The bridge to possible

Advanced ASR 9000 Operation & Troubleshooting

Mike Mikhail, Delivery Architect mamikhai@cisco.com
Aleksandar Vidakovic, Principal Engineer avidakov@cisco.com
Thomas Wang, Technical Leader pewang@cisco.com

cisco Live!

TECSPG-3204

Cisco Webex App

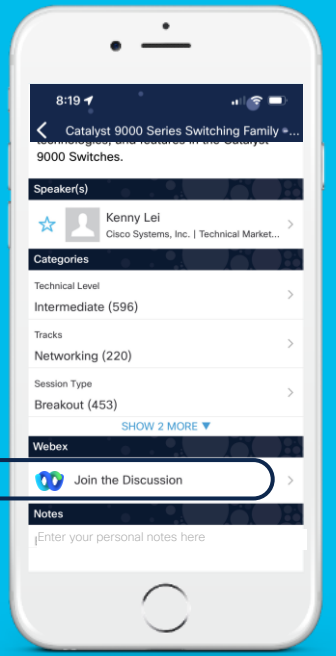
Questions?

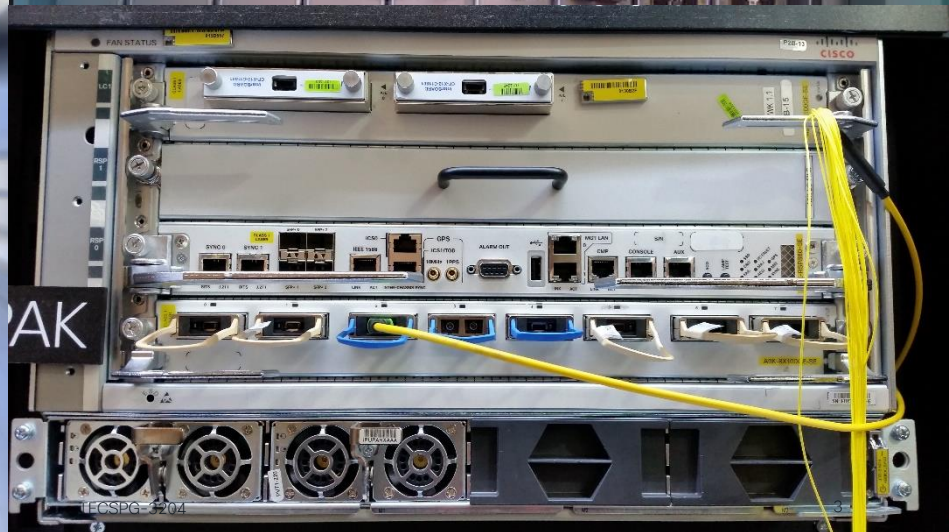
Use Cisco Webex App to chat
with the speaker after the session

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 until February 24, 2023.




FAN STATUS
CISCO
GPS
ALARM OUT
MGT LAN
CNS
CONSOLE
AUX

Agenda

- **System Architecture**: System anatomy & health
- **Operating System & Configuration**: IOS-XR & configuration models
- **Control, Management, Security**: Processing of control & exceptions
- **Transit Packet/Frame Journey**: Life of L3/L2 unicast/multicast
- **MPLS Operation**: Processing, forwarding & L3/L2 service operation
- **Troubleshooting**: Diagnostics, counters, drops, and packet capture

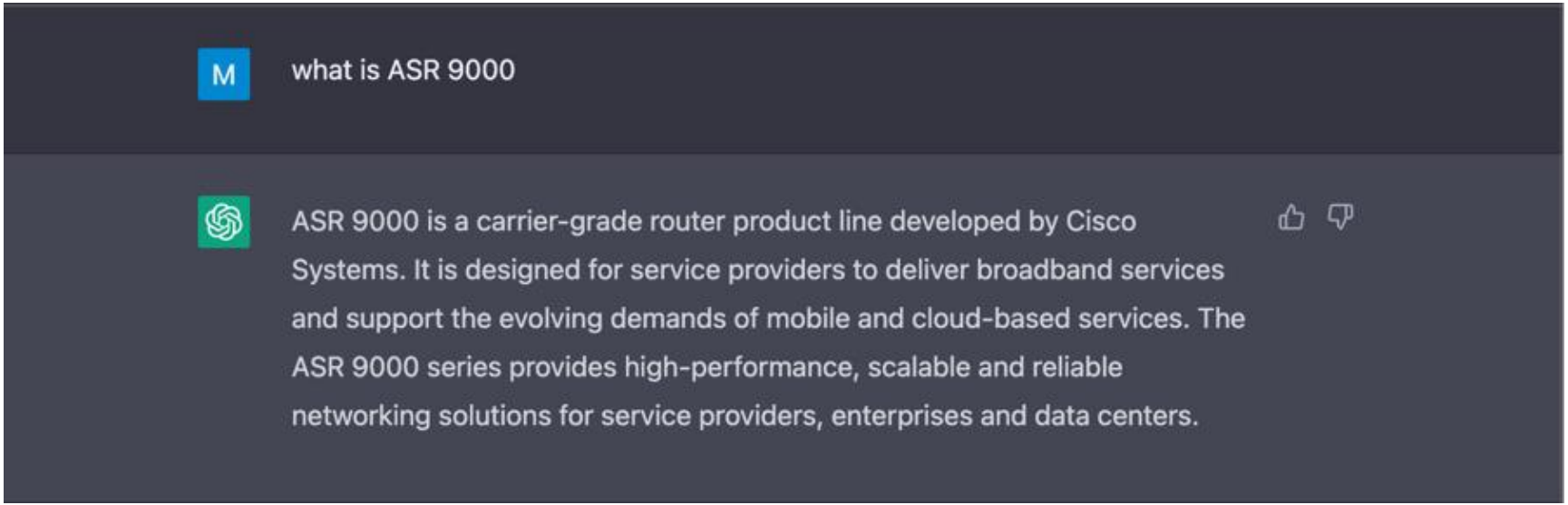
Not covered

- We can have separate discussions about:
  - Clustering
  - Satellites in pdf
  - 1st generation hardware
  - 2nd generation (Typhoon) is in pdf
  - SIP-700 and TDM
  - ISM & VSM
  - Technology and protocol troubleshooting [will focus on platform specifics]

Introduction

- About us
  - Mike Mikhail, Delivery Architect, mamikhai@cisco.com
  - Aleks Vidakovic, Principal Engineer, avidakov@cisco.com
  - Thomas Wang, Technical Leader, pewang@cisco.com
  - Available at “Meet the Engineer” for 1:1 discussions
  - Interests: SRv6, SP platforms & technologies, API’s, ML, Telemetry
- ASR 9000 today
  - The Best SP & WAN edge router. High bandwidth Ethernet services platform
  - HW forwarding, high density: ~3B pps per LC, >4Tbps/slot
  - Distributed processing and distributed forwarding
  - Continued dev, rich roadmap, a wealth of new features

ASR 9000 is known as



Agenda

- **System Architecture**: System anatomy & health
- **Operating System & Configuration**: IOS-XR & configuration models
- **Control, Management, Security**: Processing of control & exceptions
- **Transit Packet/Frame Journey**: Life of L3/L2 unicast/multicast
- **MPLS Operation**: Processing, forwarding & L3/L2 service operation
- **Troubleshooting**: Diagnostics, counters, drops, and packet capture

1 System Architecture

System Components & Sub-Systems

ASR 9900 Family

Chassis, cards, power, air flow

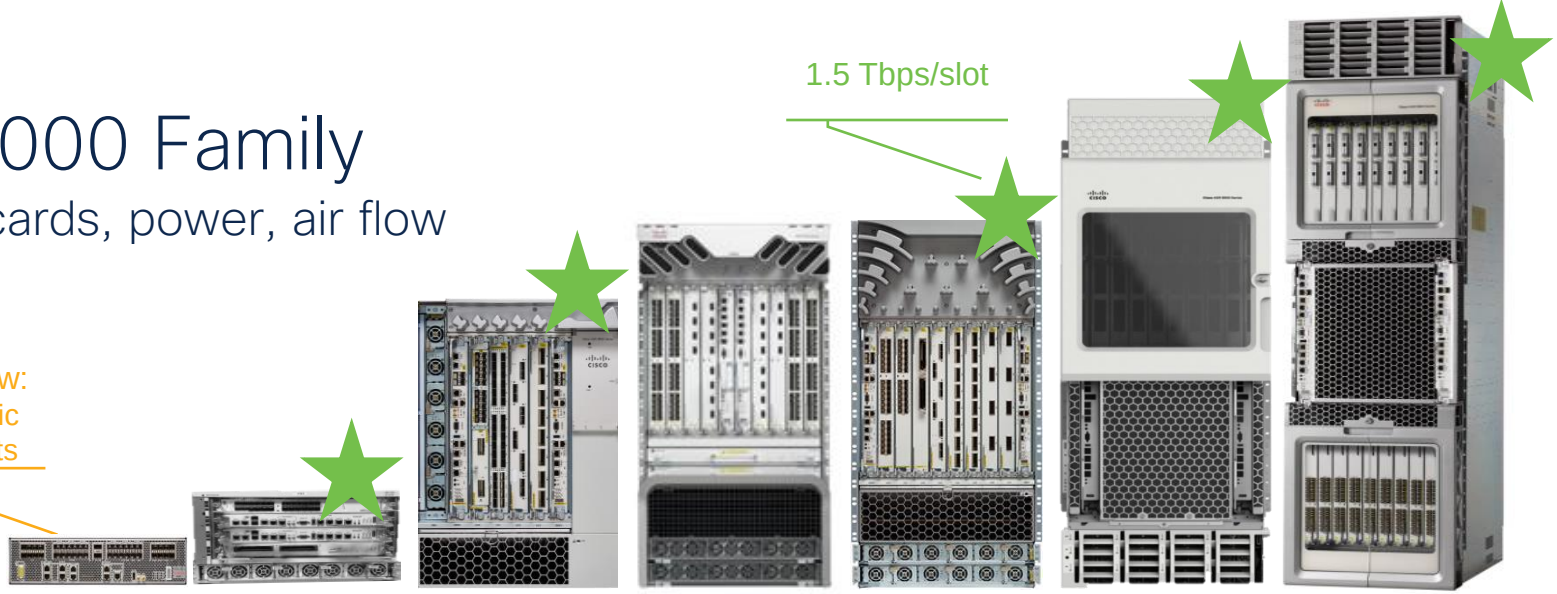
	ASR 9901	ASR 9902	ASR 9903	ASR 9904	ASR 9906	ASR 9910	ASR 9912	ASR 9922
RP	Built-in	1+1 RP	1+1 RP	1+1 RSP	1+1 RSP	1+1 RSP	1+1 RP	1+1 RP
Fabric	Built-in	Built-in	Built-in	2x RSP	6+1	6+1	6+1	6+1
Line cards or ports	16x1G + 24x1/10G + 2x100G	40x1/10G + 16x25G + 4x100G	20x1/10G + 16x100G + PEC	2	4	8	10	20
Rack units	2	2	3	6	14	21	30	44
Power modules	2x AC or DC	2x AC or DC	4x AC or DC	4x AC or DC	3x AC or 4x DC	8x AC or DC	12x AC or DC	16x AC or DC
Air flow	Front to back	Front to back	Front to back	Right to left	Front to back	Front to back	Front to back	Front to back

ASR 9000 Family

Chassis, cards, power, air flow

1.5 Tbps/slot

"Fixed" hw:
RP+Fabric
+LC+ports



	ASR 9901	ASR 9904	ASR 9906	ASR 9010	ASR 9910	ASR 9912	ASR 9922
RP	Built-in	1+1 RSP	1+1 RSP	1+1 RSP	1+1 RSP	1+1 RP	1+1 RP
Fabric	Built-in	2x RSP	6+1	2x RSP	6+1	6+1	6+1
Line cards or ports	16x1G + 24x1/10G + 2x100G	2	4	8	8	10	20
Rack units	2	6	14	21	21	30	44
Power modules	2x AC or 2x DC	4x AC or 4x DC	3x AC or 4x DC	8x AC or 8x DC	8x AC or 8x DC	12x AC or 12x DC	16x AC or 16x DC
Air flow	Front to back	Right to left	Front to back	Front to back	Front to back	Front to back	Front to back

