CISCO Live!

ALL IN

#CiscoLive

Troubleshooting the Cisco Catalyst 9000 Series Switches

Michel Peters
Technical Leader Engineering
BRKTRS-3090

#CiscoLive

Cisco Webex App

Questions?
Use Cisco Webex App to chat
with the speaker after the session

How

1. Find this session in the Cisco Live Mobile App
2. Click "Join the Discussion"
3. Install the Webex App or go directly to the Webex space
4. Enter messages/questions in the Webex space

Webex spaces will be moderated
by the speaker until June 17, 2022.

https://ciscolive.ciscoevents.com/ciscolivebot/#BRKTRS-3090

Agenda

- Architecture
- Tools
- Packet Drops
- Forwarding issues

Architecture

Catalyst 9k family

Cisco Catalyst
9200 Series

Cisco Catalyst
9300 Series

Cisco Catalyst
9400 Series

Cisco Catalyst
9500 Series

Cisco Catalyst
9600 Series

|	9200	9300	9400	9500	9600
UADP	2.0 mini	2.0	2.0 XL	2.0 XL & 3.0	3.0
Format	Stackable	Stackale	Chassis	Standalone	Chassis

- UADP Asics are the foundation of the Catalyst 9K switches
- UADP 2.x evolution of UADP 1.0 used on Cat3k switches
- Port Asics (UADP) responsible for data plan forwarding
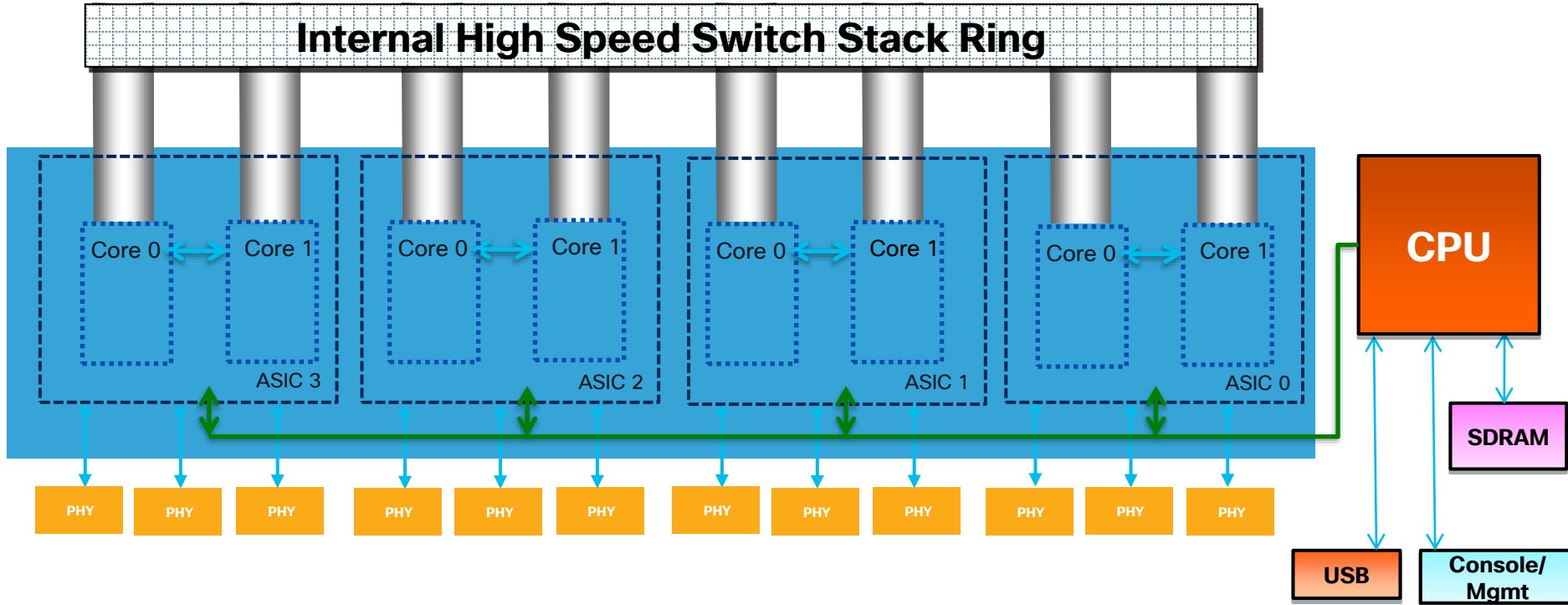- CPU's handle Control Plane

