<td>1,3,5,7,9,11,13,15</td>
</tr>
<tr>
<td>98 Multicast L3 PR replication pkt cnt</td>
<td>00000000000001000</td>
<td>1,3,5,7,9,11,13,15</td>
</tr>
<tr>
<td>99 Multicast L2 PR replication pkt cnt</td>
<td>0000000000000500</td>
<td>1,3,5,7,9,11,13,15</td>
</tr>
</tbody>
</table>

N7K-1-PeerA# show hardware internal statistics module 1 device rewrite pktflow asic-all | egrep Dev|Inst|Multi

Instance 0

96 Multicast L2 MET replication pkt cnt 0000000000000500 I1
97 Multicast L3 MET replication pkt cnt 0000000000000500 I1
98 Multicast L3 PR replication pkt cnt 0000000000000500 I1
99 Multicast L2 PR replication pkt cnt 0000000000000500 I1

Instance 1

97 Multicast L3 MET replication pkt cnt 00000000000001000 25-48 I1
98 Multicast L3 PR replication pkt cnt 00000000000001000 25-48 I1

Statistics from L3 forwarding engine

<table>
<thead>
<tr>
<th>Start Interval</th>
<th>End Interval</th>
<th>PPS*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wed Apr 13 01:17:52 2011</td>
<td>Wed Apr 13 01:18:01 2011</td>
<td>4.3 K</td>
</tr>
<tr>
<td>Wed Apr 13 01:18:02 2011</td>
<td>Wed Apr 13 01:18:11 2011</td>
<td>4.3 K</td>
</tr>
</tbody>
</table>

1. 1x500 – Instance 0 Ingress (4/1)
2. 1x500 – Instance 0 MD to Fabric
3. 1x500 – Instance 0 L3 OIF (Po66)
4. 1x500 – Instance 1 MD from Fabric
5. 1x500 – Instance 1 L3 OIF (Po4093)
6. 1x500 – Instance 1 MD from Fabric (in other vDC)
7. 1x500 – Instance 1 L3 OIF (4/25 in other vDC)
8. 1x500 – Control-plane and other traffic

L2/L3 MET – number of packets sent to replication
L2/L3 PR - number of copies created
Multicast L2 and L3 Forwarding

L3 — Data Collection

N7K-1-VDC3# show tech-support ip pim | grep "`show "
`show running-config pim`
`show ip pim group-range vrf all`
`show ip pim interface vrf all`
`show ip pim neighbor vrf all`
`show ip pim route vrf all`
`show ip pim route internal vrf all`
`show ip pim rp vrf all`
`show ip pim df vrf all`
`show ip pim statistics vrf all`
`show system internal sysmgr service name pim`
[snip]

N7K-1-VDC3# show tech-support forwarding l3 multicast | grep "`show "
`show forwarding multicast outgoing-interface-list`
`show forwarding ip multicast route summary vrf all`
`show system internal forwarding ip multicast route summary`
`show forwarding ipv6 multicast route summary vrf all`
`show system internal forwarding adjacency multicast`
`show forwarding internal mem-stats detail`
`show forwarding internal errors`
`show forwarding internal multicast debugs`
`show forwarding internal multicast count`
`show forwarding multicast outgoing-interface-list`
[snip]

N7K-1-VDC3# show tech-support ip multicast | grep "`show "
`show tech-support ip igmp`
`show running-config igmp`
`show ip igmp route vrf all`
[snip]
`show tech-support ip msdp`
Agenda

- Before You Get Started
  - Traditional Versus NX-OS Troubleshooting Approach
  - Nexus 7000 Module and Forwarding Engine Architecture Overview
  - Built-in Troubleshooting Tools

- Troubleshooting
  - CPU
  - Control-Plane
  - Memory Utilization
  - vPC
  - Unicast Layer 2 and Layer 3 Forwarding and ARP
  - Multicast Layer 2 and Layer 3 Forwarding
  - Switch Fabric
  - ACL
  - QoS
Switch Fabric
Load-Sharing

- Ingress fabric interface asic knows all active paths through 3-stage xbar to each destination.
- Unicast traffic is 'sprayed' (2.5kB superframe broken to small chunks) across all active paths to egress fabric interface asic.
- Multicast traffic selects one of the active paths to egress fabric interface asic based on hash result calculated based on L2/L3/L4 information (same as EC hash but not configurable).
- First and next fragments may take different path due to missing L4 information in next fragments.
### Switch Fabric Utilization, Capacity, VoQ