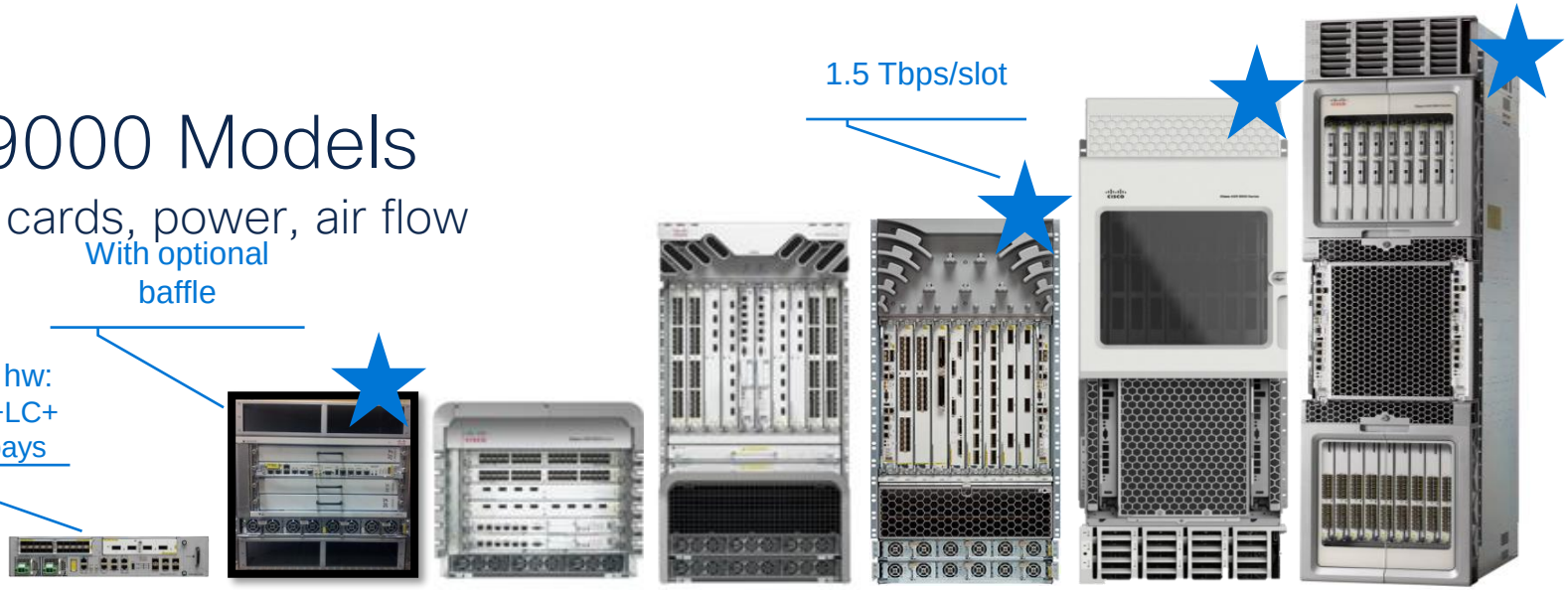
ASR 9000 Models

Chassis, cards, power, air flow

1.5 Tbps/slot

With optional baffle

"Fixed" hw: RP+SP+LC+ ports+bays



	ASR 9001	ASR 9904	ASR 9006	ASR 9010	ASR 9910	ASR 9912	ASR 9922
RP	Built-in	1+1 RSP	1+1 RSP	1+1 RSP	1+1 RSP	1+1 RP	1+1 RP
Fabric	Built-in	2x RSP	2x RSP	2x RSP	6+1	6+1	6+1
Line cards & ports	4x SFP+ 2x MPA	2	4	8	8	10	20
Rack units	2	10U 2-post 23"	10	21	21	30	44
Power modules	2x AC or 2x DC	4x AC or 4x DC	4x AC or 4x DC	8x AC or 8x DC	8x AC or 8x DC	12x AC or 12x DC	16x AC or 16x DC
Air flow	Right to left	Front to back	Right to back	Front to back	Front to back	Front to back	Front to back

System Health: Inventory

components and serial numbers

- Subscribe to Cisco notifications
  - www.cisco.com/cisco/support/notifications.html
- Field notices?
  - www.cisco.com/en/US/partner/support/tsd_products_field_notice_summary.html
- Test and approve
- Track and audit

```
RP/0/RSP0/CPU0:rasr9906-1y#admin show platform
Sun Jan  1 10:05:42.505 PST
Location  Card Type                 HW State      SW State      Config State
------------------------------------------------------------------------------
0/0       A99-8X100GE-CM            OPERATIONAL   OPERATIONAL   NSHUT
0/1       A99-48X10GE-1G-SE         OPERATIONAL   OPERATIONAL   NSHUT
0/2       A99-10X400GE-X-SE         OPERATIONAL   OPERATIONAL   NSHUT
0/RSP0    A9K-RSP5-SE               OPERATIONAL   OPERATIONAL   NSHUT
0/RSP1    A9K-RSP5-SE               OPERATIONAL   OPERATIONAL   NSHUT
0/FC0     A99-SFC3-T                OPERATIONAL   N/A           NSHUT
0/FC1     A99-SFC3-T                OPERATIONAL   N/A           NSHUT
0/FC2     A99-SFC3-T                OPERATIONAL   N/A           NSHUT
0/FC3     A99-SFC3-T                OPERATIONAL   N/A           NSHUT
0/FC4     A99-SFC3-T                OPERATIONAL   N/A           NSHUT
0/FT0     ASR-9906-FAN              OPERATIONAL   N/A           NSHUT
0/FT1     ASR-9906-FAN              OPERATIONAL   N/A           NSHUT
0/PT0     A9K-AC-PEM-V3             OPERATIONAL   N/A           NSHUT
RP/0/RSP0/CPU0:rasr9906-1y#admin show inventory
Sun Jan  1 10:08:28.025 PST
 Name: Rack 0               Descr: ASR 9906 4 Line Card Slot Chassis
 PID: ASR-9906              VID: V01                    SN: FOX2247P3QM
 Name: 0/0                  Descr: ASR 9000 8-port 100GE Consumption Model Line Card
 PID: A99-8X100GE-CM        VID: V02                    SN: FOC2045N384
 Name: 0/1                  Descr: ASR 9000 48-port 10GE & 1GE dual rate Service Edge LC
 PID: A99-48X10GE-1G-SE     VID: V01                    SN: FOC2335NLX3
 Name: 0/2                  Descr: ASR 9900 10-port 400GE X SE linecard
 PID: A99-10X400GE-X-SE     VID: V01                    SN: FOC2512NES0
 Name: 0/RSP0               Descr: ASR 9000 Route Switch Processor 5 for Service Edge 40G
 PID: A9K-RSP5-SE           VID: V01                    SN: FOC2248NDVF
 Name: 0/RSP1               Descr: ASR 9000 Route Switch Processor 5 for Service Edge 40G
 PID: A9K-RSP5-SE           VID: V01                    SN: FOC2246NLGE
 Name: 0/FC0                Descr: ASR 9906 Switch Fabric Card 3
 PID: A99-SFC3-T            VID: V01                    SN: FOC2245N5UN
```

64Bit eXR

System Health: Power

Installed power, status, consumption, and redundancy

```
RP/0/RSP0/CPU0:rasr9906-1y#admin show environment power
Sun Jan  1 10:20:37.742 PST
==============================================================================
CHASSIS LEVEL POWER INFO: 0
==============================================================================
   Total output power capacity (N + 1)             :    6000W +       0W
   Total output power required                     :    4865W
   Total power input                               :    3327W
   Total power output                              :    3118W

Power Shelf 0:
==============================================================================
   Power       Supply      --------Input-------   ----Output----       Status
   Module      Type       Volts A/B   Amps A/B   Volts     Amps
==============================================================================
   0/PT0-PM0   6kW-AC       0.0/0.0     0.0/0.0     0.0      0.0    FAILED or NO PWR
   0/PT0-PM1   6kW-AC     209.4/0.0     7.9/0.0    53.7     29.0    OK
   0/PT0-PM2   6kW-AC     209.1/0.0     8.0/0.0    54.0     28.9    OK

Total of Power Shelf 0:      3327W/  (15.9/ 0.0)A    3118W/ 57.9A

==============================================================================
   Location     Card Type            Power       Power       Status
                                     Allocated   Used
                                     Watts       Watts
==============================================================================
   0/0          A99-8X100GE-CM         915         778       ON
   0/1          A99-48X10GE-1G-SE      470         336       ON
   0/2          A99-10X400GE-X-SE     1370         852       ON
   0/3                -                 10           -       RESERVED
```

```
0/RSP0      A9K-RSP5-SE      480      256      ON
0/RSP1      A9K-RSP5-SE      480      252      ON
0/FC0       A99-SFC3-T       108       47      ON
0/FC1       A99-SFC3-T       108       48      ON
0/FC2       A99-SFC3-T       108       48      ON
0/FC3       A99-SFC3-T       108       48      ON
0/FC4       A99-SFC3-T       108       48      ON
0/FT0       ASR-9906-FAN     300        -      ON
0/FT1       ASR-9906-FAN     300        -      ON
```

- Use Telemetry or SNMP to check on power population and status
- Use Power Calculator http://tools.cisco.com/cpc/launch.jsp . If changing chassis power version/type [v2/v1, AC/DC]: http://www.cisco.com/en/US/docs/routers/asr9000/hardware/installation/guide/asr9kIGmaintaining.html#wp1323197

System Health: Heat Dissipation & Alarms

Fans, temperature sensors, and alarms

```
RP/0/RSP0/CPU0:rasr9906-1y# admin show environment temperatures
Sun Jan  1 10:33:08.316 PST
================================================================================
Location  TEMPERATURE                 Value   Crit Major Minor Minor Major  Crit
          Sensor                    (deg C)   (Lo) (Lo)  (Lo)  (Hi)  (Hi)  (Hi)
--------------------------------------------------------------------------------
0/RSP0
          AIR_Outlet                   33    -10    -5    0     80    85    100
          Inlet                        24    -10    -5    0     70    85    100
          Hotspot                      38    -10    -5    0     90    93     95
0/RSP1
          AIR_Outlet                   34    -10    -5    0     80    85    100
          Inlet                        26    -10    -5    0     70    85    100
          Hotspot                      36    -10    -5    0     90    93     95
0/FC1
          SKB0_HOTSPOT                 37    -10    -5    0     80    83     85
          Inlet                        22    -10    -5    0     60    65     80
          DIE_FabSwitch0               52    -10    -5    0    115   125    140
0/FC2
          SKB0_HOTSPOT                 38    -10    -5    0     80    83     85
          Inlet                        22    -10    -5    0     60    65     80
          DIE_FabSwitch0               52    -10    -5    0    115   125    140
...
0/FT0
          Inlet                        27    -10    -5    0    105   115    120
          Hotspot                      28    -10    -5    0    105   115    120
0/FT1
          Inlet                        30    -10    -5    0    105   115    120
          Hotspot                      31    -10    -5    0    105   115    120
0/PT0-PM0
          PM0-Inlet Temperature         -    -10    -5    0     61    65     70
          PM0-Outlet Temperature        -    -10    -5    0     80    92    105
          PM0-Heat Sink Temperature     -    -10    -5    0    105   112    120
0/PT0-PM1
          PM1-Inlet Temperature        19    -10    -5    0     61    65     70
          PM1-Outlet Temperature       43    -10    -5    0     80    92    105
          PM1-Heat Sink Temperature    56    -10    -5    0    105   112    120
```

```
RP/0/RSP0/CPU0:rasr9906-1y# admin show environment fan
Sun Jan  1 10:50:53.013 PST
================================================================================
                          Fan speed (rpm)
Location      FRU Type          FAN_0  FAN_1  FAN_2  FAN_3  FAN_4  FAN_5  FAN_6
--------------------------------------------------------------------------------
0/FT0         ASR-9906-FAN       7974   7976   7984   7978   7978   7905   7980
0/FT1         ASR-9906-FAN       8032   7995   8059   8010   7970   7977   7974

RP/0/RSP0/CPU0:rasr9906-1y# admin show environment leds
RP/0/RSP0/CPU0:rasr9906-1y# admin show led
Sun Jan  1 11:08:32.125 PST
================================================================
Location  LED Name                          Mode          Color
================================================================
0/0
          0/0-Status LED                    WORKING       GREEN
0/1
          0/1-Status LED                    WORKING       GREEN
0/2
          0/2-Status LED                    WORKING       GREEN
0/RSP0
          0/RSP0-Fail LED                   WORKING       OFF
          0/RSP0-ACO LED                    WORKING       OFF
          0/RSP0-Alarm Minor LED            WORKING       OFF
          0/RSP0-Alarm Major LED            WORKING       RED
          0/RSP0-Alarm Critical LED         WORKING       OFF
          0/RSP0-FC Fault LED               WORKING       OFF
```

RP front status LED

ASR 9000 System & Switch Fabric Overview

	SFC2	SFC3
Fabric Capacity per SFC	215G	600G
Fabric Capacity Per Line Card Slot	1.29T N+1 1.51T N+0	3.6T N+1 4.2T N+0
Fabric Redundancy	N+1	N+1
LC Support	Typhoon Tomahawk	Tomahawk LightSpeed LightSpeed Plus