Catalyst 9200/9300 Stackable Switches
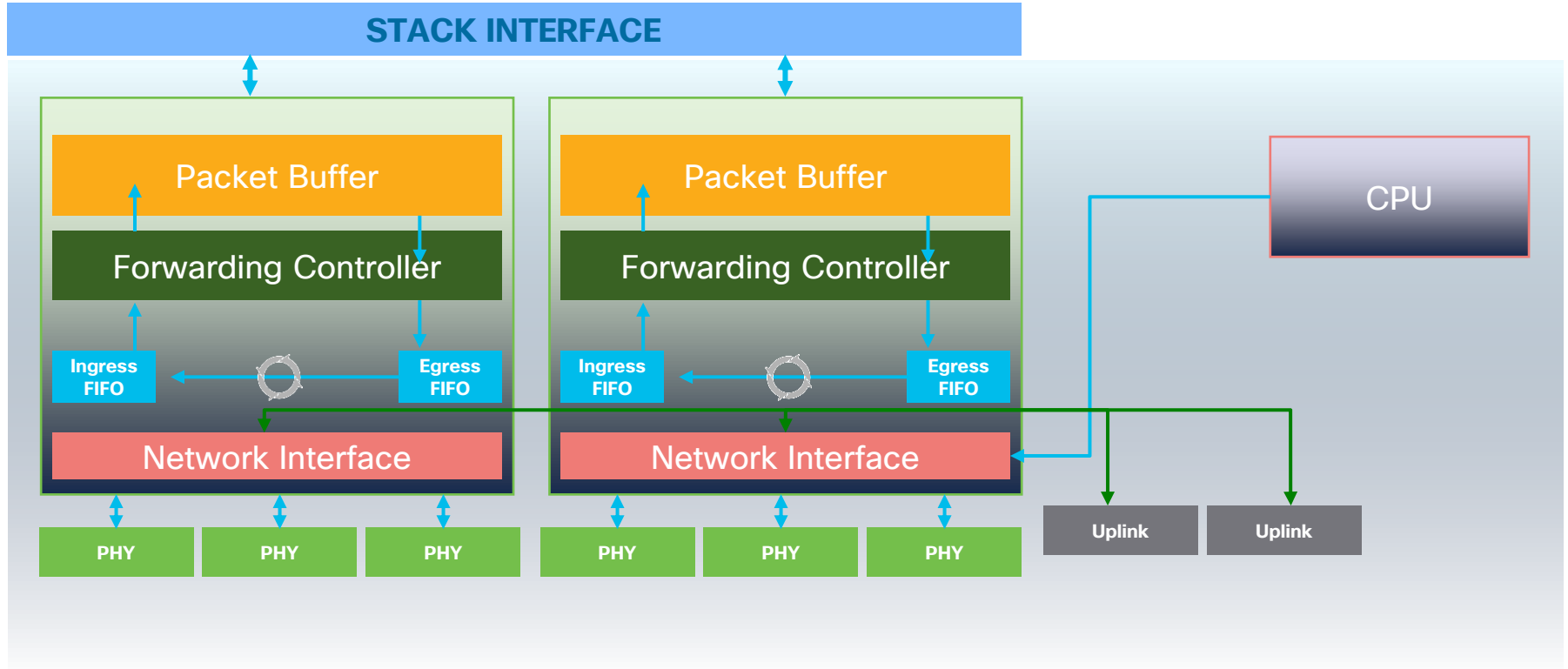
Catalyst 9400/9600 Chassis based
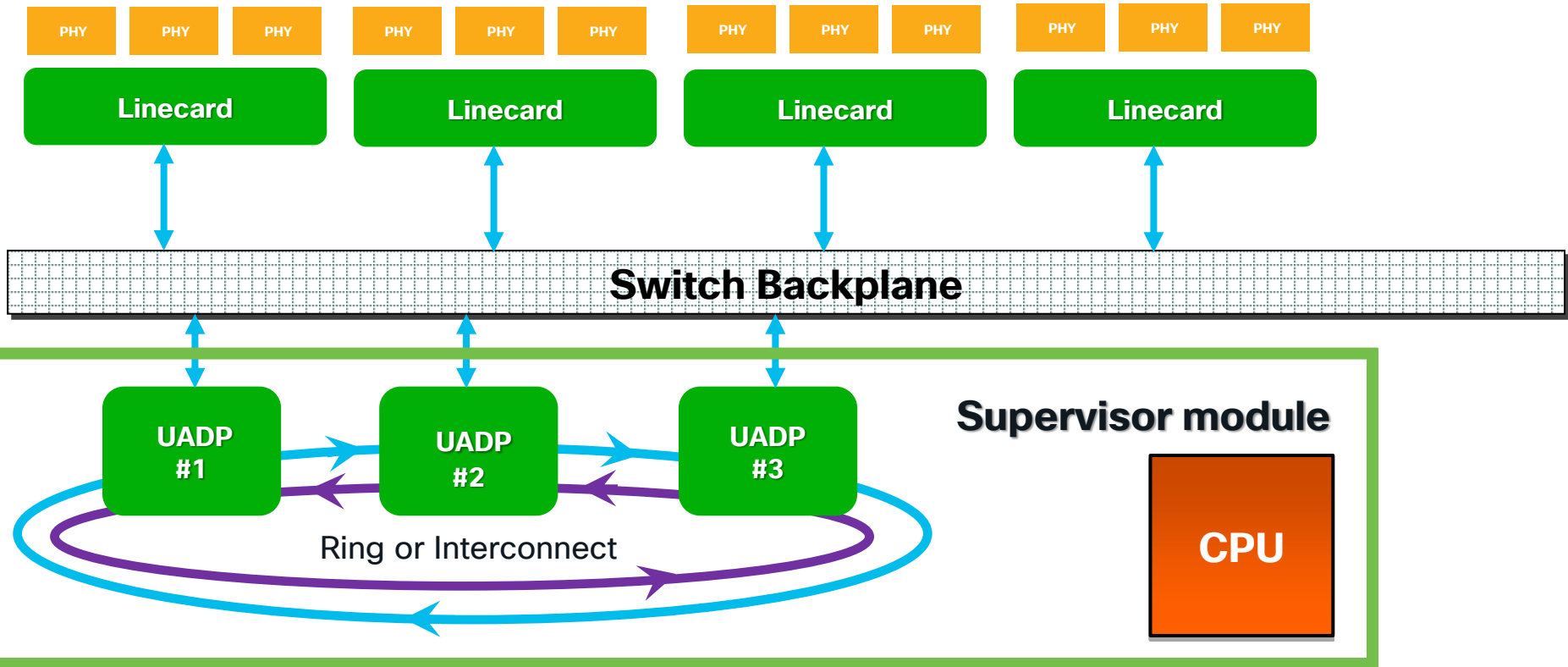
Catalyst 9500 Fixed Switches

Interface Internal Mappings

Interface to ASIC mapping important to understand data flows

```
9300_1#show platform software fed switch active ifm mappings
Interface              IF_ID     Inst Asic Core Port SubPort Mac   Cntx LPN   GPN   Type Active
GigabitEthernet1/0/1   0x8        1   0    1    0    0       26    6    1     1     NIF  Y
GigabitEthernet1/0/2   0x4c       1   0    1    1    0       6     7    2     2     NIF  Y
GigabitEthernet1/0/3   0x4d       1   0    1    2    0       28    8    3     3     NIF  Y
```

Internally used interface addressing:

- LPN   : Local Port Number
- GPN   : Global Port Number
- IF_ID   : Interface Identification, used for many fed CLI
- Type   : Type of interface, NIF = Network Interface
- Inst   : Instance : ASIC + Core
- Port   : Asic Ports
- Active   : Is Interface Active , multipurpose ports

IFM Mappings 9400/9500/9600

```
9500_1#sh platform software fed switch active ifm mappings | inc 1/./[1] |Int
Interface                  IF_ID    Inst Asic Core Port SubPort Mac   Cntx LPN  GPN   Type Active
TenGigabitEthernet1/0/1    0x2c     1    0    1    0    0       11    0    1    1     NIF  Y
TenGigabitEthernet1/1/1    0x18     0    0    0    16   0       0     19   17   17    NIF  N
FortyGigabitEthernet1/1/1  0x3c     0    0    0    24   0       4     4    25   25    NIF  Y


9600_1#sh platform software fed active ifm mappings | inc /0/[12][5] |Int
Interface                  IF_ID    Inst Asic Core Port SubPort Mac   Cntx LPN  GPN   Type Active
FortyGigabitEternet1/0/15  0x17     3    1    1    24   0       30    1    15   115   NIF  Y
HundredGigE1/0/25          0x21     0    0    0    0    0       0     0    25   125   NIF  Y
TwentyFiveGigE2/0/15       0x47     1    0    1    21   0       21    5    15   215   NIF  Y
TwentyFiveGigE2/0/25       0x51     3    1    1    7    0       7     7    25   225   NIF  Y


C9407R#sh platform software fed active ifm mappings | inc /0/1 |Int
Interface                  IF_ID    Inst Asic Core Port SubPort Mac   Cntx LPN  GPN   Type Active
GigabitEthernet2/0/1       0x7      1    0    1    0    0       7     7    1    201   NIF  Y
TenGigabitEthernet3/0/1    0x7f     4    2    0    19   0       19    3    1    301   NIF  Y
TenGigabitEthernet4/0/1    0x7e     5    2    1    3    0       11    3    1    401   NIF  Y
GigabitEthernet5/0/1       0x4b     3    1    1    0    0       4     4    1    501   NIF  Y
```
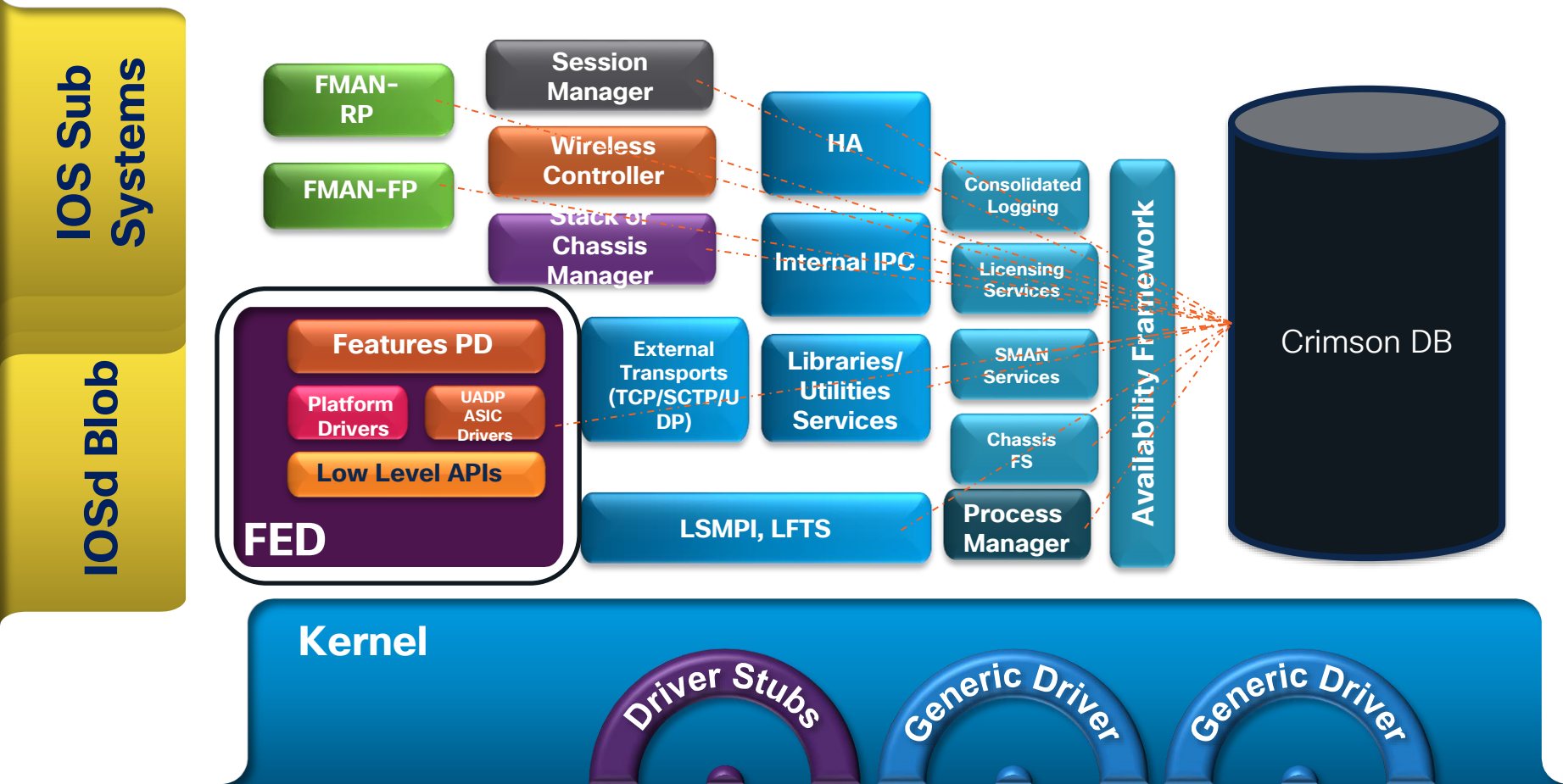
Different mappings on different models