#### N7K-1-VDC3# show hardware fabric-utilization

<table>
<thead>
<tr>
<th>Slot</th>
<th>Total Fabric Bandwidth</th>
<th>Utilization Ingress %</th>
<th>Egress %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>138 Gbps</td>
<td>4.0</td>
<td>2.0</td>
</tr>
<tr>
<td>2</td>
<td>138 Gbps</td>
<td>0.0</td>
<td>1.0</td>
</tr>
<tr>
<td>4</td>
<td>138 Gbps</td>
<td>0.0</td>
<td>1.0</td>
</tr>
<tr>
<td>5</td>
<td>69 Gbps</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>6</td>
<td>69 Gbps</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>7</td>
<td>138 Gbps</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>10</td>
<td>138 Gbps</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

#### N7K-3# show hardware fabric-utilization

<table>
<thead>
<tr>
<th>Slot</th>
<th>Total Fabric Bandwidth</th>
<th>Utilization Ingress %</th>
<th>Egress %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>230 Gbps</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2</td>
<td>230 Gbps</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>3</td>
<td>230 Gbps</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>9</td>
<td>115 Gbps</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>10</td>
<td>115 Gbps</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>15</td>
<td>230 Gbps</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>17</td>
<td>230 Gbps</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>18</td>
<td>230 Gbps</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

**4 Virtual Output Queues (VOQ) to every egress port ASIC (every 12 x 1GE ports or 4 x 10GE ports in shared mode or 1 x 10GE port in dedicated mode or 2 x 1/10GE ports)**

**Unicast** traffic access to fabric is arbitrated (arbiter on active supervisor provides grant when there is enough bandwidth available to destination VoQ)

**Multicast** traffic access to fabric is non-arbitrated

- 10 slot chassis with 3 xbar modules
- % utilization calculated based on 138Gbps (6 x 23G) overall slot capacity
- 18 slot chassis with 5 xbar modules (10 x 23G = 230G) capacity per slot

**Same RBH calculation algorithm as for EC**

#### N7K-1-VDC3# show hardware forwarding multicast fabric-path ingress e1/1 src-ip 172.23.25.64 dst-ip 239.28.28.64 src-port 1964 dst-port 1967 src-mac 001d.4632.3c00 dst-mac 0100.5e1c.1c40

Missing params will be substituted by 0's.

Module 1: **RBH: 0x7**

Xbar link instance: **2**, Fabric Slot: **1** Port: **3**
### Switch Fabric

Unicast Traffic Across Fabric, Utilization Details

```
N7K-1-VDC3# show hardware fabric-utilization detail | egrep -v "A --|B --|Fabric Planes"
```

<table>
<thead>
<tr>
<th>I/O Slot</th>
<th>Fab Mod</th>
<th>Fab Ins</th>
<th>Fab Chnl</th>
<th>Fab Link</th>
<th>Ingress%</th>
<th>Egress%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>12</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>12</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td>12</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td>12</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td>12</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>1</td>
<td>12</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>1</td>
<td>12</td>
<td>5</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

**Diagram:**

- **E1/1**
- **E1/2**
- **E1/25-26**
- **E2/25-26**
- **PeerA**

**Legend:**

- **I/O Slot** – I/O Module Number
- **Fab Mod** – Xbar Module Number
- **Fab Ins** – Fabric Asic Instance (10 slot chassis xbar has one, 18 slot chassis xbar has 2)
- **Fab Chnl** – Physical Fabric Channel (Port) Number
- **Fab Link** – I/O Module Stage 1 or 3 to Xbar Stage 2 logical Link number
- **Fab Plane** – Logical Fabric Data Plane
- **Plane A** – Unicast Data Plane
- **Plane B** – Multicast Data Plane
Switch Fabric
Multicast Traffic Across Fabric, Utilization Details

N7K-1-PeerA# show hardware fabric-utilization detail

Fabric Planes:
A -- Unicast fabric interface
B -- Multicast/Multidestination fabric interface

Unidirectional Fabric Bandwidth per Fab Link is 23 Ggps (A+B)

<table>
<thead>
<tr>
<th>I/O Slot</th>
<th>Fab Mod</th>
<th>Ins Chnl Link Plane</th>
<th>Fabric Utilization</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1 5 0 A</td>
<td>Ingress% Egress%</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1 5 0 B</td>
<td>0 0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1 3 1 A</td>
<td>0 0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1 3 1 B</td>
<td>42 21</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1 5 2 A</td>
<td>0 0</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1 5 2 B</td>
<td>0 0</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1 3 3 A</td>
<td>0 0</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1 3 3 B</td>
<td>0 0</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>1 5 4 A</td>
<td>0 0</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>1 5 4 B</td>
<td>0 0</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>1 3 5 A</td>
<td>0 0</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>1 3 5 B</td>
<td>0 0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1 4 0 A</td>
<td>0 0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1 4 0 B</td>
<td>0 0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1 12 1 A</td>
<td>0 0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1 12 1 B</td>
<td>0 21</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1 4 2 A</td>
<td>0 0</td>
</tr>
</tbody>
</table>