Integrated Fabric on RSP

ASR 9904 ASR 9006 ASR 9010

Hybrid Systems

ASR 9906 ASR 9910

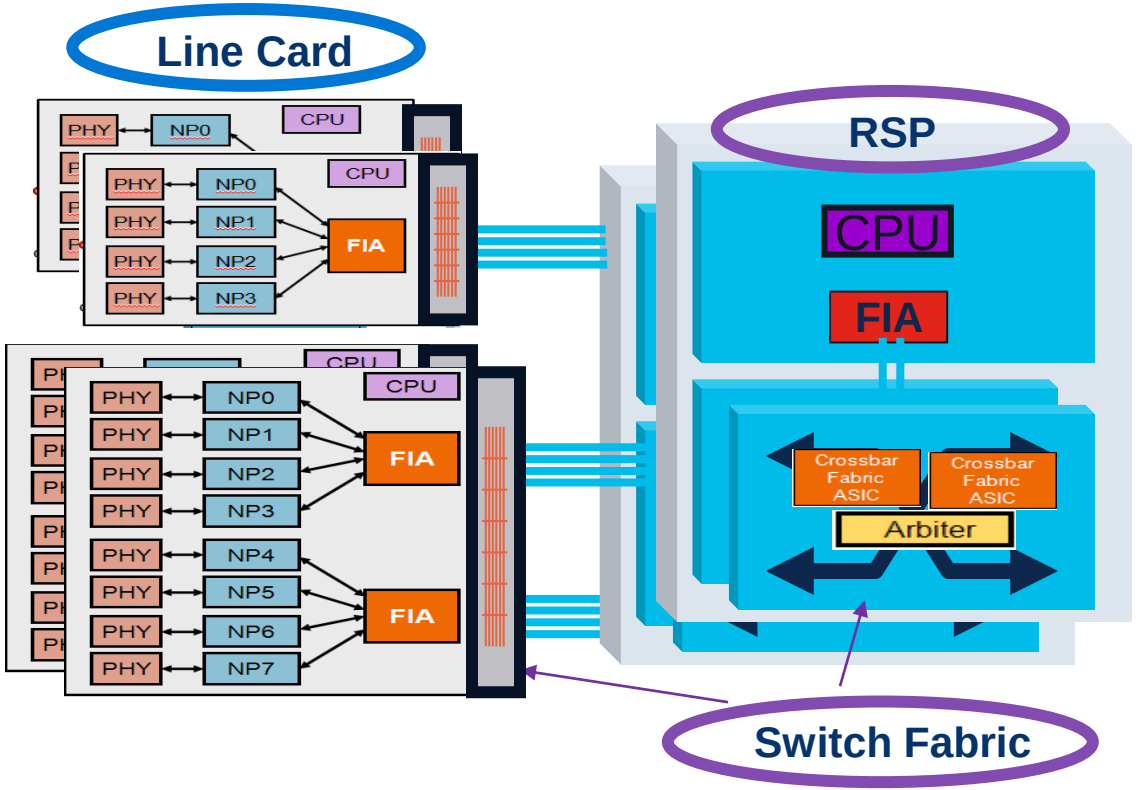
Dedicated Fabric Cards

ASR 9912 ASR 9922

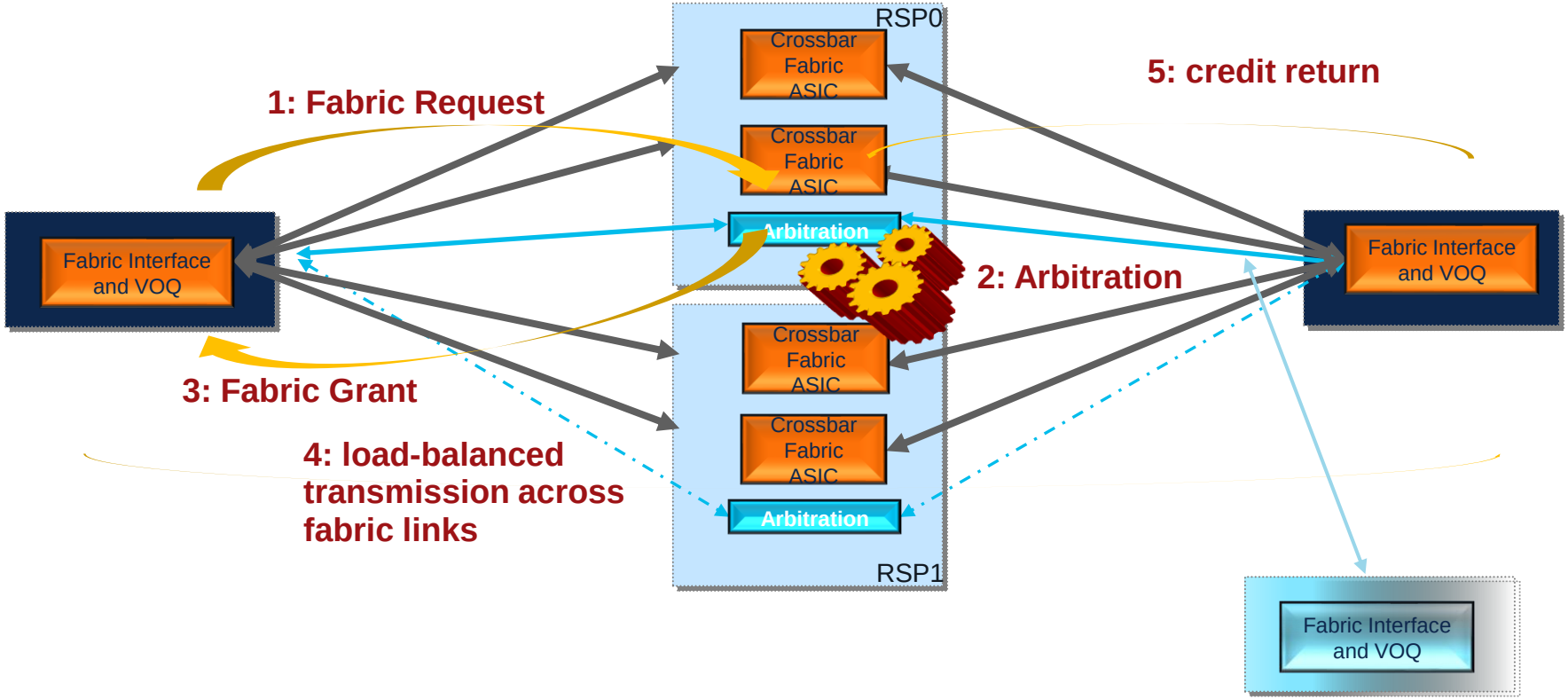
System Architecture

9904, 9006, 9010

- Distributed control plane
  - L2 protocols, ARP, BFD, CFM, Netflow run on LC CPU
- Distributed data plane
  - Forwarding distributed to NP's
- Active-active switch fabric
  - Each RSP houses "half" of the fabric



RP/RSP Fabric Arbitration Diagram



RSP880

Third generation RP and fabric

~~1G/10G SFP+~~
~~FOBC ports for nV Cluster~~

USB Type A

LEDs
Status, Alarm

Aux

IEEE 1588, GPS
SyncE, IEEE1588 master and slave
10/100M Copper Ethernet

Console

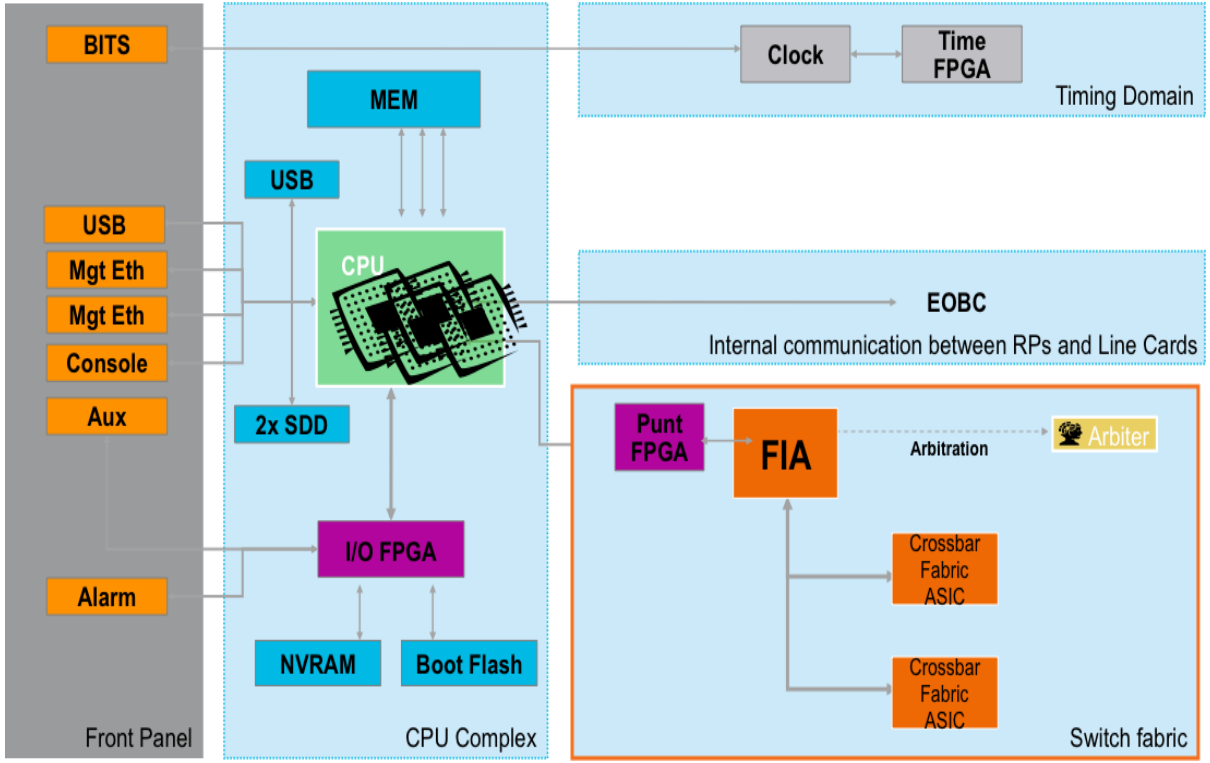
BITS/J.211
Sync 0, Sync 1
RJ45

Management
Ethernet

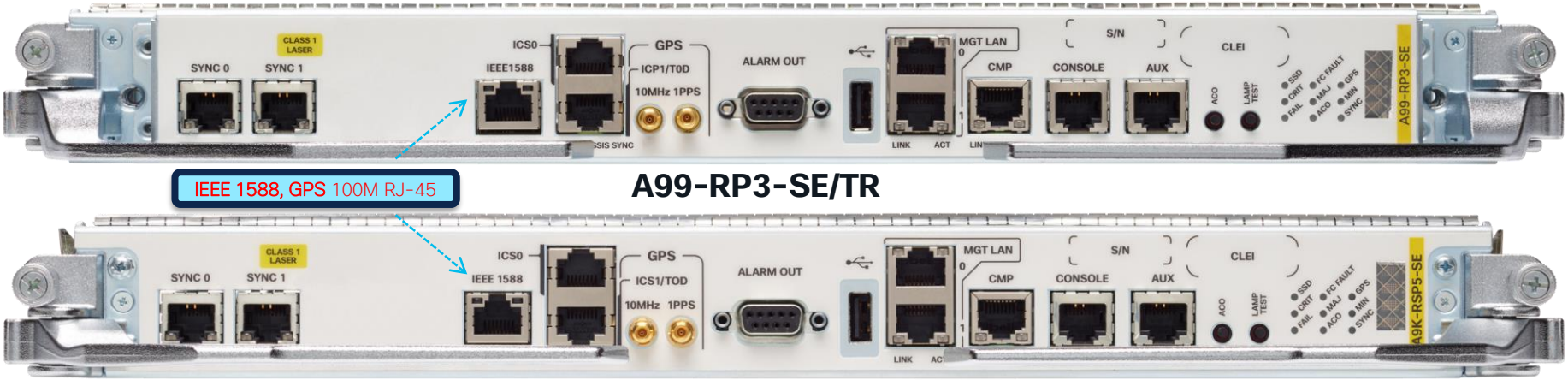
RSP880 Architecture

The RP and fabric

- Intel x86 8-core proc. @2 GHz
- SE and TR versions
  - SE: 32 GB RAM
  - TR: 16 GB RAM
- Storage
  - 2x 32 GB SSD
  - 1x 8 GB USB
- EOBC on panel
  - 4 SFP+ EOBC ports for clustering
- Fabric on each RSP
  - 440 Gbps per slot per RSP
  - 0 packet loss switchover



RSP5/RP3 Front Panel

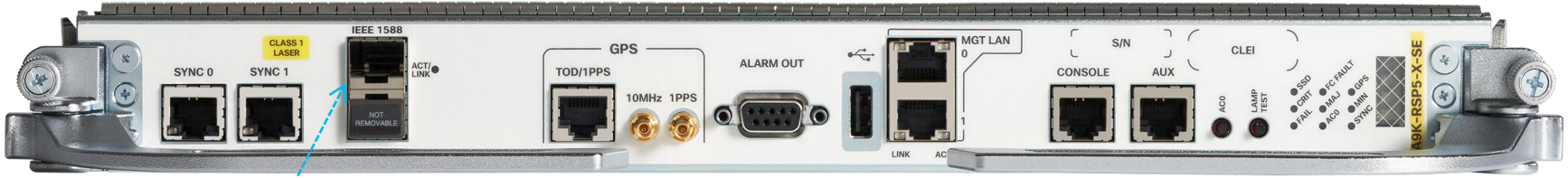


A9K-RSP5-SE/TR

- RAM: -TR 24G; -SE 40G
- 2x BITS ports on RJ-45
- 100Mbps, 1588 port – RJ-45
- TOD – RJ-45
- 10Mhz on SMA
- 1PPS on SMA

- Alarm output serial port
- 1x USB, 1x CMP
- 2x Management ports on RJ-45
- AUX & Console on RJ-45 connectors
- LED's for major/critical and normal oper alarms or states

RSP5-X/RP3-X Front Panel



1588: GPS changed from RJ45 (RSP5/RP3) to 10/1G optical port (RSP5-X/RP3-X)

A9K-RSP5-X-SE/TR

A99-RP3-X-SE/TR

- **New PIDs:** A9K-RSP5-X-TR / SE and A99-RP3-X-TR /SE (-TR RAM 24G; -SE RAM 48G)
- **Supported Release:** IOS-XR 64-bit 7.6.2 (August 2022)
- **RSP5-X & RP3-X brings Class C timing compliant to ASR 9K Modular Chassis**

  5Th Generation modular Line card already supports Class C timing

- Use case for 5G applications (Mobile-Back Haul & 5G transport) requiring Class C timing

Route Switch Processors and Route Processors

	RSP5-X (IOS-XR 64Bit)	RSP5 (IOS-XR 64Bit)	RSP880-LT (IOS-XR 32/64Bit)
Description	RP and Fabric ASIC	RP and Fabric ASIC	RP and Fabric ASIC
Switch Fabric Bandwidth	1.2T + 1.2T (9006/9010) 1.8T + 1.8T (9904) 600G + 600G + 3.0T (9906/9910)*	1.2T + 1.2T (9006/9010) 1.8T + 1.8T (9904) 600G + 600G + 3.0T (9906/9910)*	400G + 400G (9006/9010) 700G + 700G (9904) 200G per RSP/SFC (9906/9910)
Processor	Intel (Skylake EP) 8 Cores 2GHz	Intel Skylake EP 8 Cores 2GHz	Intel (Ivy Bridge EP) 4 Cores, 2.4GHz
RAM	-TR: 24GB -SE: **48GB**	-TR: 24GB -SE: 40GB	-TR: 16GB -SE: 32GB
SSD	2 x 128GB SSD	2 x 128GB SSD	2 X 128GB SSD
Punt BW	40GE	40GE	40GE
Timing	LS+: **Class C** Available at XR762	Class B	Class B
Platforms Supported	ASR 9904, 9006, 9010, 9910, 9906	ASR 9904, 9006, 9010, 9910, 9906	ASR 9904, 9006, 9010, 9910, 9906
3rd Party App Support	Supported	Supported	Not Supported

*For line card slots equipped with Tomahawk line cards, switch fabric bandwidth is equal to RSP880-LT
RSP used in ASR9910/9906/9904/9006/9010, RP in ASR9922/9912