IOS-XE 16 & IOS-XE 17 graphical overview

IOSd CPU utilization

```
9300_1#show processes cpu sorted
CPU utilization for five seconds: 1/0% one minute: 0%; five minutes: 0%
 PID Runtime(ms)       Invoked      uSecs    5Sec    1Min    5Min TTY Process
 434        8197       1230039          6   0.07%   0.01%   0.00%   0 MMON MENG
 203       10890        614953         17   0.07%   0.03%   0.01%   0 VRRS Main thread
 287           8            46        173   0.07%   0.00%   0.00%   0 Exec
 221       12377       1226864         10   0.07%   0.02%   0.00%   0 IP ARP Retry Age
 113       11806         20043        589   0.07%   0.03%   0.02%   0 Crimson flush tr
 218       12527       1226864         10   0.07%   0.02%   0.01%   0 IPAM Manager
 238       11425        393615         29   0.07%   0.02%   0.01%   0 UDLD
```

- Many process like still running as processes inside IOSd process (BGP, RIP, CEF, ARP, UDLD, CDP , etc)
- CPU utilization 1%/0% showing Utilization for Processes and Utilization for Interrupt
- IOSd does not provide multicore architecture

Kernel CPU information

```
9300_1#sh proc cpu platform sorted location switch active R0
CPU utilization for five seconds:  1%, one minute:  1%, five minutes:  1%
Core 0: CPU utilization for five seconds:  3%, one minute:  2%, five minutes:  2%
Core 7: CPU utilization for five seconds:  1%, one minute:  1%, five minutes:  0%
                   5Sec     1Min     5Min  Status          Size  Name
--------------------------------------------------------------------------
 14416   13034      6%       6%       6% S             223452  fed main event
 10014    9623      1%       1%       1% S              52212  sif_mgr
  9738    9215      1%       1%       1% S             818660  linux_iosd-imag
```

- Underlaying kernel shows Multi Core Architecture
- IOSd runs as process on kernel-> linux_iosd-image
- Not all processes run on all switches

```
Switch_1#sh processes cpu platform sorted location switch 1 R0 | in fman
 27777   26990      0%       0%       0% S         314179584  fman_rp
 19145   17642      0%       0%       0% S         296591360  fman_fp_image
Switch_1#sh processes cpu platform sorted location switch 3 R0 | in fman
 20643   19400      0%       0%       0% S         296599552  fman_fp_image
```

IOSd Memory

```
9300_1#sh processes memory sorted
Processor Pool Total: 1445417856 Used:   290878080  Free: 1154539776
 lsmpi_io Pool Total:    6295128 Used:     6294296  Free:         832

 PID TTY  Allocated      Freed      Holding     Getbufs      Retbufs Process
   0   0  288949984   53306504   214634744           0           0 *Init*
   4   0   22511680     100104    22277344           0           0 RF Slave Main Th
  81   0   24107152    2420648    13717584           0           0 IOSD ipc task
 472   0    4133424     105760     4069608      849828           0 EEM ED Syslog
   0   0   62739512   58194512     2978824    23259559      382788 *Dead*
 609   0    6717728    3968088     2803904           0           0 ISIS Upd
 490   0    1719800      90880     1659432           0           0 EEM Server
```

- IOSd runs as a process , but does still provides some kernel features like memory management for all processes running inside IOSd
- Processor Pool: Pool for Processes on IOSd
- lsmpi_io : Linux Shared Memory Punt Interface memory , IO buffers

Platform Memory

```
9300_1#sh processes memory platform  sorted location switch 1 R0
System memory: 7711304K total, 2388036K used, 5323268K free,
Lowest: 5321604K
    Pid     Text       Data    Stack    Dynamic       RSS                 Name
------------------------------------------------------------------------
   9738   151604     817392      136        408    817392     linux_iosd-imag
  14416      152     224856      136      85308    224856      fed main event
  21595      238     168844      136       2780    168844                 dbm
  20109      176     128988      136       5404    128988            sessmgrd
  21225     6825     117488      136       3004    117488             fman_rp
  22061      260      93396      136        124     93396           cli_agent
  22864      600      91152      136      16444     91152               smand
```

IOSd

FED

SMD

Forwarding manager

- Kernel memory utilization is available per switch
- linux_iosd-image process is IOSd
- Resident Set Size(RSS), memory occupied by each Process

Catalyst IOS-XE Software release schedule

	Amsterdam 17.3	Bengaluru 17.6	Cupertino 17.7	Cupertino 17.8	Cupertino 17.9
Next planned release	17.3.6	17.6.4	None	None	17.9.1
9200/9300/940 9500/9600	Yes	Yes	Yes	Yes	Yes
9300X	No	Yes	Yes	Yes	Yes
9500X/9600X	No	No	Yes	Yes	Yes
Maintenance Throttle	Yes	Yes	No	No	Yes

Maintenance throttles receive more rebuilds and thus would be recommended over feature releases

Not all SKU support all IOS-XE versions, above table indicative

Recommended releases: https://www.cisco.com/c/en/us/support/docs/switches/catalyst-9300-series-switches/214814-recommended-releases-for-catalyst-9200-9.html

Tools

Platform Specific Show commands

- Platform Independent Show commands are similar like on any IOS platform. Ex. Show cdp, show bgp, show udld

- IOS-XE specific troubleshooting commands are under show platform software/hardware

```
Switch#show platform software ip switch active R0 cef
ASR_1k#show platform software ip rp active cef
```

- *Show platform [hardware|software] fed* contain Platform Dependent troubleshooting for the Catalyst 9k (FED layer)

```
9300#show platform software fed switch active ifm mappings
9300#show platform software fed switch 5 ifm mappings
9400#show platform software fed active ifm mappings
```

- Outputs might vary depending on if it executed on active/standy and or/on Forwarding Processor(FP) or Route Processor(RP)

Debugging, non-IOSd Processes

- IOSd cannot do real time debugging on processes outside IOSd

- To facilitate debugging/logging trace logs are available per process

- Tracing levels set with granularity (default notice). Tracing always on

- Common processes: smd , fed, forwarding-manager

```
Switch#set platform software trace smd switch active R0 dot1x-all verbose
Switch#show platform software trace level smd switch active R0 | inc dot1x
dot1x                          Notice
dot1x-all                      Verbose
dot1x-redun                    Notice
Switch#set platform software trace all notice
```

- Processes *do not* run on just active switch but potentially all

Always on Tracing usage

- Tracelog files are stored in crashinfo:/logs in binary format.