Use keyword ‘timestamp’ to see maximum fabric channel utilization time stamp

Fabric utilization displayed is always from xbar perspective for given module

Multicast fabric path calculated RBH hash selecting Fabric module 1, Fabric channel 3 and Fabric link 2 (this output is zero based)
Switch Fabric

Unicast and Multicast Traffic Across Fabric, VoQ-to-Xbar Link Mapping Details

N7K-1-PeerA# show hardware fabric-utilization detail module 2 | grep -a 21 STAGE-1

<table>
<thead>
<tr>
<th>I/O Mod-Fab</th>
<th>Mod-Fab Fabric Utilization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Link Instance Channel ID</td>
<td>Plane</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
</tr>
</tbody>
</table>

module-2# show hardware internal xbar-driver local inst 1 driver_info | grep -a 15 Port-Enabled

<table>
<thead>
<tr>
<th>Port-Enabled</th>
<th>Connected-To</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>Octo 1 Port 7</td>
</tr>
<tr>
<td>01</td>
<td>Fabric 2 Link 1</td>
</tr>
<tr>
<td>03</td>
<td>Octo 0 Port 1</td>
</tr>
<tr>
<td>04</td>
<td>Octo 0 Port 0</td>
</tr>
<tr>
<td>05</td>
<td>Octo 0 Port 3</td>
</tr>
<tr>
<td>07</td>
<td>Fabric 3 Link 1</td>
</tr>
<tr>
<td>08</td>
<td>Octo 1 Port 6</td>
</tr>
<tr>
<td>09</td>
<td>Octo 0 Port 2</td>
</tr>
<tr>
<td>10</td>
<td>Fabric 1 Link 0</td>
</tr>
<tr>
<td>11</td>
<td>Fabric 3 Link 0</td>
</tr>
<tr>
<td>15</td>
<td>Fabric 2 Link 0</td>
</tr>
<tr>
<td>16</td>
<td>Fabric 1 Link 1</td>
</tr>
<tr>
<td>17</td>
<td>Octo 1 Port 5</td>
</tr>
<tr>
<td>18</td>
<td>Octo 1 Port 4</td>
</tr>
</tbody>
</table>

I/O Link – Logical Link Instance between I/O Module Fabric Asic and VoQ asic
Mod-Fab Instance – I/O Module Fabric Asic Instance
Mod-Fab Channel ID – I/O Module Fabric Asic physical channel (port) ID

Connection to 3 Xbars modules (Stage 2, 2 links per Xbar module)

Unicast and Multicast traffic flowing across switch fabric
Switch Fabric
Data Collection (No show tech-support xbar Exists)

show hardware internal xbar-driver event-history errors
show hardware internal xbar-driver event-history msgs
show logging onboard internal xbar
show event-history xbar
show hardware internal qengine asic 0|1 ➔ Collect for VoQ to Fabric interaction issues

N7K-1-PeerA# show system internal xbar ?
  all                  Show xbar all data
  dyn-mcast-info      Show xbar dynamic multicast info
  dynamic-mc          Show xbar dynamic multicast table
  event-history       Show internal event history
  flood-mc            Show xbar flood multicast table
  get-mi-slotmask     Enter the slotmask
  mc                   Show xbar sw multicast table
  mem-stats           Show xbar allocation statistics
  static-mc           Show xbar static multicast table
  sw                   Show xbar sw data
  sync-loss-threshold Enable setting sync-loss handling params
  vqi-info            Show internal vqi-info
  vqi-map             Show vqi-map information

For any packet loss related issues use first show hardware internal error module X and when you see any potentially related counters moving use show hardware internal statistics module X device <device> <category> asic-all to filter out unnecessary output (cli may produce very long output difficult to read)
Agenda

- **Before Troubleshooting**
  - Brief Nexus 7000 Module and Forwarding Engine Architecture Overview
  - Build in Troubleshooting Tools
  - System Access, File System Management

- **Troubleshooting**
  - CPU, Control-Plane and Memory Utilization
  - vPC
  - Unicast L2 and L3 Forwarding and ARP
  - Multicast Layer 2 and Layer 3 Forwarding
  - Switch Fabric
  - ACL
  - QoS
ACL — Operation

- Atomic/hitless update of existing applied ACL while modified
  - temporary label swap (no use of default-result)
  - two acl copies in tcam, if there is no enough space, process fails
- ACL tcam banks chaining supported
- L4OPs/LOUs only used for expansion beyond 5 lines, configurable

```
N7K-1-VDC3# show system internal access-list globals module 1
Atomic Update : ENABLED
Default ACL : PERMIT
Bank Chaining : DISABLED
LOU Threshold Value : 5
```

```
N7K-1(config)# hardware access-list resource ?
  pooling Enable ACL programming across TCAM banks
```

```
N7K-1(config)# hardware access-list update ?
  atomic Enable atomic update of access-list in hardware
  default-result Default access-list result during non-atomic hardware update
```

```
N7K-1(config)# hardware access-list lou resource threshold 10
```