Route Switch Processors and Route Processors

RSP used in ASR9910/9906/9904/9006/9010, RP in ASR9922/9912

	RSP880 A99-RSP	RP2	RSP5	RP3
Description	3rd Gen RP and Fabric ASIC		4th Gen RP and Fabric ASIC	
Switch Fabric Bandwidth	400G + 400G (9006/9010) 700G + 700G (9904) 215G + 215G + 1.07T (9906/9910)	1.2Tb + 215G (separated fabric card)	900G + 900G (9006/9010) 1.8T + 1.8T (9904) 600G + 600G + 3.0T (9906/9910)	3.6Tb + 600G (separated fabric card)
Processor	Intel x86 (Ivy Bridge EP) 8 Core 2GHz		Intel x86 (Skylake EP) 8 Core 2GHz	
RAM	-TR: 16GB -SE: 32GB		-TR: 16GB -SE: 40GB	
SSD	2 x 32GB Slim SATA		2 x 128GB Slim SATA	
Punt BW	40GE		40GE	

RP

Processor and storage: RSP5-SE

```
RP/0/RSP0/CPU0:rasr9906-1y# show version
Sun Jan  1 13:27:47.999 PST
Cisco IOS XR Software, Version 7.7.2
Copyright (c) 2013-2022 by Cisco Systems, Inc.

Build Information:
 Built By     : ingunawa
 Built On     : Wed Oct 26 12:57:07 PDT 2022
 Built Host   : iox-ucs-057
 Workspace    : /auto/srcarchive14/prod/7.7.2/asr9k-x64/ws
 Version      : 7.7.2
 Location     : /opt/cisco/XR/packages/
 Label        : 7.7.2

cisco ASR9K () processor
System uptime is 5 weeks 15 hours 17 minutes
```

```
RP/0/RSP0/CPU0:rasr9906-1y# show filesystem
Sun Jan  1 13:28:11.786 PST
File Systems:

      Size(b)       Free(b)      Type  Flags  Prefixes
 4060278784    4036325376 flash-disk     rw  disk0:
20507914240   20433276928 flash-disk     rw  apphost:
           0             0    network     rw  tftp:
           0             0    network     rw  ftp:
  1015304192    1011585024      flash     rw  /misc/config
23653642240   18958163968   harddisk     rw  harddisk:
 3921260544    3913465856   harddisk     rw  harddiska:
 3921260544    3913465856   harddisk     rw  harddiskb:
```

USB

SSD

RP

Redundancy

```
RP/0/RSP0/CPU0:rasr9906-1y# show redundancy
Sun Jan  1 13:26:57.892 PST
Redundancy information for node 0/RSP0/CPU0:
==========================================
Node 0/RSP0/CPU0 is in ACTIVE role
Partner node (0/RSP1/CPU0) is in STANDBY role
Standby node in 0/RSP1/CPU0 is ready
Standby node in 0/RSP1/CPU0 is NSR-ready

Reload and boot info
----------------------
A9K-RSP5-32G reloaded Tue Nov 29 22:10:31 2022: 5 weeks, 15 hours, 16 minutes ago
Active node booted Tue Nov 29 22:10:31 2022: 5 weeks, 15 hours, 16 minutes ago
Standby node boot Tue Nov 29 22:10:39 2022: 5 weeks, 15 hours, 16 minutes ago
Standby node last went not ready Fri Dec 23 13:05:07 2022: 1 week, 5 days, 21 minutes ago
Standby node last went ready Fri Dec 23 13:05:07 2022: 1 week, 5 days, 21 minutes ago
Standby node last went not NSR-ready Fri Dec 23 12:03:54 2022: 1 week, 5 days, 1 hour, 23 minutes ago
Standby node last went NSR-ready Fri Dec 23 12:06:15 2022: 1 week, 5 days, 1 hour, 20 minutes ago
There have been 0 switch-overs since reload

Active node reload "CARD_SHUTDOWN"
Standby node reload "CARD_SHUTDOWN "
```

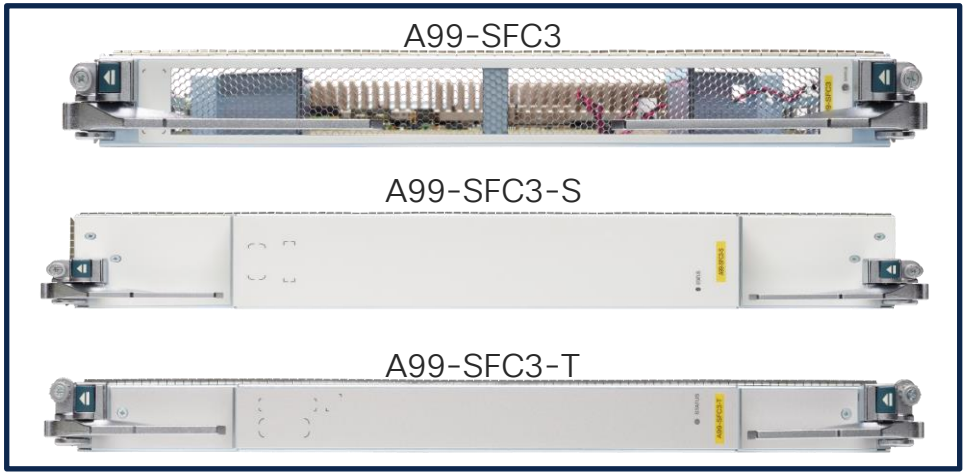
RP

Redundancy failover example: RSP0 removed

```
RP/0/RSP1/CPU0:Dec 16 20:57:47.049 : ce_switch_srv[53]: %PLATFORM-CE_SWITCH-6-UPDN : Interface 8
(Peer_RSP) is down
RP/0/RSP1/CPU0:Dec 16 20:57:47.091 : sc_reddrv[392]: %PLATFORM-REDDRV-6-RESET_STANDBY : Active
is going to reset standby peer node node0_RSP0_CPU0
RP/0/RSP1/CPU0:Dec 16 20:57:47.093 : rmf_svr[386]: %HA-REDCON-6-GO_ACTIVE : this card going
active
RP/0/RSP1/CPU0:Dec 16 20:57:47.094 : sysmgr[93]: %OS-SYSMGR-5-NOTICE : This standby node is
going active at Mon Dec 16 20:57:47 2013
RP/0/RSP1/CPU0:Dec 16 20:57:47.400 : dao_tmp[56]: Card Removed 0/RSP0/CPU0
RP/0/RSP1/CPU0:Dec 16 20:57:47.449 : sysmgr[93]: %OS-SYSMGR-5-NOTICE : This node is active now
at Mon Dec 16 20:57:47 2013
RP/0/RSP1/CPU0:Dec 16 20:57:47.450 : sysmgr[93]: %OS-SYSMGR-5-NOTICE : Critical failover elapsed
time 0.353 seconds (86.080% idle)
RP/0/RSP1/CPU0:Dec 16 20:57:47.457 : isis[1003]: %ROUTING-ISIS-6-INFO_STARTUP_START : Cold
controlled start beginning
RP/0/RSP1/CPU0:Dec 16 20:57:47.468 : ospf[1011]: %ROUTING-OSPF-5-HA_NOTICE : Process 100:
Attempting Cisco NSF-enabled restart of vrf default
RP/0/RSP1/CPU0:Dec 16 20:57:47.470 : ospf[1011]: %ROUTING-OSPF-5-HA_NOTICE : Process 100:
Signaled PROC_AVAILABLE
RP/0/RSP1/CPU0:Dec 16 20:57:47.784 : mpls_ldp[1043]: %ROUTING-LDP-3-ERR_LPTS_FILT : Failed to
remove LPTS filter for local 10.101.111.1 remote=10.101.112.1: Host is down
.
```

ASR 9000 5th Gen Fabric PIDs

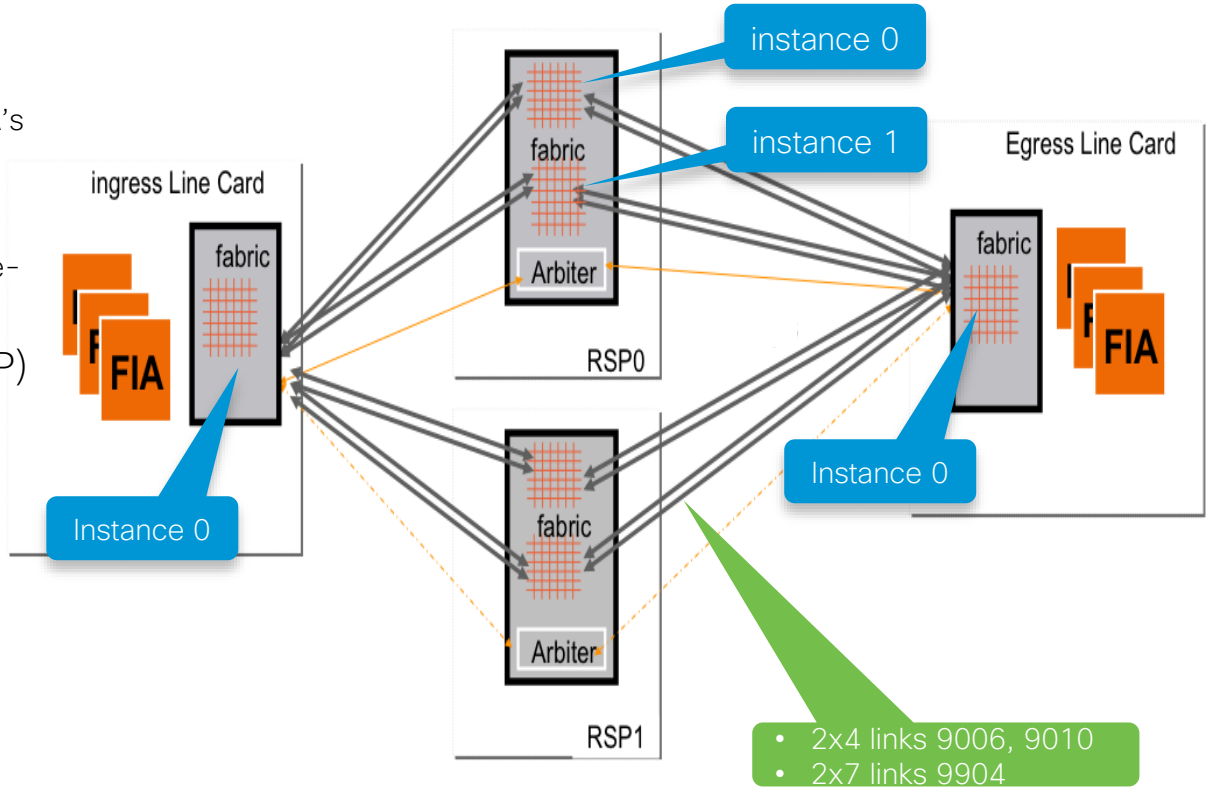
- *ASR 9922*
  - A99-RP3-SE/TR
  - A99-SFC3
  - ASR-9922-FAN-V3
- *ASR 9912*
  - A99-RP3-SE/TR
  - A99-SFC3
- *ASR 9910*
  - A9K-RSP5-SE/TR
  - A99-SFC3-S
- *ASR 9906*
  - A9K-RSP5-SE/TR
  - A99-SFC3-T
- *ASR 9904 / ASR 9010 / ASR 9006*
  - A9K-RSP5-SE/TR



The Integrated Switch Fabric

ASR 9010, 9006, 9904

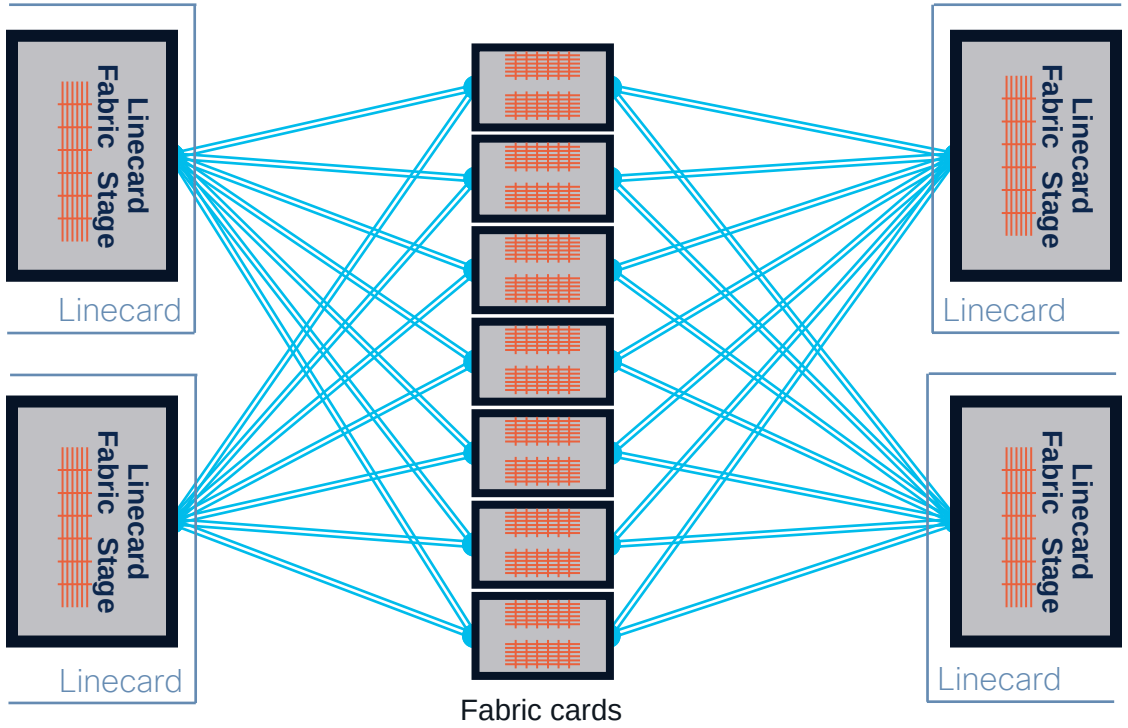
- 3-Stage fabric
  - Allows for variable number of FIA's and FIA links on LC
- Super-framing for unicast
  - Super-frame same-priority same-egress frames in a jumbo frame
- Bandwidth per slot (dual RSP)
  - RSP880: 880G
    - asr9904: 1.5T
  - RSP5: 1.8T
    - asr9904: 3.6T
- Fabric load sharing
  - Unicast: per super-frame
  - Multicast: per flow