- Traces can be displayed using show platform software trace or show logging process command

- Archive of traces can be created using the command
  "request platform software trace archive"

- Archives contain binary files, not readable with text viewer

```
Edge_1#sh logging process smd | inc RADIUS
2022/06/06 23:24:03.268912 {smd_R0-0}{1}: [radius] [24732]: (info): RADIUS: Send Accounting-Request to
10.48.91.222:1813 id 1813/184, len 850
2022/06/06 23:24:03.268937 {smd_R0-0}{1}: [radius] [24732]: (info): RADIUS:  authenticator e5 d1 b7 4d 8b e9 d5
06 - 14 b9 8d b6 8c 29 93 94
2022/06/06 23:24:03.268945 {smd_R0-0}{1}: [radius] [24732]: (info): RADIUS:  Vendor, Cisco        [26]   211
2022/06/06 23:24:03.268954 {smd_R0-0}{1}: [radius] [24732]: (info): RADIUS:   Cisco
AVpair      [1]    205  "cts-pac-opaque="
2022/06/06 23:24:03.268960 {smd_R0-0}{1}: [radius] [24732]: (info): RADIUS:  Vendor, Cisco        [26]    36
2022/06/06 23:24:03.268966 {smd_R0-0}{1}: [radius] [24732]: (info): RADIUS:   Cisco
AVpair      [1]     30  "dc-profile-name=Cisco-Device"
```

Embedded Packet Capture (EPC)

- EPC provides insight into both Data Plane and Control Plane traffic

- Captures can be done on Interfaces or Control Plane

- Data can be buffered and exported or stored directly in flash

- Data capture implemented on Port Asics, traffic copied to EPC process.

- EPC process provides deeper packet capture and display filtering

- Analysis can be done off-box or on box using included packet dissectors

- Packet capture rate limited

Running Embedded Packet Capture

```
Switch#monitor capture CL interface GigabitEthernet 1/0/2 both
Switch#monitor capture CL match ipv4 any any
Switch#monitor capture CL start
Switch#monitor capture CL stop
Switch#show monitor capture CL buffer

Starting the packet display ......... Press Ctrl + Shift + 6 to exit

   1   0.000000 10.254.111.100 -> 10.254.254.1 TCP 74 734 b^F^R 2049 [SYN]
Seq=0 Win=29200 Len=0 MSS=1460 SACK_PERM=1 TSval=445826583 TSecr=0 WS=128
   2   0.000501 10.254.254.1 -> 10.254.111.100 TCP 60 2049 b^F^R 734 [RST, ACK]
Seq=1 Ack=1 Win=0 Len=0
   3   1.001299 10.254.111.100 -> 10.254.254.1 TCP 74 711 b^F^R
Seq=0 Win=29200 Len=0 MSS=1460 SACK_PERM=1 TSval=445826833 TSec
   4   1.001582 10.254.254.1 -> 10.254.111.100 TCP 60 2049 b^F^R 711 [RST, ACK]
Seq=1 Ack=1 Win=0 Len=0

Switch#monitor capture CL export location flash:cl.pcap
```

Where and what to capture

Start and stop the capture

Displays capture buffer

Export capture to file

Displaying packet captures

```
Switch#show monitor capture file flash:cl.cap brief
Starting the packet display ........ Press Ctrl + Shift + 6 to exit
  1   0.000000 10.200.10.100 -> 10.200.10.200 TCP 66 59498 b^F^R 80 [ACK] Seq=1 Ack=1
Switch#show moni capture file flash:cl.pcap packet-number 1 detailed | be Transmission
Transmission Control Protocol, Src Port: 59498 (59498), Dst Port: 80 (80), Seq: 1, Ack:
1,   Source Port: 59498
    Destination Port: 80
    Sequence number: 1     (relative sequence number)
    Acknowledgment number: 1     (relative ack number)
    Header Length: 32 bytes
    Flags: 0x010 (ACK)
        000. .... .... = Reserved: Not set
        ...0 .... .... = Nonce: Not set
        .... 0... .... = Congestion Window Reduced (CWR): Not set
        .... .0.. .... = ECN-Echo: Not set
        .... ..0. .... = Urgent: Not set
        .... ...1 .... = Acknowledgment: Set
        .... .... 0... = Push: Not set
        .... .... .0.. = Reset: Not set
        .... .... ..0. = Syn: Not set
        .... .... ...0 = Fin: Not set
        [TCP Flags: *******A****]
    Window size value: 24464
```

Details packet decodes

Off Box Analysis using Wireshark also possible

Control Plane Policing Statistics

```
9300_1#show policy-map control-plane
 Control Plane
  Service-policy input: system-cpp-policy
    Class-map: system-cpp-police-ios-routing (match-any)
      0 packets, 0 bytes
      5 minute offered rate 0000 bps, drop rate 0000 bps
      Match: none
      police:
          rate 13000 pps, burst 3173 packets
        conformed 379638519 bytes; actions:
          transmit
        exceeded 0 bytes; actions:
          drop
    Class-map: system-cpp-police-ios-feature (match-any)
      0 packets, 0 bytes
      5 minute offered rate 0000 bps, drop rate 0000 bps
      Match: none
      police:
          rate 6000 pps, burst 1464 packets
        conformed 20422413 bytes; actions:
          transmit
        exceeded 0 bytes; actions:
          drop
```

- Control Plane Traffic by default protected with control plane policing
- Multiple Queues might map to one class map
- Rates/actions are configurable, use caution when modifying

Control Plane Policing HW stats

```
9300_1#show plat hardware fed switch active qos queue stats internal cpu policer
                       CPU Queue Statistics
                                       (default) (set)     Queue         Queue
QId PlcIdx  Queue Name             Enabled  Rate   Rate    Drop(Bytes)  Drop(Frames)
-----------------------------------------------------------------------------------
0   11      DOT1X Auth             Yes    1000    1000    0            0
1   1       L2 Control             Yes    2000    2000    0            0
2   14      Forus traffic          Yes    4000    4000    0            0
                       CPU Queue Policer Statistics
Policer      Policer Accept    Policer Accept  Policer Drop  Policer Drop
  Index         Bytes            Frames          Bytes         Frames
-----------------------------------------------------------------------------
0            17261371          11408           0             0
1            17682901          52775           0             0
2            357304765         43037           0             0
                       CPP Classes to queue map
=================================================================================
PlcIdx CPP Class                             :  Queues
-----------------------------------------------------------------------------
0      system-cpp-police-data                :  ICMP GEN/ BROADCAST/ ICMP Redirect/
10     system-cpp-police-sys-data            :  Openflow/ Exception/ EGR Exception/ NFL SAMPLED DATA/ RPF
1      system-cpp-police-l2-control          :  L2 Control/
```