Note: All below cli is available in default vDC only as it applies to system wide resources

TCAM chaining (2x32K TCAMs, 2 banks each)
Disable atomic update if there is not enough space in TCAM
Hidden cli available in 5.1.X code

NOTE: Operation in progress, please check the status using 'show hardware access-list lou resource threshold' command
ACL — Operation

- Use ‘configure session’ to create and modify large ACLs (dry run)
  - 32 sessions, 100k lines
- High CPU usage during large ACL processing is expected
- CPU/RP (netstack process) has its own database of egress ACLs for software switched traffic cases

```
N7K-3-VDC3(config)# interface ethernet1/9
N7K-3-VDC3(config-if)# ip access-group 20k_ace_out out

N7K-3-PeerB# show system internal processes cpu
top - 17:26:02 up 2 days, 1:47, 3 users, load average: 0.76, 0.52, 0.47
Tasks: 419 total, 1 running, 418 sleeping, 0 stopped, 0 zombie
Cpu(s): 4.3%us, 5.6%sy, 0.1%ni, 85.9%id, 0.6%wa, 0.1%hi, 3.4%si, 0.0%st
Mem: 4115232k total, 3496772k used, 618460k free, 52344k buffers
Swap: 0k total, 0k used, 0k free, 1653252k cached

PID USER PR NI VIRT RES SHR S %CPU %MEM TIME+ COMMAND
4447 root 20 0 186m 61m 18m S 97.0 1.5 96:39.99 netstack

module-3# show processes cpu sort
PID Runtime(ms) Invoked uSecs 1Sec Process
----- --------- ------ ------ ------- ---------
1940 442620 894399 494 93.1 aclqos
```

Netstack processing egress access-list
Ingress ACLs are programmed only to required I/O modules (localization support)

Egress access-lists are programmed to all I/O modules as they are executed on ingress

ACL statistics in software must be enabled via configuration

```
N7K-1-VDC3# show run interface ethernet 1/1 | i access
    ip access-group tcp_flags in
    ip access-group test_punt out

N7K-1-VDC3# show hardware access-list interface ethernet 1/1 input config module 1

    Policy id:  1, Type: QoS, Protocol: IPv4   Name: *
    Policy id:  3, Type: RACL, Protocol: IPv4   Name: tcp_flags

permit tcp 0.0.0.0/0.0.0.0 0.0.0.0/0.0.0.0 syn log
permit tcp 0.0.0.0/0.0.0.0 0.0.0.0/0.0.0.0 ack log
permit ip 0.0.0.0/0.0.0.0 0.0.0.0/0.0.0.0
deny ip 0.0.0.0/0.0.0.0 0.0.0.0/0.0.0.0 *

N7K-1-VDC3# show hardware access-list interface ethernet 1/1 input config module 2
no policy found

ACL TCAM on I/O Module 2 does not contain access-lists configured on I/O Module 1 interface
```
### ACL — Hardware Programming

#### egress ACL hw configuration

N7K-1-VDC3# `show hardware access-list interface ethernet 1/1 output config module 1`

- **Policy id:** 2, **Type:** QoS, **Protocol:** IPv4, **Name:** *
- **Policy id:** 5, **Type:** RACL, **Protocol:** IPv4, **Name:** test_punt

```
permit udp 172.222.222.64/255.255.255.255 172.31.31.250/255.255.255.255 log
permit icmp 9.9.9.9/255.255.255.255 172.31.31.250/255.255.255.255 log
permit ip 0.0.0.0/0.0.0.0 0.0.0.0/0.0.0.0
deny ip 0.0.0.0/0.0.0.0 0.0.0.0/0.0.0.0 *
```

N7K-1-VDC3# `show hardware access-list interface ethernet 1/1 output config module 2`

- **Policy id:** 4, **Type:** RACL, **Protocol:** IPv4, **Name:** test_punt

```
permit udp 172.222.222.64/255.255.255.255 172.31.31.250/255.255.255.255 log
permit icmp 9.9.9.9/255.255.255.255 172.31.31.250/255.255.255.255 log
permit ip 0.0.0.0/0.0.0.0 0.0.0.0/0.0.0.0
deny ip 0.0.0.0/0.0.0.0 0.0.0.0/0.0.0.0 *
```