The Dedicated Switch Fabric

ASR 9922, 9912

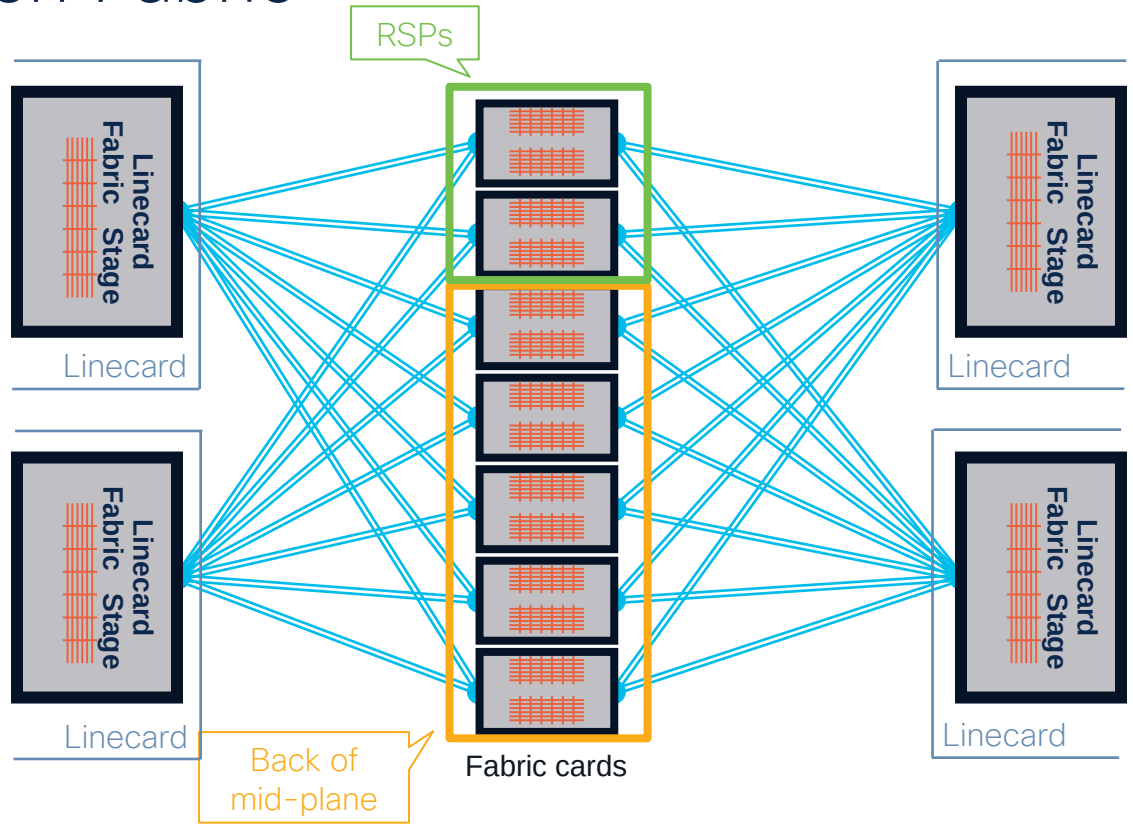
- 3-Stage fabric
  - Allows for variable number of FIA's and FIA links on LC
- Super-framing for unicast
  - Super-frame same-priority same-egress frames in a jumbo frame
- Bandwidth per slot:
  - SFC1: 110G [Typhoon gen]
  - SFC2: 215G [Tomahawk gen]
  - SFC3: 600G [Lightspeed gen]
- Fabric load sharing
  - Unicast: per super-frame
  - Multicast: per flow



The Hybrid Switch Fabric

ASR 9910, 9906

- 3-Stage fabric
  - Allows for variable number of FIA's and FIA links on LC
- Super-framing for unicast
  - Super-frame same-priority same-egress frames in a jumbo frame
- Bandwidth per slot:
  - SFC1: 110G [Typhoon gen]
  - SFC2: 215G [Tomahawk gen]
  - SFC3: 600G [Lightspeed gen]
- Fabric load sharing
  - Unicast: per super-frame
  - Multicast: per flow



Slot Address

Internal slot address, slot mask, and fabric group ID

Slot		Slot Mask	
Logical	Physical	Binary	Hex
LC19	21	10 0000 0000 0000 0000 0000	0x10 0000
LC4-18	6-20		
LC3	5	00 0010 0000	0x0020
LC2	4	00 0001 0000	0x0010
LC1	3	00 0000 1000	0x0008
LC0	2	00 0000 0100	0x0004
RSP1	1	00 0000 0010	0x0002
RSP0	0	00 0000 0001	0x0001

- Follows the sequence of slots in chassis
- 4/6/12/22 RP slots: 000011 which is 0x0003, decimal 3
- Exception: 9010 RSP slots: 0000110000 which is 0x0030, decimal 48

Slot Address

Internal slot address, slot mask, and fabric group ID

Location	Slot		Slot Mask	
	Logical	Physical	Binary	Hex
0/19/CPU0	LC19	21	10 0000 0000 0000 0000 0000	0x10 0000
0/4/CPU0 – 0/18/CPU0	LC4-18	6-20		
0/3/CPU0	LC3	5	00 0010 0000	0x0020
0/2/CPU0	LC2	4	00 0001 0000	0x0010
0/1/CPU0	LC1	3	00 0000 1000	0x0008
0/0/CPU0	LC0	2	00 0000 0100	0x0004
0/RSP1/CPU0	RSP1/RP1	1	00 0000 0010	0x0002
0/RSP0/CPU0	RSP0/RP0	0	00 0000 0001	0x0001

- Follows the sequence of slots in chassis
- 4/6/12/22 RP slots: 000011 which is 0x0003, decimal 3
- Exception: 9010 RSP slots: 0000110000 which is 0x0030, decimal 48

Slot Address

Internal slot address, slot mask, and fabric group ID

Slot		Slot Mask	
Logical	Physical	Binary	Hex
LC7	9	1000000000	0x0200
LC6	8	0100000000	0x0100
LC5	7	0010000000	0x0080
LC4	6	0001000000	0x0040
RSP0	5	0000100000	0x0020
RSP1	4	0000010000	0x0010
LC3	3	0000001000	0x0008
LC2	2	0000000100	0x0004
LC1	1	0000000010	0x0002
LC0	0	0000000001	0x0001

9906

9010

9910/12/22

Slot		Slot Mask	
Logical	Physical	Binary	Hex
LC3	5	0000100000	0x0020
LC2	4	0000010000	0x0010
LC1	3	0000001000	0x0008
LC0	2	0000000100	0x0004
RSP1	1	0000000010	0x0002
RSP0	0	0000000001	0x0001

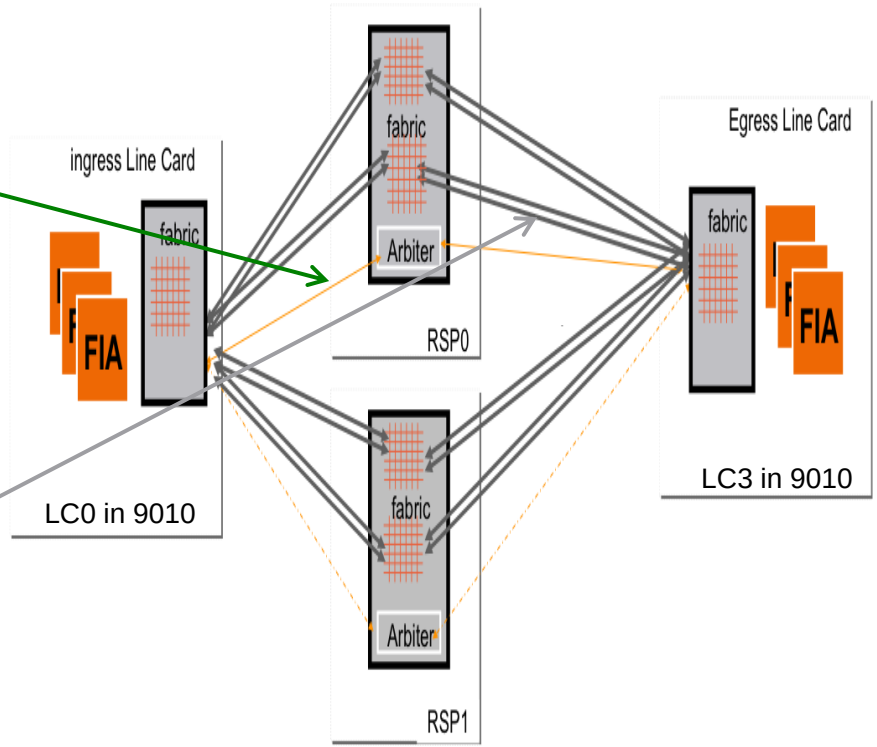
Slot		Slot Mask	
Logical	Physical	Binary	Hex
LC19	21	10000 00000000 00000000	0x10 0000
LC1-18	3-20		
LC0	2	0000000100	0x0004
RP1	1	0000000010	0x0002
RP0	0	0000000001	0x0001

- Follows the sequence of slots in chassis
- 4/6/12/22 RP slots: 000011 which is 0x0003, decimal 3
- 9010 RSP slots: 0000110000 which is 0x0030, decimal 48

Fabric Channels

Link status

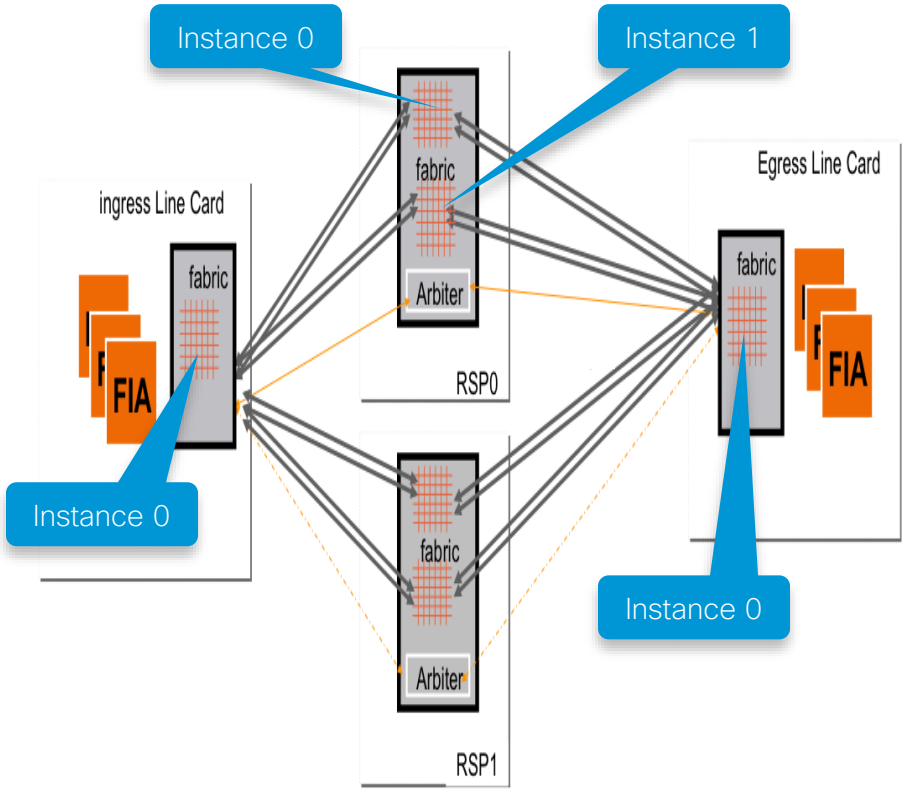
```
RP/0/RSP0/CPU0:rasr9000-2w-a#show controllers fabric
arbiter link-status location 0/RSP0/CPU0
Port     Remote Slot   Remote Elem     Remote Inst   Status
=========================================================
08       0/1/CPU0            ARB             0         Up
12       0/2/CPU0            ARB             0         Up
14       0/0/CPU0            ARB             0         Up
20       0/RSP0/CPU0         FIA             0         Up
21       0/RSP1/CPU0         FIA             0         Up


RP/0/RSP0/CPU0:rasr9000-2w-a#show controllers fabric
crossbar link-status instance 0 location 0/RSP0/CPU0
Tue Dec 17 02:59:30.110 EST
PORT     Remote Slot   Remote Inst     Logical ID   Status
=========================================================
02         0/3/CPU0          00              0         Up
16         0/0/CPU0          00              0         Up
18         0/2/CPU0          00              1         Up
20         0/2/CPU0          00              0         Up
24         0/3/CPU0          00              1         Up
```



Fabric Statistics

Link statistics

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show controllers fabric
crossbar statistics instance 1 location 0/RSP0/CPU0
Tue Dec 17 02:59:36.376 EST
Port statistics for xbar:1 port:2
==============================
Hi priority stats (unicast)
==========================
    Ingress Packet Count Since Last Read       : 17347357500
    Ingress Channel Utilization Count          : 5
    Output Buffer Queued Packet Count          : 1
    Egress Packet Count Since Last Read        : 19006087016
    Egress Channel Utilization Count           : 4
.