Traffic punted to CPU is punted to a Queue and a Policer Index

Determining Cause of Inband traffic

```
9500_1#show plat software fed switch active punt rates interfaces
Punt Rate on Interfaces Statistics
Packets per second averaged over 10 seconds, 1 min and 5 mins
==============================================================================
                            |        | Recv | Recv | Recv | Drop | Drop | Drop
 Interface Name             | IF_ID  | 10s  | 1min | 5min | 10s  | 1min | 5min
==============================================================================
 TenGigabitEthernet1/0/1      0x0000002e     1      1      1      0      0      0
 FortyGigabitEthernet1/1/1    0x00000032     1      1     18      0      0      0

9500_1#show platform software fed switch active punt cpuq rates
Punt Rate CPU Q Statistics
Packets per second averaged over 10 seconds, 1 min and 5 mins
==============================================================================
Q  |         Queue     | Rx   | Rx   | Rx   | Drop   | Drop   | Drop
no |         Name      | 10s  | 1min | 5min | 10s    | 1min   | 5min
==============================================================================
 2   CPU_Q_FORUS_TRAFFIC       0      0     17      0      0      0
 4   CPU_Q_ROUTING_CONTROL     3      3      3      0      0      0
```

Active interfaces sending to cpu

Per Queue Statistics

Show Tech Enhancements

- Show tech contains lot of generic information, not feature specific

- For more focused information gathering show tech <keyword>

- Scripted command generation based on provided parameters

- Examples:
  - show tech-support cts
  - show tech-support cef
  - show tech-support acl
  - show tech-support fabric

> Show tech can be large, redirect to flash

```
9300_1#sh tech identity mac 0001.0001.0001 interface Gi 1/0/1 | redirect flash:shtech.log
9300_1#dir flash:shtech.log
Directory of flash:/shtech.log
671754   -rw-          1504931  Jun 10 2019 00:07:47 +00:00  shtech.log
11353194496 bytes total (9337597952 bytes free)
```

Packet Drops

Ethernet Interface Utilization

```
Switch#show interfaces | inc line|rate
Vlan1 is up, line protocol is up , Autostate Enabled
  5 minute input rate 0 bits/sec, 0 packets/sec
  5 minute output rate 0 bits/sec, 0 packets/sec
GigabitEthernet0/0 is administratively down, line protocol is d
  5 minute input rate 0 bits/sec, 0 packets/sec
  5 minute output rate 0 bits/sec, 0 packets/sec
GigabitEthernet1/0/1 is up, line protocol is up (connected)
  5 minute input rate 103000 bits/sec, 174 packets/sec
  5 minute output rate 3879000 bits/sec, 324 packets/sec
GigabitEthernet1/0/2 is down, line protocol is down (notconnect)
  5 minute input rate 0 bits/sec, 0 packets/sec
  5 minute output rate 0 bi
GigabitEthernet1/0/3 is dow
  5 minute input rate 0 bit
  5 minute output rate 0 bi
```

Input/output rates show average over 5 minutes.

```
9500#show controllers utilization
Port            Receive Utilization   Transmit Utilization
Te1/0/1                   3                    2
Te1/0/2                   2                    3
.
Te1/0/16                  0                    0
Fo1/1/1                   0                    0
Fo1/1/2                   0                    0
Total Ports : 18
Total Ports Receive Bandwidth Percentage Utilization  : 0
Total Ports Transmit Bandwidth Percentage Utilization : 0
Average Switch  Percentage Utilization : 0
```

Bandwidth in % Current load

Ethernet Interface Utilization

```
Switch#sh controllers ethernet-controller gi 5/0/48
Transmit                    GigabitEthernet5/0/48 Receive
   1562496684 Total bytes                 2968958225 Total bytes
      5032561 Unicast frames                 6004241 Unicast frames
    700808558 Unicast bytes               1807110661 Unicast bytes
      1269484 Multicast frames               2789759 Multicast frames
    861688062 Multicast bytes             1161847500 Multicast bytes
            1 Broadcast frames                     1 Broadcast frames
            0 Cos 0 Pause frames                   0 Cos 0 Pause frames
      1236978 Minimum size frames             871517 Minimum size frames
      1892419 65 to 127 byte frames          2181611 65 to 127 byte frames
      1941967 128 to 255 byte frames         2712229 128 to 255 byte frames
       685594 256 to 511 byte frames         1260418 256 to 511 byte frames
        20261 512 to 1023 byte frames         900135 512 to 1023 byte frames
       524827 1024 to 1518 byte frames        868091 1024 to 1518 byte frames
            0 8192 to 16383 byte frames            0 8192 to 16383 byte frames
            0 16384 to 32767 byte frame             0 16384 to 32767 byte frame
            0 > 32768 byte frames                   0 > 32768 byte frames
            0 Late collision frames                 0 SymbolErr frames
            0 Excess Defer frames                   0 Collision fragments
            0 Good (1 coll) frames                  0 ValidUnderSize frames
            0 Good (>1 coll) frames                 0 InvalidOverSize frames
            0 Deferred frames                       0 ValidOverSize frames
LAST UPDATE 361 msecs AGO
```

Ethernet controller statistics give more detailed port statistics

Tail Drops

- Tail drops occur when exceeding buffer thresholds on overloaded links

```
9300_1#show interfaces gigabitEthernet 1/0/1 | inc output drops
   Input queue: 0/2000/0/0 (size/max/drops/flushes); Total output drops: 1277
9300_1#show controllers ethernet-controller gig 1/0/1 | inc Excess Def
         1277 Excess Defer frames              0 Collision fragments
SNMP:
SNMPv2-SMI::enterprises.9.2.2.1.1.27.8 = INTEGER: 1277
```

- Buffer allocation per class can be changed inside service-policy

- Global multiplier to increase buffers up to 1200%
    qos softmax-queue-multiplier <percentage>

- Increasing buffers increases maximum, buffers allocated based on availability

- 17.1.1 introduces CLI to monitor high watermark utilization on ports

QoS Hardware configuration

```
9300_1#sh plat hard fed switch active qos queue config interface gi 1/0/1
Asic:0 Core:1 DATA Port:0 GPN:1 LinkSpeed:0x1
----------------------------------------------------------------
   DTS   Hardmax   Softmax    PortSMin  GlblSMin   PortStEnd
   -----  --------  --------    --------  --------   ---------
0  1  5    200        3200    5   500    0    0    6  9600
1  1  4     0   13   4800    5   750    2   300   6  9600
Priority      Shaped/shared    weight    shaping_step    sharpedWeight
--------     -------------    ------    -----------    -------------
0     0       Shared             50            0              0
1     0       Shared             75            0              0
Port          Port             Port      Port
Priority      Shaped/shared    weight    shaping_step
--------     -------------    ------    -----------
      2       Shaped            254           255
   Weight0 Max_Th0 Min_Th0 Weight1 Max_Th1 Min_Th1  Weight2 Max_Th2 Min_Th2
   ------- ------- ------- ------- ------- -------  ------- ------- ------
0     0    2709        0       0    3028        0       0    3400       0
1     0    3825        0       0    4275        0       0    4800       0
```

- **Hardmax.**
  Reserved buffers
- **Softmax.**
  Non Reserved

- Queue mode
- Queue limit:
  Step/weight * speed

- Drop thresholds
  per queue in buffers

- QoS configured using service policies on interfaces
- Applied service-policy translated into Hardware settings that match HW capabilities

CISCO Live!