- Specific applications (dhcp, bfd) may install their own ACLs which must merge with user configured racl,vacl,pacl
- Some combination of ACL based applications may not be supported
- Data collection: `show tech-support aclmgr detail`
N7K-1-VDC2# show hardware access-list vlan 33 input statistics module 1
Tcam 1 resource usage:
-----------------------
Label_b = 0x3
  Bank 0
------
IPv4 Class
  Policies: [DHCP Snooping() BFD() [Merged]]
Entries:
  [Index] Entry [Stats]
  ---------------------
[0014] redirect(0x43024) udp 0.0.0.0/0 0.0.0.0/0 eq 3785 ttl eq 254 [185050]
[0015] redirect(0x43024) udp 0.0.0.0/0 0.0.0.0/0 eq 3784 ttl eq 255 [5783]
[0016] redirect(0x800) udp 0.0.0.0/0 255.255.255.255/32 eq 68 [0]
[0017] redirect(0x800) udp 0.0.0.0/0 255.255.255.255/32 eq 67 [0]
[0018] redirect(0x800) udp 0.0.0.0/0 eq 68 255.255.255.255/32 [0]
[0019] redirect(0x800) udp 0.0.0.0/0 eq 67 255.255.255.255/32 [0]
[0020] permit ip 0.0.0.0/0 0.0.0.0/0 [240021]

N7K-1-VDC2-CS1# show hardware access-list vl 33 input l4ops module 1
Tcam 1 resource usage:
-----------------------
Lou usage:
  [Lou] sw_id l4op_bit ref_count Operation
  ---------------------
  2(A) 0 0 1 IPTTL EQ(255)
  2(B) 1 1 1 IPTTL EQ(254)

TCP flags usage: none
ACL — Hardware Resource Usage
Feature ACLs and RACL Merge

N7K-1-PeerA# show hardware access-list vlan 33 input statistics module 1
Tcam 1 resource usage:
----------------------
Label b = 0x8

Bank 0

 refined

IPv4 Class
Policies: RACL(test_lou) DHCP Snooping() BFD() [Merged],
Entries:

[0013] permit tcp 1.1.1.0/24 2.2.2.0/24 fragment [0]
[0014] permit tcp 1.1.1.0/24 2.2.2.0/24 eq 179 [0]
[0015] permit tcp 1.1.1.0/24 eq 179 2.2.2.0/24 [0]
[0016] deny routed udp 0.0.0.0/0 0.0.0.0/0 range 2000 2300 [0]
[0017] deny routed tcp 10.0.0.0/8 20.0.0.0/24 range 1500 1900 [0]
[0054] redirect(0x43035) udp 0.0.0.0/0 0.0.0.0/0 eq 3785 ttl eq 254 [152]
[0055] redirect(0x43035) udp 0.0.0.0/0 0.0.0.0/0 eq 3784 ttl eq 255 [3]
[0056] redirect(0x800) udp 0.0.0.0/0 255.255.255.255/32 eq 68 [0]
[0057] redirect(0x800) udp 0.0.0.0/0 255.255.255.255/32 eq 67 [0]
[0058] redirect(0x800) udp 0.0.0.0/0 eq 68 255.255.255.255/32 [0]
[0059] redirect(0x800) udp 0.0.0.0/0 eq 67 255.255.255.255/32 [0]
[0060] permit ip 0.0.0.0/0 0.0.0.0/0 [124]

N7K-1-PeerA# show hardware access-list vlan 33 output merge module 1
Tcam 1 resource usage:
----------------------
Lou usage:

<table>
<thead>
<tr>
<th>sw_id</th>
<th>l4op_bit</th>
<th>ref_count</th>
<th>Operation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1(AB)</td>
<td>4</td>
<td>0</td>
<td>dest-port: RANGE(2000, 2300)</td>
</tr>
<tr>
<td>4(AB)</td>
<td>5</td>
<td>1</td>
<td>dest-port: RANGE(1500, 1900)</td>
</tr>
<tr>
<td>0(A)</td>
<td>0</td>
<td>2</td>
<td>IPTTL EQ(255)</td>
</tr>
<tr>
<td>0(B)</td>
<td>1</td>
<td>3</td>
<td>IPTTL EQ(254)</td>
</tr>
</tbody>
</table>

TCP flags usage: none

ACEs would be expanded to more than 5 lines and therefore LOU were used instead
### ACL — Hardware Resource Usage