Port statistics for xbar:1 port:16
==============================
Hi priority stats (unicast)
==========================
    Ingress Packet Count Since Last Read       : 49365
    Egress Packet Count Since Last Read        : 323

Low priority stats (multicast)
==========================
    Ingress Packet Count Since Last Read       : 1623
    Egress Packet Count Since Last Read        : 716

.
Total Unicast In:    63038489128
Total Unicast Out:   63038489275
Total Multicast In:  1625
Total Multicast Out: 1252
```



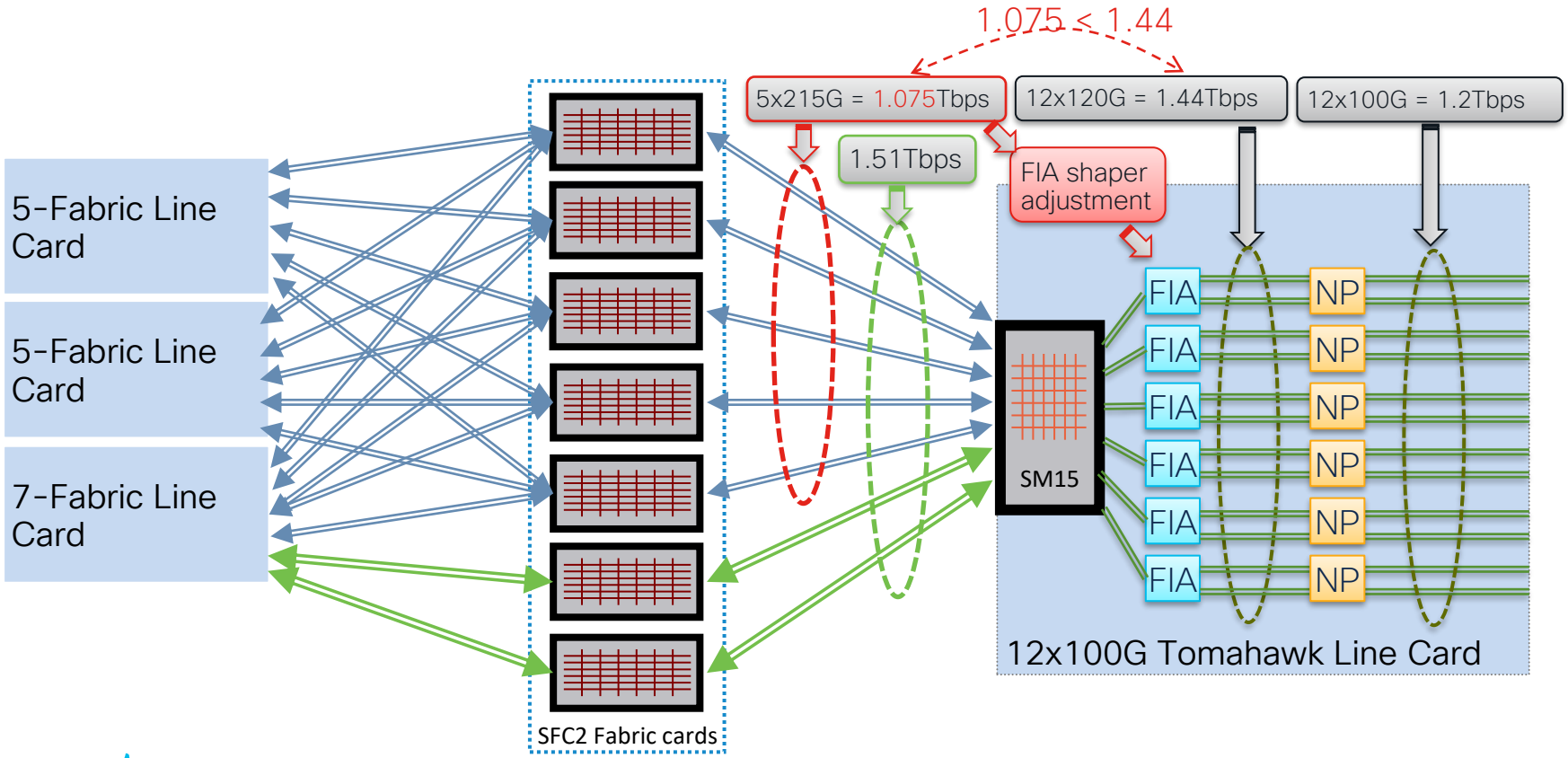
Fabric Mode

High bandwidth, or mixed

- Default:
  - Max 1024 VQI per system
  - Multicast traffic uses the first 5 fabric cards
- highbandwidth:
  - Max 2048 VQI per system (➔ only 3rd Generation and later allowed)
  - Multicast traffic uses the first 5 fabric cards
- A99-highbandwidth:
  - Max 2048 VQI per system
  - Multicast traffic uses all 7 fabric cards (➔ only A99 3rd Generation and later allowed)

```
RP/0/RSP1/CPU0:ASR9K-2(admin-config)#fabric enable mode ?
  A99-highbandwidth  A99 High bandwidth cards only
  highbandwidth      High bandwidth cards only
```

ASR 9000 3rd Gen 5 & 7-Fabric LC Interworking

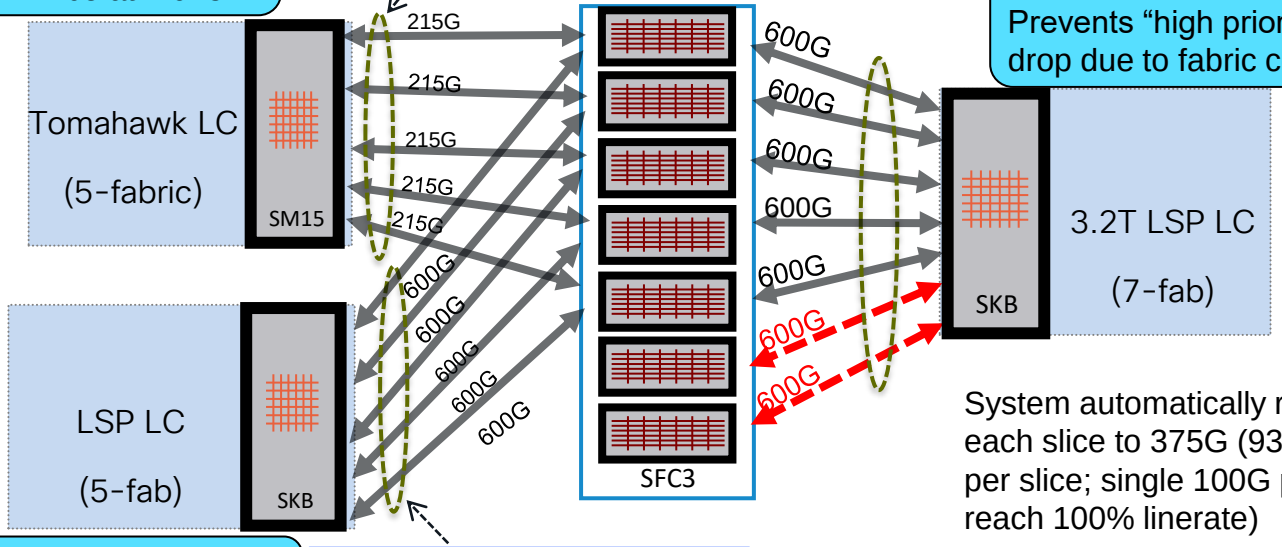


ASR 9000 Mixed 5 & 7-Fabric LC Interworking

- 7-fabric LC uses 5 fabrics if there is a 5-fab LightSpeed Plus or Tomahawk LC in system

All Tomahawk LCs interoperate at full throughput w/ fabric redundancy; except for 12x100GE LC, which will be at 1.075T

1.075T with 5 fabrics
860G w/ 4 fabrics



3T with 5-fabrics
Prevents "high priority" packet drop due to fabric congestion

System automatically rate-limits each slice to 375G (93.75% linerate per slice; single 100G port can reach 100% linerate)

For 100G linerate on all ports:

- Shut down one slice

2.0T LC at line-rate w/ fabric redundancy
800G LC at line-rate w/ fabric redundancy

3T w/ 5 fabrics
2.4T w/ 4 fabrics

Fabric Interworking: 5-Fab LC to 12x100G LC

- FIA shaper is applied by default on 12x100G line cards
- A99 chassis with 5 fabric cards or more:
  - 83Gbps per 100G port (total of 996 Gbps; fabric conn 5x230Gbps = 1.15Tbps Gbps)
- Any chassis with 4 fabric cards (asr9010, asr9006 with dual RSP880):
  - 71Gbps per 100G port (total of 852 Gbps; fabric connection 4x230Gbps = 920 Gbps)
- Syslog:
  - `LC/0/0/CPU0:Dec 27 12:05:16.429 EST: pfm_node_lc[299]: %FABRIC-FIA-1-RATE_LIMITER_ON : Set|fialc[163907]|0x1072000|Insufficient fabric capacity for card types in use - FIA egress rate limiter applied`
- Checking the shaper rate:
  - `show controllers fabric fia information location <location>`
  - `show controller fabric fia trace location <location> | include “shape_RL"`

ASR 9000 HW Offer: 3rd Generation LC Portfolio

200G to 1.2T per slot

A99-12X100GE

A9K-4X100GE

A9K-8X100GE-TR/SE

A9K-4X100GE-TR/SE

A9K-48X10GE-1G-TR/SE

A9K-24X10GE-1G-TR/SE

A9K-MOD400-TR/SE

A9K-MOD200-TR/SE

Ports	Bandwidth	Mode	MACSec	
12 Ports of QSFP28 4 Ports of QSFP28	1.2 Tbps 400 Gbps	LAN	No	Class B
8 Ports CPAK 100GE 4 Ports CPAK 100GE	800 Gbps 400 Gbps	LAN, OTN, WAN	Yes	Class B
48 Ports SFP+ 24 Ports SFP+	400 Gbps 200 Gbps	LAN	No	Class B
Mixed Speed based on the MPA	400 Gbps 200 Gbps	LAN, OTN, WAN	Yes*	Class B

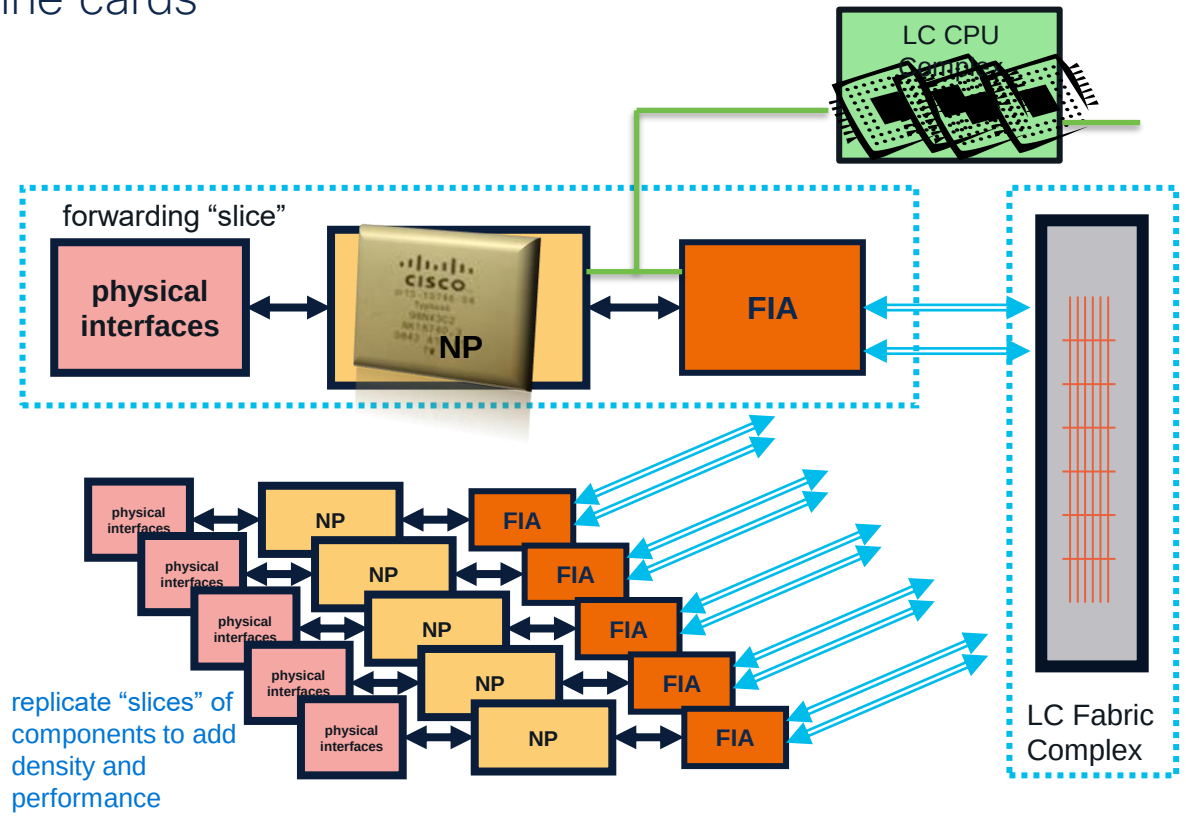
*Depending on MPAs

cisco Live!

3rd Generation Line Card Architecture

3rd generation Ethernet line cards

- Physical 1/10/40/100 Gbps
  - No frame processing there
- Tomahawk NP
  - FIB, MAC, ACL, QoS, encap/decap, LPTS, all hw features, buffer, i/f stats
  - 240 Gbps 150 Mpps bidirectional
  - Or 480 Gbps unidirectional
- Fabric Interface ASIC
  - 2PQ+1BE into fabric, VOQ
  - System priority queueing
  - Separate unicast and multicast queueing
  - Super-framing and buffering



ASR 9000 3rd Gen TomaHawk 100G Options

	4x100GE Octane 8x100GE Octane	MOD200/400 + MPA	4x100GE SkyHammer 12x100GE SkyHammer
PIDs	A9K-4X100GE-TR A9K-4X100GE-SE A9K-8X100GE-TR A9K-8X100GE-SE	A9K-MOD200-TR/SE w/ MPA A9K-MOD400-TR/SE w/ MPA A9K-MPA-1X100GE A9K-MPA-2X100GE	A9K-4X100GE A99-12X100GE
Interface	4x 100G 8x 100G	2x 100G 4x 100G	4x 100G 12x 100G
Power	1000W (Typical) 1100W (Typical)	445W (Typical) 660W (Typical)	450W (Typical) 1350W (Typical)
Chassis Support	ASR 99xx & 90xx	ASR 99xx & 90xx	ASR 99xx & 90xx
Modes	LAN/WAN/OTN	LAN/WAN/OTN	LAN
MACSec Capabale	Yes	Yes	No
Optics	CPAK. Breakouts option available	CFP2 and CPAK (w/ CFP2-CPAK adapter). Breakout option available	QSFP28. 4x10GE breakout available
Fabric variants	5-fabric only	5-fabric only	5- & 7-fabric
Scale	SE and TR scale	SE and TR scale	SE Scale except for features dependent on TCAM

ASR 9000 3rd Gen TomaHawk 10G Options

	MOD400	48p PowerGlide	MOD200	24p PowerGlide
PIDs	A9K-MOD400-TR/SE w/ A9K-MPA-20X10GE (2)	A9K-48X10GE-1G-TR/SE or A99-48X10GE-1G-TR/SE	A9K-MOD200-TR/SE w/ A9K-MPA-20X10GE	A9K-24X10GE-1G-TR or A9K-24X10GE-1G-SE
Interface	40x 10GE	**48x** 10GE / 1GE	20x 10GE	**24x** 10GE / 1GE
Power	660W (Typical)	700W (Typical)	445W (Typical)	**420W (Typical)**
Chassis Support	ASR 99xx & 90xx	ASR 99xx & 90xx	ASR 99xx & 90xx	ASR 99xx & 90xx
Modes	**LAN/WAN/OTN**	LAN only	**LAN/WAN/OTN**	LAN only
MACSec	**Yes**	No	**Yes**	No
Fabric variants	5-fabric only	5- & **7-fabric**	5-fabric only	5-fabric only
Scale	Tomahawk SE & TR scale	Tomahawk SE & TR scale	Tomahawk SE & TR scale	Tomahawk SE & TR scale

ASR 9000 3rd Gen TomaHawk 10G/1G Options

	MOD200/400		48p PowerGlide
PIDs	A9K-MOD200-TR/SE or A9K-MOD400-TR/SE w/ A9K-MPA-20X1GE (2)	A9K-MOD200-TR/SE or A9K-MOD400-TR/SE w/ A9K-MPA-32X1GE (2)	A9K-48X10GE-1G-TR/SE or A99-48X10GE-1G-TR/SE
Interface	40x 1GE SFP	32x 1GE SFP **64x 1GE cSFP**	**48x** 10GE / 1GE
Power	445W (Typical)	445W (Typical)	700W (Typical)
Chassis Support	ASR 99xx & 90xx	ASR 99xx & 90xx	ASR 99xx & 90xx
MACSec	No	**Yes**	No
Scale	Tomahawk SE & TR scale	Tomahawk SE & TR scale	Tomahawk SE & TR scale

3rd GEN Line Cards: A9K-8X100GE-TR/SE

8x 100G CPAK Line Cards

A9K-8X100G

- SE and TR versions
  - SE System Edge: More memory, scalable QoS
  - TR Transport: Basic QoS, mostly sufficient for Transport or core-facing
- Port breakout
  - 1x 100G or 2x 40G or 10x 10G with breakout cables (or passive panel)