QoS hardware statistics

```
9300_1#sh platform hardware fed switch active qos queue stats interface gigabitEthernet 1/0/1
-----------------------------------------------------------------------------------------------
AQM Global counters
GlobalHardLimit:  7976  |  GlobalHardBufCount: 0
GlobalSoftLimit: 11872  |  GlobalSoftBufCount: 0
-----------------------------------------------------------------------------------------------

High Watermark Soft Buffers:      0 <--- clear on read
-----------------------------------------------------------------------------------------------
Asic:0 Core:1 DATA Port:0 Hardware Enqueue Counters
-----------------------------------------------------------------------------------------------
 Q Buffers        Enqueue-TH0          Enqueue-TH1          Enqueue-TH2          Qpolicer
   (Count)           (Bytes)              (Bytes)              (Bytes)            (Bytes)
-- -------   ------------------   ------------------   ------------------   ------------------
 0     0                    0               385820             46085690                    0
 1     0                    0                    0                    0                     0
Asic:0 Core:1 DATA Port:0 Hardware Drop Counters
-----------------------------------------------------------------------------------------------
 Q        Drop-TH0             Drop-TH1             Drop-TH2            SBufDrop             QebDrop            QpolicerDrop
          (Bytes)              (Bytes)              (Bytes)            (Bytes)              (Bytes)             (Bytes)
-- ------------------   ------------------   ------------------   ------------------   ------------------   ------------------
 0        0                    0                    0                    0                    0                    0
 1        0                    0                    0                    0                    0                    0
```

set plat hard fed active qos port-monitor interface <if>

- At UADP level there are 8 Queues/3 Thresholds
- Enqueue/Drop Counters available per queue/per threshold
- Buffers (count) show currently assigned buffers to Queue (256 bytes)
- High water mark counter (17.x) shows highest watermark since last output

ASIC packet forwarding drop counters

```
9300_1#sh platform hardware fed switch active fwd-asic drops exceptions
****EXCEPTION STATS ASIC INSTANCE 0 (asic/core 0/0)****

=========================================================================
 Asic/core |                  NAME               | prev  |  current  | delta
=========================================================================
0   0   NO_EXCEPTION                              35364016   35364108   92
0   0   IPV4_CHECKSUM_ERROR                       0          0          0
0   0   ROUTED_AND_IP_OPTIONS_EXCEPTION           2          2          0
0   0   CTS_FILTERED_EXCEPTION                    0          0          0
0   0   AUTH_DRIVEN_DROP                          0          0          0
0   0   PKT_DROP_COUNT                            0          3732       3732
0   0   ALLOW_DOT1Q_EXCEPTION_COUNT               0          0          0
0   0   ALLOW_PRIORITY_TAGGED_EXCEPTION_COUNT     0          0          0
0   0   IGR_EXCEPTION_L5_ERROR                    0          363        363
0   0   IP_UNICAST_TTL_REACHED_ZERO               0          0          0
0   0   MISC_FATAL_ERROR                          0          0          0
```

- Every packet passing through Port Asic gets parsed by port asic
- Exception drops are counted per Asic , no per port statistics

Forwarding

Troubleshooting Forwarding

- UADP responsible for all hardware forwarding. CPU is not directly involved in forwarding.

- IOS-XE uses Forwarding manager, Forwarding manager process manages forwarding related information but only on a Platform Independent level. Forwarding manager is on all platforms running IOS-XE16+

- FED (Forwarding Engine Driver) process is Platform Dependant layer, interfaces with Forwarding manager and responsible for all tables on UADP asics

IOSd

Forwarding Manager

FED
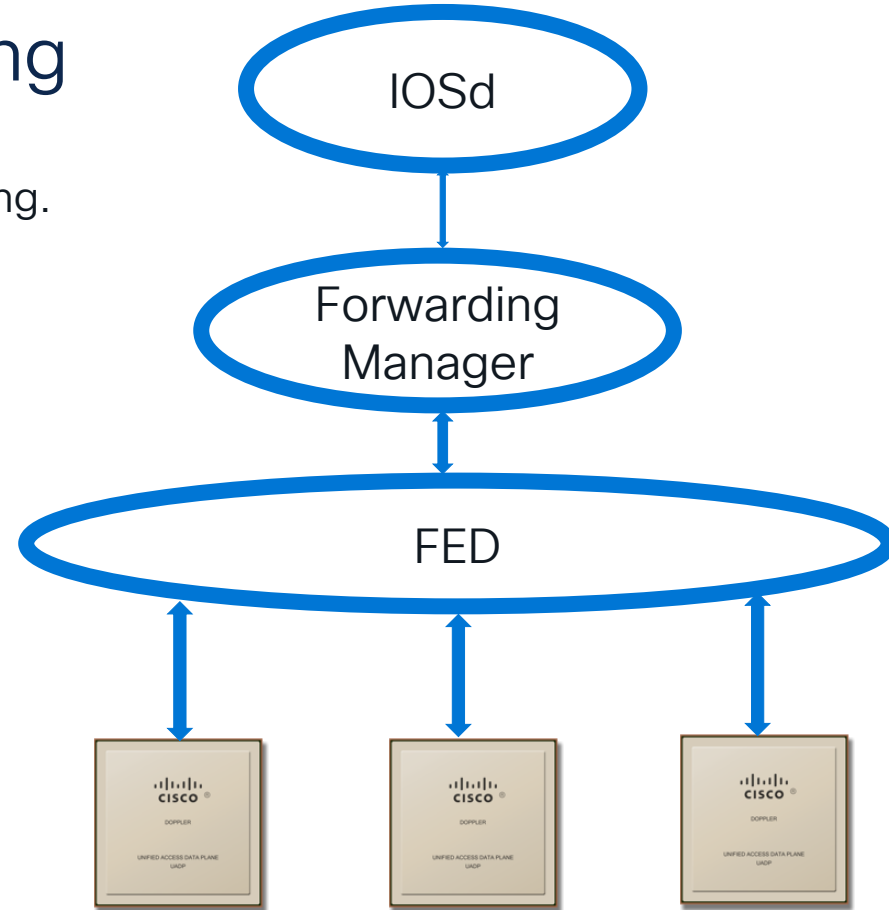
Forwarding troubleshooting, the easy way

- Show forward supported since 2900/3500XL switches , up to 3750 family only software emulation of forwarding results were used

- UADP introduced HW captures of lookup results during various stages of packet forwarding

- CLI: "*Show platform hardware fed switch <ingress switch> forward …* "

- Supports Input using packet capture file or packet parameters

- When using pcap file needs to be present on flash on ingress switch

- Readability of results greatly improved since 16.9 with summary option

- Packets introduced for captures inserted at Ingress and removed at Egress before sending out on the wire.

Running Show platform hardware fed forward

```
9300_1#sh monitor capture file flash:icmp.pcap packet-number 11
Starting the packet display ........ Press Ctrl + Shift + 6 to exit
  11   5.006009 10.100.10.100 b^F^R 10.200.10.200 ICMP 98 Echo (ping) request  id=0x262f
```

Verify capture!

- Using the packet capture the show forward can be executed
- Execute on switch where the packet ingresses pcap needs to be in that switches flash

```
9300_1#sh plat hard fed 1 forward int gi 1/0/1 pcap flash:icmp.pcap num 11 data
Show forward is running in the background. After completion, syslog will be generated.
```

- Once completed a syslog gets generated and results will be available
- Can only run one show forward at a time

```
*Jan 27 10:07:35.009: %SHFWD-6-PACKET_TRACE_DONE: Switch 1 R0/0: fed: Packet Trace
Complete:  Execute (show platform hardware fed switch <> forward last summary|detail)
```

Show platform forward results

```
9300_1#sh platform hardware fed switch active forward last summary
Input Packet Details:
###[ Ethernet ]###
  dst        = a0:f8:49:10:48:51
  src        = 00:0c:29:4d:9e:16
  type       = 0x800
###[ IP ]###
     version   = 4L
     ihl       = 5L
     tos       = 0x0
     len       = 84
     id        = 46165
     flags     = DF
     frag      = 0L
     ttl       = 64
     proto     = icmp
     chksum    = 0x5bfc
     src       = 10.100.10.100
     dst       = 10.200.10.200
     options   = ''
```