Per I/O Module Summary, VDC Wide ACL Summary

#### N7K-1-PeerA# show hardware access-list resource utilization module 1

**ACL Hardware Resource Utilization (Module 1)**

<table>
<thead>
<tr>
<th>Used</th>
<th>Free</th>
<th>Percent Utilization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tcam 0, Bank 0</td>
<td>5</td>
<td>16379</td>
</tr>
<tr>
<td>Tcam 0, Bank 1</td>
<td>3</td>
<td>16381</td>
</tr>
<tr>
<td>Tcam 1, Bank 0</td>
<td>55</td>
<td>16329</td>
</tr>
<tr>
<td>Tcam 1, Bank 1</td>
<td>151</td>
<td>16233</td>
</tr>
<tr>
<td>LOU</td>
<td>5</td>
<td>99</td>
</tr>
<tr>
<td>Both LOU Operands</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Single LOU Operands</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>LOU L4 src port:</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>LOU L4 dst port:</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>LOU L3 packet len:</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>LOU IP tos:</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>LOU IP dscp:</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>LOU ip precedence:</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>LOU ip TTL:</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>TCP Flags</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>Protocol CAM</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Mac Etype/Proto CAM</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Non L4op labels, Tcam 0</td>
<td>3</td>
<td>6140</td>
</tr>
<tr>
<td>Non L4op labels, Tcam 1</td>
<td>4</td>
<td>6139</td>
</tr>
<tr>
<td>L4 op labels, Tcam 0</td>
<td>0</td>
<td>2047</td>
</tr>
<tr>
<td>L4 op labels, Tcam 1</td>
<td>7</td>
<td>2040</td>
</tr>
<tr>
<td>Ingress Dest info table</td>
<td>2</td>
<td>510</td>
</tr>
<tr>
<td>Egress Dest info table</td>
<td>1</td>
<td>511</td>
</tr>
</tbody>
</table>

Cumulative usage of I/O Module 1 ACL TCAM hardware resources by all type of programmed access-lists

#### N7K-1-PeerA# show access-lists summary | egrep -a 5 tcp|lou

**IPV4 ACL tcp_zoom**

- Total ACEs Configured: 5
  - Configured on interfaces:
    - port-channel111 - ingress (Router ACL)
  - Active on interfaces:
    - port-channel111 - ingress (Router ACL)

**IPV4 ACL test_lou**

- Total ACEs Configured: 5
  - Configured on interfaces:
    - Vlan33 - ingress (Router ACL)
  - Active on interfaces:
    - Vlan33 - ingress (Router ACL)
Agenda

- Before Troubleshooting
  - Brief Nexus 7000 Module and Forwarding Engine Architecture Overview
  - Build in Troubleshooting Tools
  - System Acess, File System Management

- Troubleshooting
  - CPU, Control-Plane and Memory Utilization
  - vPC
  - Unicast L2 and L3 Forwarding and ARP
  - Multicast Layer 2 and Layer 3 Forwarding
  - Switch Fabric
  - ACL
  - QoS
QoS — What Is It?

- QoS provides preferential service for particular end-to-end application flows, it is enabled by default and cannot be disabled.

- General QoS Operational Model
  - Inbound Packets Scheduling - ingress queuing
  - Inbound Packet Classification - assigning CoS/ToS in 802.1p or IP header
  - Ingress policing
  - Outbound Packets Classification
  - Outbound Packet Scheduling - egress queuing
  - Egress policing, shaping

- Default Queuing (bandwidth, queue-limit) and QoS (policing) policies applied to all physical and port-channel interfaces across all vDCs

- Default interface behavior is trust and can only be overritten by applying both Queuing and QoS policies

- QoS comes to play only during over-subscription condition caused by excess of traffic or flow-control

- VoQ may be considered part of NEXUS 7000 QoS functionality
QoS — Where Is It?
Why Does Packet Loss Happen?

What is the burst size port egress buffer can sustain?
Avps = Average packet size = ~1500B
OR = Outgoing Rate = ~192Mbps
BU = Buffer size = 6.2MB
BT = Burst duration Time
BS = Burst Size

BT = BU/OR

BT = (6.2 \times 10^6 \times 8b) / 192 \times 10^6 b/s = \approx 258ms

BS = BT/(Avps*8)

BS = (258ms \times 1^9)/(1500*8) = 21.5K packets

@1500B
**QoS**

CoS-to-Queue Mapping in Hardware

```bash
N7K-1-VDC3# show system internal qos queuing config interface ethernet1/1 | b "from HW" | grep -a 9 "COS2Q Config"
COS2Q Config
Direction: egress
COS 0 Queue Number: 0
COS 1 Queue Number: 0
COS 2 Queue Number: 0
COS 3 Queue Number: 0
COS 4 Queue Number: 0
COS 5 Queue Number: 7
COS 6 Queue Number: 7
COS 7 Queue Number: 7
--
```

```bash
N7K-1-VDC3# show class-map type queuing 8q2t-in-q-default
Type queuing class-maps
========================
class-map type queuing match-any 8q2t-in-q-default
Description: Classifier for egress default queue of type 8q2t
 match cos 0-4
```

```bash
N7K-1(config-cmap-que)# match cos 5-7
```

```bash
N7K-1-VDC3# show class-map type queuing 8q2t-in-q-default
Type queuing class-maps
========================
class-map type queuing match-any 2q4t-in-q-default
Description: Classifier for ingress default queue of type 2q4t
 match cos 0-7
```
QoS
Untrusted Port Configuration (Queuing Class Changes Are Global)

N7K-1-VDC3# show ip access-lists reset-dscp-any | grep -v ^$
IP access list reset-dscp-any
  10 permit ip any any

N7K-1-VDC3# show class-map type qos reset-dscp-to-0 | grep -v ^$
Type qos class-maps
  ==============
    class-map type qos match-all reset-dscp-to-0
    match access-group name reset-dscp-any

N7K-1-VDC3# show policy-map type queuing reset-cos-to-0 | grep -v ^$
Type queuing policy-maps
  ==============
    policy-map type queuing reset-cos-to-0
      class type queuing 8q2t-in-q-default
        set cos 0
        bandwidth percent 100

N7K-1-VDC3# show policy-map interface e1/1 input type qos | grep -v ^$
Global statistics status: enabled
Ethernet1/1
  Service-policy (qos) input: reset-dscp-to-0
    policy statistics status: enabled
    Class-map (qos): reset-dscp-to-0 (match-all)
      2218264 packets
      Match: access-group reset-dscp-any
      set dscp 0

N7K-1-VDC3# show policy-map interface e1/1 input type queuing | grep -v ^$
Global statistics status: enabled
Ethernet1/1
  Service-policy (queuing) input: reset-cos-to-0
    policy statistics status: enabled
    Class-map (queuing): 8q2t-in-q-default (match-any)
      set cos 0
      bandwidth percent 100
      queue dropped pkts : 0

---

module-1(lamira-elm)#echo dbus | egrep _sa|_da|cos|tos
acos = dscp

This is egress result after rewrite

module-1(lamira-elm)# show dbus | egrep _sa|_da|cos|tos
ofe_acos = 0x30
acos = dscp
ac0 = 0x30
cos = 0x6
tos_offset = 0x0f
tos = 0xc0

No policies applied cos=6 dscp=48

Policies applied cos=0 and dscp=0
QoS
Untrusted Port Configuration

N7K-1-VDC3# show run int e1/1 | i service | grep -v ^$
service-policy type qos input reset-dscp-to-0
service-policy type queuing input reset-cos-to-0
N7K-1-VDC3# show queuing interface e1/1 summary
Egress Queuing for Ethernet1/1 [Interface]
-----------------------------------
Template: 1P7Q4T
Que# Group Bandwidth% PrioLevel Shape% CoSMap
-----------------------------------
0 - 25 - - 0-4
1 - 15 - - -
2 - 12 - - -
3 - 12 - - -
4 - 12 - - -
5 - 12 - - -
6 - 12 - - -
7 - - High - 5-7

Ingress Queuing for Ethernet1/1 [Interface]
-----------------------------------
Template: 8Q2T
Trust: Untrusted [Default CoS 0]
Que# Group Qlimit% IVL CoSMap
-----------------------------------
0 - 100 - 0-7

[snip]

N7K-1-VDC3# show run int e1/1 | i service
N7K-1-VDC3# show queuing interface e1/1 summary
Egress Queuing for Ethernet1/1 [Interface]
-----------------------------------
Template: 1P7Q4T
Que# Group Bandwidth% PrioLevel Shape% CoSMap
-----------------------------------
0 - 25 - - 0-4
1 - 15 - - -
2 - 12 - - -
3 - 12 - - -
4 - 12 - - -
5 - 12 - - -
6 - 12 - - -
7 - - High - 5-7

Ingress Queuing for Ethernet1/1 [Interface]
-----------------------------------
Template: 8Q2T
Trust: Trusted

Que# Group Qlimit% IVL CoSMap
-----------------------------------
0 - 100 - 0-7

[snip]

Be very careful when remapping CoS into different tx queues as queuing policies apply to all physical ports and queues may not have enough buffer space allocated by default!

Policies not applied and port is trusted

Policies applied and port is untrusted

© 2011 Cisco and/or its affiliates. All rights reserved.
### Ingress Port Type Default QoS Functionality