Input packet used in show platform forward

Show platform forward results

```
Ingress:
    Port                    : GigabitEthernet1/0/1
    Global Port Number      : 1
    Local Port Number       : 1
    Asic Port Number        : 0
    Asic Instance           : 1
    Vlan                    : 100
    Mapped Vlan ID          : 6
    STP Instance            : 4
    BlockForward            : 0
    BlockLearn              : 0
    L3 Interface            : 50
        IPv4 Routing        : enabled
        IPv6 Routing        : enabled
        Vrf Id              : 0
    Adjacency:
        Station Index       : 185
        Destination Index   : 21358
        Rewrite Index       : 34
        Replication Bit Map : 0x10   ['coreData']
```

Ingress port

Layer 2 forwarding parameters

Layer 3 forwarding parameters

Forwarding Information

Show platform forward results

```
Adjacency:
      Station Index       : 185
      Destination Index   : 21358
      Rewrite Index       : 34
      Replication Bit Map  : 0x10   ['coreData']
  Decision:
      Destination Index   : 21358
      Rewrite Index       : 34
      Dest Mod Index      : 0        [IGR_FIXED_DMI_NULL_VALUE]
      CPU Map Index       : 0        [CMI_NULL]
      Forwarding Mode     : 3        [Other or Tunnel]
      Replication Bit Map  :          ['coreData']
      Winner              :          L3FWDIPV4 LOOKUP
      Qos Label           : 1
      SGT                 : 0
      DGTID               : 0
```

> **Destination Index determines egress interface**

> **Rewrite Index determines packet processing action**

> **Winner indicates what lookup was used to determine final forwarding decision.**

Show platform forward results

```
Egress:
    Possible Replication         :
        Port                     : TenGigabitEthernet1/1/7
    Output Port Data             :
        Port                     : TenGigabitEthernet1/1/7
        Global Port Number       : 59
        Local Port Number        : 59
        Asic Port Number         : 58
        Asic Instance            : 0
        Unique RI                : 34
        Rewrite Type             : 9        [L3_UNICAST_IPV4]
        Mapped Rewrite Type      : 9        [L3_UNICAST_IPV4]
        Vlan                     : 0
        Mapped Vlan ID           : 0
```

Last section shows Egress Packet

```
Output Packet Details:
    Port                       : TenGigabitEthernet1/1/7
###[ Ethernet ]###
    dst       = 00:50:56:92:48:d8
    src       = a0:f8:49:10:48:66
    type      = 0x800
###[ IP ]###
    version   = 4L
    ihl       = 5L
  tos         = 0x0
    len       = 84
    id        = 46165
    flags     = DF
    frag      = 0L
    ttl       = 63
    proto     = icmp
    chksum    = 0x5cfc
    src       = 10.100.10.100
    dst       = 10.200.10.200
    options   = ''
```

**Packet will not egress switch.
Dropped at last stages of processing.**

Show platform hardware fed forward detail results

- Flash contains shfwd<>.log with detailed information

- Same information gathered with
  show platform hardware fed switch <switch> forward last detail

- Detail info containing raw information regarding lookup results

```
9300_1#sh platform hardware fed switch 1 forward last detail
--------------------------------------------------
Starting IPP capture
=============================
IppDefaultClientTable[4]
------------------------------------
  defaultClientLeIndex          1
===================================
IppClientLeAd[1]
--------------------------------------------------------
                LEAD_CLIENT_ANCHORED            0
         LEAD_CLIENT_AUTH_BEHAVIOR_TAG          0
              LEAD_CLIENT_CLIENT_GROUP          0
```

TCAM utilization

- Hardware forwarding occurs only when tables download into Hardware.

- Running out of Hardware resources results in possible performance issues

- Multiple SDM templates available on most platforms to get different allocation

```
Switch#show platform hardware fed active fwd-asic resource tcam utilization
Codes: EM - Exact_Match, I - Input, O - Output, IO - Input & Output, NA - Not Applicable
CAM Utilization for ASIC  [0]
Table                 Subtype        Dir     Max      Used    %Used      V4        V6       MPLS      Other
-----------------------------------------------------------------------------------------------------------
Mac Address Table     EM             I      65536       53    0.08%       0         0         0         53
Mac Address Table     TCAM           I       1024       21    2.05%       0         0         0         21
L3 Multicast          EM             I      16384        7    0.04%       7         0         0          0
L3 Multicast          TCAM           I       1024       12    1.17%       6         6         0          0
L2 Multicast          EM             I      16384        0    0.00%       0         0         0          0
L2 Multicast          TCAM           I       1024       11    1.07%       3         8         0          0
IP Route Table        EM             I      49152       72    0.15%      55         6        11          0
IP Route Table        TCAM           I      65536      107    0.16%      86        18         2          1
```

Layer 2 Forwarding. Verifying STP state

- Show spanning tree gives IOSd view of Spanning Tree
- Hardware forwarding states can be checked *per switch* on FED layer
- Outputs will show what interface are in forwarding state and if traffic will be tagged or untagged
- Flood list indicates what Ports will receive flooded traffic on this switch

```
9300_1#show platform hardware fed switch 1 vlan 100 egress
VLAN STP State in hardware
vlan id is:: 100
Interfaces in forwarding state: : Te1/1/7(Untagged), Gi1/0/1(Untagged)
9300_1#show platform hardware fed switch 1 vlan 100 ingress
VLAN STP State in hardware
vlan id is:: 100
Interfaces in forwarding state: : Te1/1/7(Untagged), Gi1/0/1(Untagged)
flood list: : Te1/1/7, Gi1/0/1
```

Layer 2 Forwarding, IOSd mac address tables

```
9300_1#sh mac address-table vlan 100
          Mac Address Table
-------------------------------------------
Vlan     Mac Address        Type        Ports
----     -----------        --------    -----
 100     000c.294d.9e16     DYNAMIC     Gi1/0/1
 100     0050.5692.adb3     DYNAMIC     Gi1/0/1
 100     0050.5692.e9aa     DYNAMIC     Gi1/0/1
 100     a0f8.4910.4851     STATIC      Vl100
Total Mac Addresses for this criterion: 4
```

- Show mac address table contains a system wide mac table on IOSd
- Types can be static, dynamic, drop
- Mac Address of SVI interfaces also showing in mac address table

FED MATM Mac Address Table

```
9300_1#sh platform software fed switch 1 matm macTable vlan 100
VLAN    MAC                 Type    Seq#   EC_Bi  Flags  machandle          siHandle          diHandle          *a_time   *e_time  ports
-------------------------------------------------------------------------------------------------------------------------------------------
100     a0f8.4910.4851      0x8002     0   99817     64  0x7f91986dfcd8     0x7f9198dadb78    0x0               0         0        Vlan100
100     0050.5692.e9aa      0x1      347      0       0  0x7f9199054668     0x7f9199020798    0x7f91986e4a58    300       234      Gi1/0/1
100     0050.5692.adb3      0x1      352      0       0  0x7f91990144a8     0x7f9199020798    0x7f91986e4a58    300       71       Gi1/0/1
100     000c.294d.9e16      0x1      364      0       0  0x7f919900e9d8     0x7f9199070018    0x7f91986e4a58    300       290      Gi1/0/1

*a_time=aging_time(secs)   *e_time=total_elapsed_time(secs)
Type:
MAT_DYNAMIC_ADDR          0x1    MAT_STATIC_ADDR          0x2    MAT_CPU_ADDR          0x4    MAT_DISCARD_ADDR          0x8
MAT_ALL_VLANS             0x10   MAT_NO_FORWARD           0x20   MAT_IPMULT_ADDR       0x40   MAT_RESYNC                0x80
MAT_DO_NOT_AGE            0x100  MAT_SECURE_ADDR          0x200  MAT_NO_PORT           0x400  MAT_DROP_ADDR             0x800
MAT_DUP_ADDR             0x1000  MAT_NULL_DESTINATION     0x2000  MAT_DOT1X_ADDR       0x4000  MAT_ROUTER_ADDR          0x8000
MAT_WIRELESS_ADDR       0x10000  MAT_SECURE_CFG_ADDR     0x20000  MAT_OPQ_DATA_PRESENT  0x40000  MAT_WIRED_TUNNEL_ADDR 0x80000
MAT_DLR_ADDR           0x100000  MAT_MRP_ADDR           0x200000  MAT_MSRP_ADDR       0x400000  MAT_LISP_LOCAL_ADDR   0x800000
MAT_LISP_REMOTE_ADDR 0x1000000  MAT_VPLS_ADDR         0x2000000
```