<table>
<thead>
<tr>
<th>L3-&gt;L3</th>
<th>L3-&gt;L3-SVI</th>
<th>L2-trunk-&gt;L2-trunk</th>
<th>L2-access-&gt;L2-access</th>
</tr>
</thead>
<tbody>
<tr>
<td>module-1(lamira-elam)# show dbus</td>
<td>module-1(lamira-elam)# show dbus</td>
<td>module-1(lamira-elam)# show dbus</td>
<td>module-2(lamira-elam)# show dbus</td>
</tr>
<tr>
<td>_sa</td>
<td>_da</td>
<td>_smac</td>
<td>_dmac</td>
</tr>
<tr>
<td>cos_1q = 0x0</td>
<td>cos_1q = 0x0</td>
<td>cos_1q = 0x0</td>
<td>cos_1q = 0x0</td>
</tr>
<tr>
<td>ipv4_tos = 0xc0</td>
<td>ipv4_tos = 0xc0</td>
<td>ipv4_tos = 0xc0</td>
<td>ipv4_tos = 0xa0</td>
</tr>
<tr>
<td>ipv4_sa = 172.222.222.064</td>
<td>ipv4_sa = 172.222.222.064</td>
<td>ipv4_sa = 172.222.222.064</td>
<td>ipv4_sa = 192.251.065.151</td>
</tr>
<tr>
<td>ipv4_da = 172.032.032.250</td>
<td>ipv4_da = 172.032.032.250</td>
<td>ipv4_da = 172.032.032.250</td>
<td>ipv4_da = 192.251.065.161</td>
</tr>
<tr>
<td>ipv4_dmac = 00.23.ac.64.46.c3</td>
<td>ipv4_dmac = 00.23.ac.64.46.c2</td>
<td>ipv4_dmac = 00.23.ac.64.46.c2</td>
<td>ipv4_dmac = 00.00.16.01.16.01</td>
</tr>
<tr>
<td>ipv4_smac = 00.1d.46.32.3c.00</td>
<td>ipv4_smac = 00.23.ac.64.46.c3</td>
<td>ipv4_smac = 00.23.ac.64.46.c3</td>
<td>ipv4_smac = 00.00.15.01.15.01</td>
</tr>
<tr>
<td>ipv4_ce_cos = 0x0</td>
<td>ofe_acos = 0x30</td>
<td>ofe_acos = 0x30</td>
<td>ofe_acos = 0x28</td>
</tr>
<tr>
<td>acos = 0x30</td>
<td>acos = 0x30</td>
<td>acos = 0x30</td>
<td>acos = 0x28</td>
</tr>
<tr>
<td>cos = 0x6</td>
<td>cos = 0x6</td>
<td>cos = 0x6</td>
<td>cos = 0x0</td>
</tr>
<tr>
<td>vlan = 0</td>
<td>vlan = 32</td>
<td>vlan = 32</td>
<td>vlan = 65</td>
</tr>
<tr>
<td>tos_offset = 0x0f</td>
<td>tos_offset = 0x0f</td>
<td>tos_offset = 0x0f</td>
<td>tos_offset = 0x0f</td>
</tr>
<tr>
<td>tos = 0xc0</td>
<td>tos = 0xc0</td>
<td>tos = 0xc0</td>
<td>tos = 0xa0</td>
</tr>
</tbody>
</table>
QoS
I/O Module Resources Utilization and Internal Mapping Example

N7K-1-VDC3# show system internal qos resource utilization module 4

<table>
<thead>
<tr>
<th>Resource</th>
<th>Module</th>
<th>Total</th>
<th>Used</th>
<th>Free</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregate policers:</td>
<td>4</td>
<td>12288</td>
<td>12</td>
<td>12276</td>
</tr>
<tr>
<td>Distributed policers:</td>
<td>4</td>
<td>4096</td>
<td>0</td>
<td>4096</td>
</tr>
<tr>
<td>Policer Profiles:</td>
<td>4</td>
<td>1024</td>
<td>12</td>
<td>1012</td>
</tr>
</tbody>
</table>

N7K-1-VDC3# show system internal ipqos global-defaults | grep -a 12 dscp-cos-map
table-map: dscp-cos-map (len: 12)
default copy
Bit array:
Values set:

DSCP values 16 – 23

0 0 0 0 0 0 0 0 0
1 1 1 1 1 1 1 1 1
2 2 2 2 2 2 2 2 2
3 3 3 3 3 3 3 3 3
4 4 4 4 4 4 4 4 4
5 5 5 5 5 5 5 5 5
6 6 6 6 6 6 6 6 6
7 7 7 7 7 7 7 7 7

N7K-1-VDC3# show system internal ipqos global-defaults | grep -a 12 precedence-dscp-map
table-map: precedence-dscp-map (len: 19)
default copy
Bit array:
Values set:

DSCP values 16 – 23 are mapped to CoS values 2

IPP value 3 is mapped to DSCP value 24-31
Conclusion — You Can Now Troubleshoot!

- Things to remember …
  - Data required to identify your issue is very likely to be in one of the device internal logs … just find it and collect it
  - 90% of issues can be resolved or at least identified analyzing data present in already existing logs
  - Permanent debugs output may fill up internal logs quickly so adjust the size of feature log files as you see fit
  - Use built-in tools to make your troubleshooting effective, be familiar with what show tech-supports are available and what component do they cover
  - Evidence collected using internal tools is most relevant
Complete Your Online Session Evaluation

- Receive 25 Cisco Preferred Access points for each session evaluation you complete.

- Give us your feedback and you could win fabulous prizes. Points are calculated on a daily basis. Winners will be notified by email after July 22nd.

- Complete your session evaluation online now (open a browser through our wireless network to access our portal) or visit one of the Internet stations throughout the Convention Center.

- Don’t forget to activate your Cisco Live and Networkers Virtual account for access to all session materials, communities, and on-demand and live activities throughout the year. Activate your account at any internet station or visit www.ciscolivevirtual.com.
Visit the Cisco Store for Related Titles
http://theciscostores.com
Thank you.