- Every FED has its own Mac address table.
- Type Field indicates the type of mac address using a bitmap
- Sequence number of an entry changing would indicated relearning

Layer 3 Forwarding. Routing protocols

```
9300_1#ping 10.48.91.151
Sending 5, 100-byte ICMP Echos to 10.48.91.151, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 1/1/1 ms
9300_1#sh ip route 10.48.91.151
Routing entry for 10.48.91.128/25
  Known via "isis", distance 115, metric 40, type level-2
  Redistributing via isis
  Last update from 172.31.250.30 on TenGigabitEthernet1/1/6, 6d22h ago
  Routing Descriptor Blocks:
  * 172.31.250.30, from 172.31.255.21, 6d22h ago, via TenGigabitEthernet1/1/6
      Route metric is 40, traffic share count is 1
9300_1#sh ip arp 172.31.250.30
Protocol  Address         Age (min)  Hardware Addr   Type   Interface
Internet  172.31.250.30        56    2c4f.523b.c142  ARPA   TenGigabitEthernet1/1/6
```

- Check Routing Table for correct next hop (Routing Information Base)
- ARP table display rewrite information for next hop (destination mac)

Cisco Express Forwarding (the FIB)

```
9300_1#sh ip cef 10.48.91.128/25 internal
10.48.91.128/25, epoch 7, RIB[I], refcnt 6, per-destination sharing
  sources: RIB
  feature space:
    IPRM: 0x00028000
    Broker: linked, distributed at 4th priority
  ifnums:
    TenGigabitEthernet1/1/6(85): 172.31.250.30
  path list 7F3B3265DE78, 139 locks, per-destination, flags 0x4D [shble, hvsh, rif, hwcn]
    path 7F3B32181A60, share 1/1, type attached nexthop, for IPv4
      nexthop 172.31.250.30 TenGigabitEthernet1/1/6, IP adj out of TenGigabitEthernet1/1/6,
      addr 172.31.250.30 7F3B33B02738
  output chain:
    IP adj out of TenGigabitEthernet1/1/6, addr 172.31.250.30 7F3B33B02738
9300_1#sh adjacency 172.31.250.30 detail
Protocol Interface                Address
IP       TenGigabitEthernet1/1/6  172.31.250.30(89)
                                  0 packets, 0 bytes
                                  epoch 0
                                  sourced in sev-epoch 11
                                  Encap length 14
                                  2C4F523BC142A0F8491048500800
                                  L2 destination address byte offset 0
                                  L2 destination address byte length 6
                                  Link-type after encap: ip
                                  ARP
```

Source of route, RIB

Next hop

Adjacency -> rewrite info

Platform CEF tables (RP)

```
9300_1#sh platform software ip switch ac R0 cef prefix 10.48.91.128/25
Forwarding Table
Prefix/Len                        Next Object        Index
------------------------------------------------------------------
10.48.91.128/25                   OBJ_ADJACENCY      0x24
9300_1#sh platform software adjacency switch active R0 index 0x24
Number of adjacency objects: 17
Adjacency id: 0x24 (36)
  Interface: TenGigabitEthernet1/1/6, IF index: 85, Link Type: MCP_LINK_IP
  Encap: 2c:4f:52:3b:c1:42:a0:f8:49:10:48:50:8:0
  Encap Length: 14, Encap Type: MCP_ET_ARPA, MTU: 9100
  Flags: no-l3-inject
  Incomplete behavior type: None
  Fixup: unknown
  Fixup_Flags_2: unknown
  Nexthop addr: 172.31.250.30
  IP FRR MCP_ADJ_IPFRR_NONE 0
  OM handle: 0x348066bc48
```

Prefix points to Adjacency 0x24

Similar output should be present on standby RP and the FP processes

FED Routing tables

```
9300_1#sh platform software fed switch 1 ip route 10.48.91.128/25
vrf    dest              htm             flags   SGT    DGID MPLS Last-modified
---    ----              ---             -----   ---    ---- ---- -----------------------
0      10.48.91.128/25  0x7f9199010f78 0x0      0      0         2020/01/14 13:49:42.054
  FIB: prefix_hdl:0x5a00026d, mpls_ecr_prefix_hdl:0
  ========== OCE chain =====
  ADJ:objid:36 {link_type:IP ifnum:0x55, si:0x9b00003d, IPv4:   172.31.250.30 }
  ==============
  MPLS info: mpls_ecr_scale_prefix_adj:0, mpls_lspa_hdl:0
9300_1#sh platform software fed switch 1 ip adj | inc dest|--|172.31.250.30
dest            if_name      dst_mac       si_hdl   ri_hdl   pd_flags adj_id  Last-modified
----            -------      -------       ------   ------   -------- -----   ----------------------
172.31.250.30   Te1/1/6      2c4f.523b.c142  0x7f9198 0x7f9198 0x0      0x24    14:09:12.058
```

- FED layer has its own copy of the IP routing table and rewrite information
- In a stacked enviroment every switch has its own FED process.
  Important to check Ingress and Egress switch
- FED programs TCAM to facilitate forwarding
- Every VRF has its own unique number, 0 is Global Routing Table

Related sessions:

- BRKTRS-2811a & BRKTRS-2811b :
   Overview of Packet Capturing Tools in Cisco Switches and Routers
- LABTRS-2391:
   Packet Capturing Tools in Enterprise Switching Environments

Technical Session Surveys

- Attendees who fill out a minimum of four session surveys and the overall event survey will get Cisco Live branded socks!

- Attendees will also earn 100 points in the Cisco Live Game for every survey completed.

- These points help you get on the leaderboard and increase your chances of winning daily and grand prizes.

Cisco Networking Bot

https://cnb.cisco.com/

Empowering Users by Digitizing Cisco Product & Adoption Experiences

EoS/EoL

Product Migration/Adoption

Config / Tshoot Guides

Security Advisories

Release Recommendations

HW-SW Compatibility

Transforming Customer Experience with Cisco AI Chatbots

Got a Question on Enterprise Products?
Chat with CN Bot now

CISCO

Cisco Learning and Certifications

From technology training and team development to Cisco certifications and learning plans, let us help you empower your business and career. www.cisco.com/go/certs

Pay for Learning with Cisco Learning Credits

(CLCs) are prepaid training vouchers redeemed directly with Cisco.

Learn

Cisco U.
IT learning hub that guides teams and learners toward their goals

Cisco Digital Learning
Subscription-based product, technology, and certification training

Cisco Modeling Labs
Network simulation platform for design, testing, and troubleshooting

Cisco Learning Network
Resource community portal for certifications and learning

Train

Cisco Training Bootcamps
Intensive team & individual automation and technology training programs

Cisco Learning Partner Program
Authorized training partners supporting Cisco technology and career certifications

Cisco Instructor-led and Virtual Instructor-led training
Accelerated curriculum of product, technology, and certification courses

Certify

Cisco Certifications and Specialist Certifications
Award-winning certification program empowers students and IT Professionals to advance their technical careers

Cisco Guided Study Groups
180-day certification prep program with learning and support

Cisco Continuing Education Program
Recertification training options for Cisco certified individuals

Here at the event? Visit us at The Learning and Certifications lounge at the World of Solutions

Continue your education

- Visit the Cisco Showcase for related demos

- Book your one-on-one Meet the Engineer meeting

- Attend the interactive education with DevNet, Capture the Flag, and Walk-in Labs

- Visit the On-Demand Library for more sessions at www.CiscoLive.com/on-demand

CISCO Live!

Thank you