```
Router(config)#hw-module 0/2/cpu0 port 3 breakout 10xTenGigE
```

- Power control per slice [110W/slice]

```
Router(config)#hw-module power saving slice 3 location 0/2/cpu0
```

Line Cards: 3rd Generation [Tomahawk NP]

12x 100G QSFP Line Card

QSFP-100G-LR4-S

QSFP-100G-SR4-S

QSFP-100G-CWDM4-S

QSFP-100G-PSM4-S

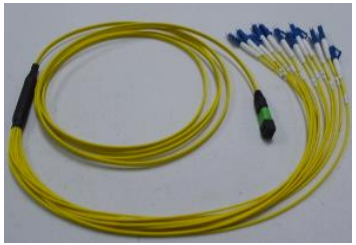
Line Cards: 3rd Generation [Tomahawk NP]

8x & 4x 100G CPAK Line Cards

A9K-8X100G

A9K-4X100G

CPAK



MPO24 TO 10X DUPLEX LC SM

MPO24 TO 10X DUPLEX LC MM

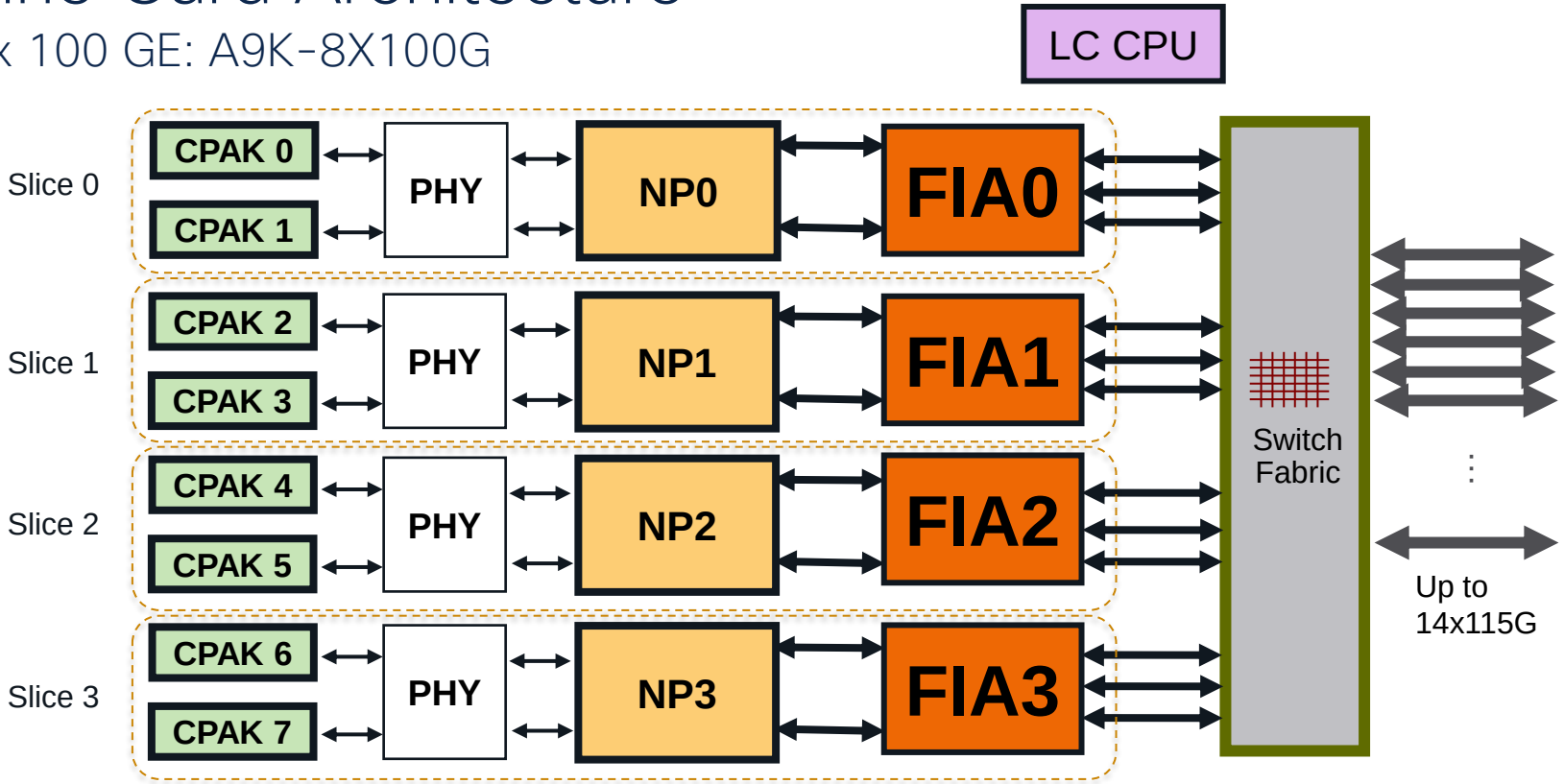
Tomahawk Line Card

Notes on 8x 100G line card

- SE and TR versions
  - SE System Edge: More memory, scalable QoS
  - TR Transport: Basic QoS, mostly sufficient for Transport or core-facing
- CPU and memory
  - Processor: Six core processor
  - RAM: SE 24 GB – TR 12 GB
- Port breakout
  - 1x 100G or 2x 40G or 10x 10G with breakout cables (or passive panel)
    - Router(config)#hw-module 0/2/cpu0 port 3 breakout 10xTenGigE
- Power control per slice [110W/slice]
  - Router(config)#hw-module power saving slice 3 location 0/2/cpu0
- Consumption Model [CM]: Pay as you go per slice

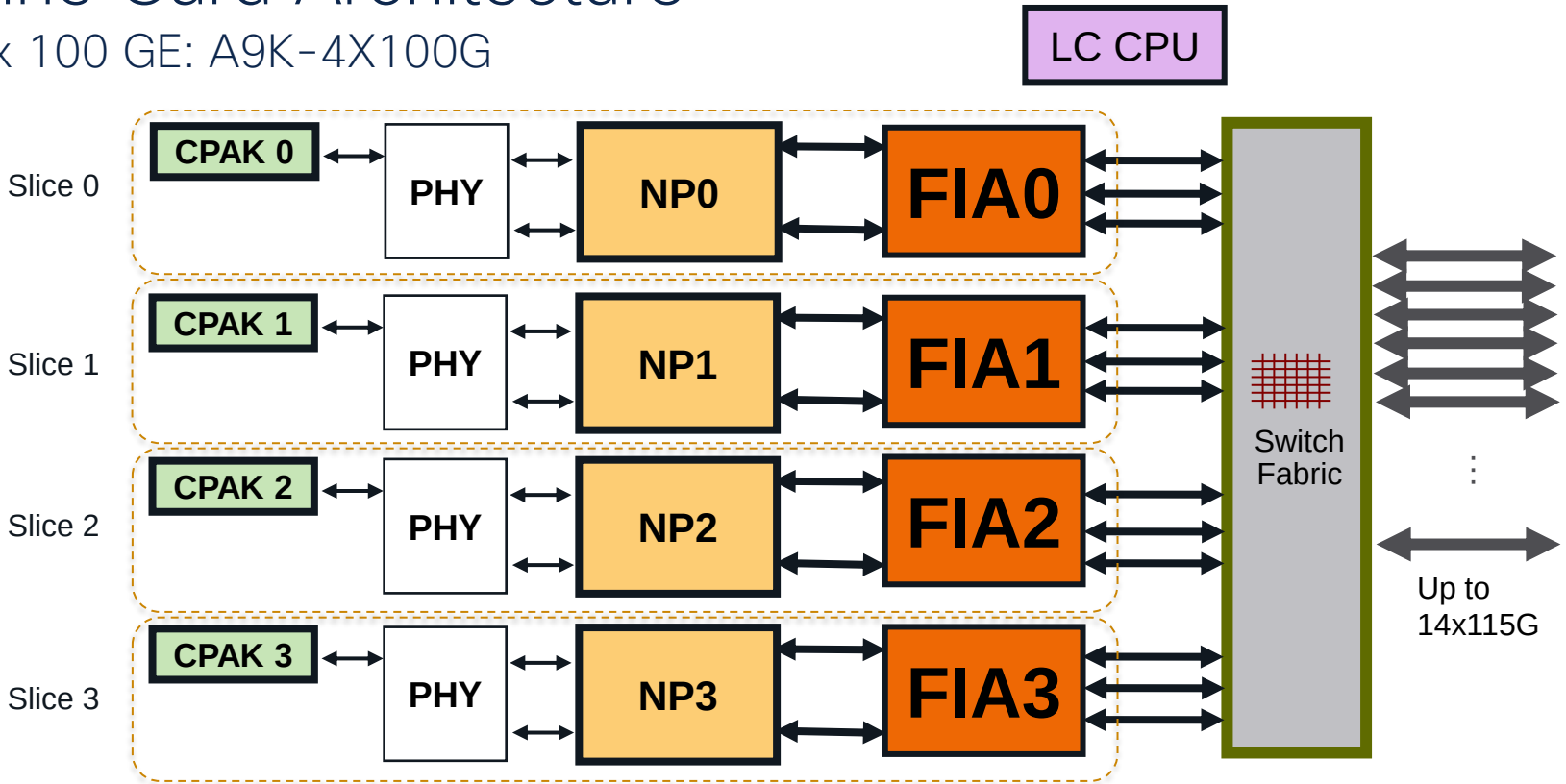
Line Card Architecture

8x 100 GE: A9K-8X100G



Line Card Architecture

4x 100 GE: A9K-4X100G



Tomahawk Interface Flexibility

Single CPAK Product ID → Three SW selectable Options

Configurable 100GE Interconnect Options for 10GE interfaces:

```
Interface HunGigE 0/x/y/z
```

Breakout Interface Convention:
Rack/Slot/Bay/Port (phy)/Breakout#

10GE Interconnect Options

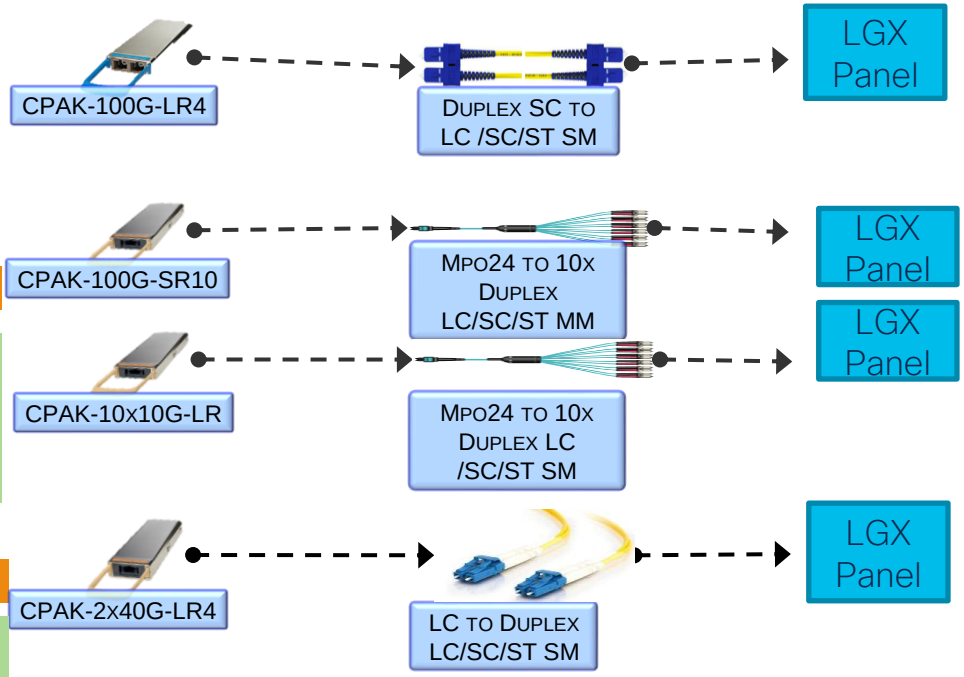
```
hw-module 0/x/cpu0 port z breakout 10xTenGigE
```

```
Interface TenGigE 0/x/y/z/0
Interface TenGigE 0/x/y/z/1
…
Interface TenGigE 0/x/y/z/9
```

40GE Interconnect Options

```
hw-module 0/x/cpu0 port z breakout 2xFortyGigE
```

```
Interface FortyGigE 0/x/y/z/0
Interface FortyGigE 0/x/y/z/1
```



Line Cards: 3rd Generation [Tomahawk NP]

Modular Line Card [2 MPA Bays]

A9K-MOD400

CFP2-DCO optics

IPoDWDM A9K-MPA-1X200GE

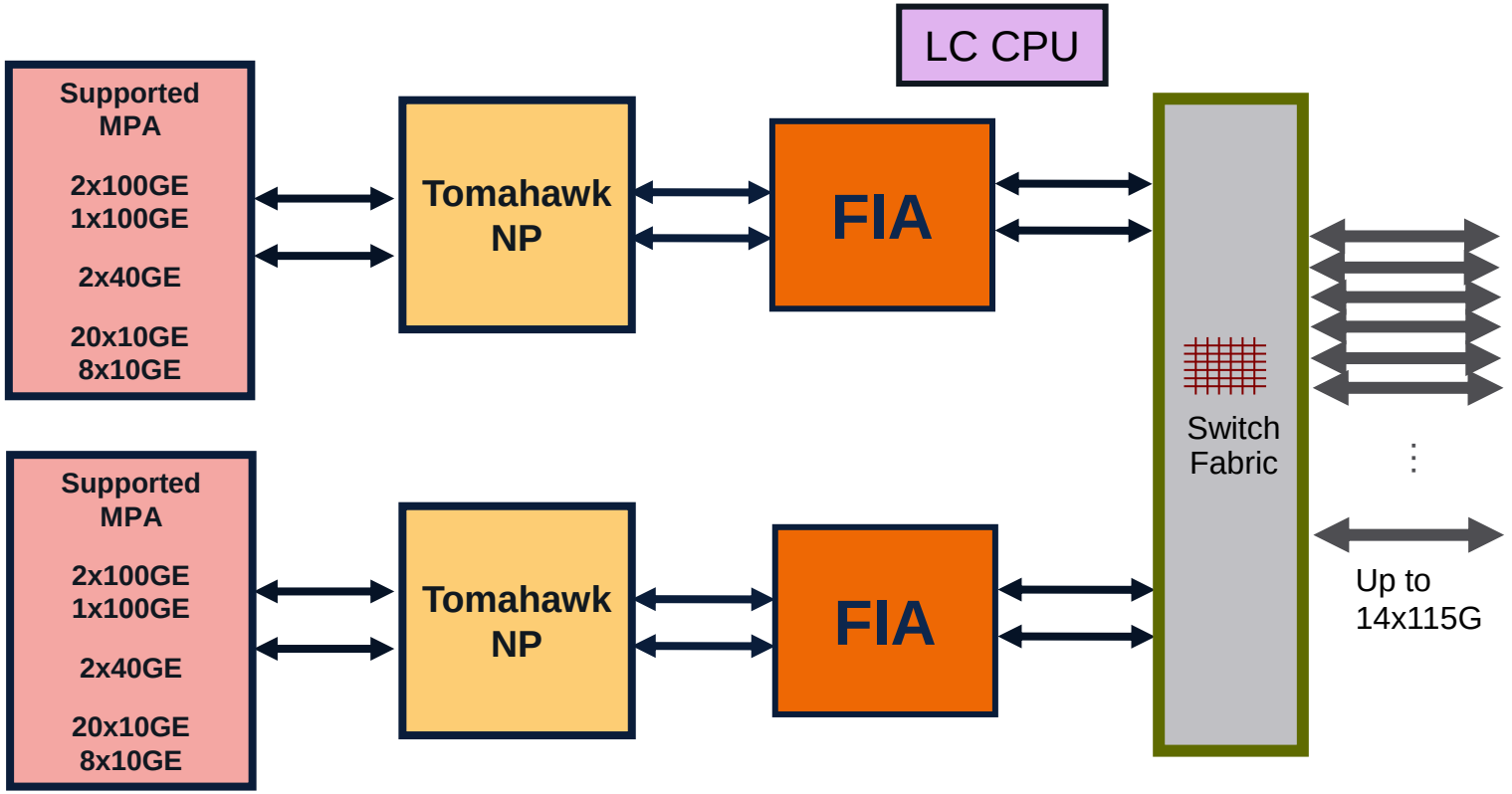
A9K-MPA-2X100GE

A9K-MPA-20X10GE

MPAs
20x1GE
2x10GE
4x10GE
8x10GE
1x40GE
2x40GE

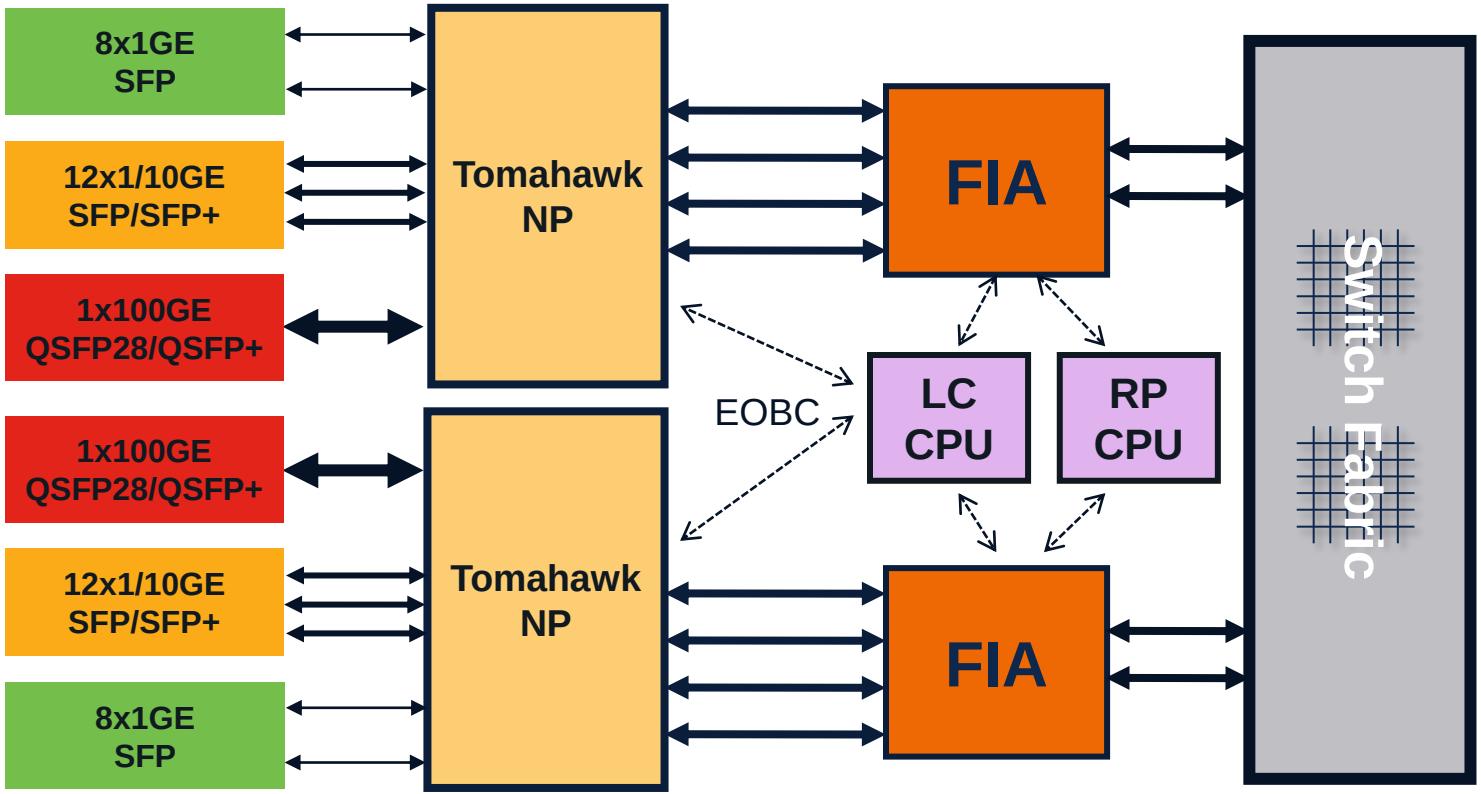
Line Card Architecture

Modular line card: A9K-MOD400-SE and A9K-MOD400-TR

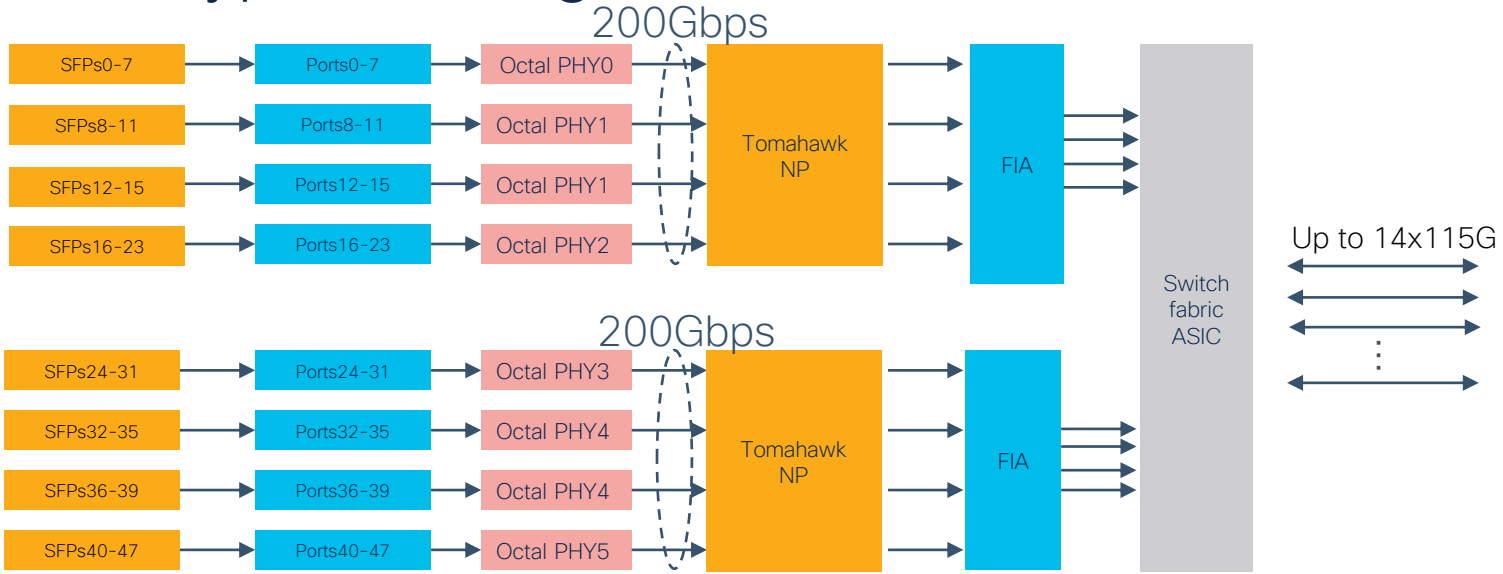


ASR 9901 Architecture

A router + LC in 2 RU



Dense 1G/10GE Dual Rate Tomahawk LCs Trident/Typhoon migration to Tomahawk



- 24-port has a single NPU; 48-port card has two NPUs
- 200G per NPU equally across 24 ports
- Graceful capacity re-distribution across ports in case of oversubscription (Support from 6.2.2)
- No MacSec/OTN PHY features
- Equivalent TCAM size compared to Typhoon → scale parity with Typhoon

48x and 24x10G/1G Port Configuration Rules

- By default all ports are 10G
- Set of 12 ports are grouped together as follows {P0 to P11}, {P12 to P23}, {P24 to P35}, {P36 to P47}.
- 1G port configuration rule: If the first port in any set of 12 ports is configured as 1GE, then the entire group of 12 ports have to be configured as 1GE
- 10G Ports have to be configured in sets of four. Starting ports could be P0, P4, P8, P12 and so on
  - Valid examples: {P0, P1, P2, P3}, {P4, P5, P6,P7}, {P8,P9,P10,P11}.
  - Invalid examples: (P1, P2, P3, P4), {P2, P3, P4, P5} and so on.
- Port configuration command: hw-module location <location> port-mode <string>
  - Example: hw-module location 0/5/CPU0 port-mode 24x10,24x1

48x/24x 1G/10G Port Configuration Examples

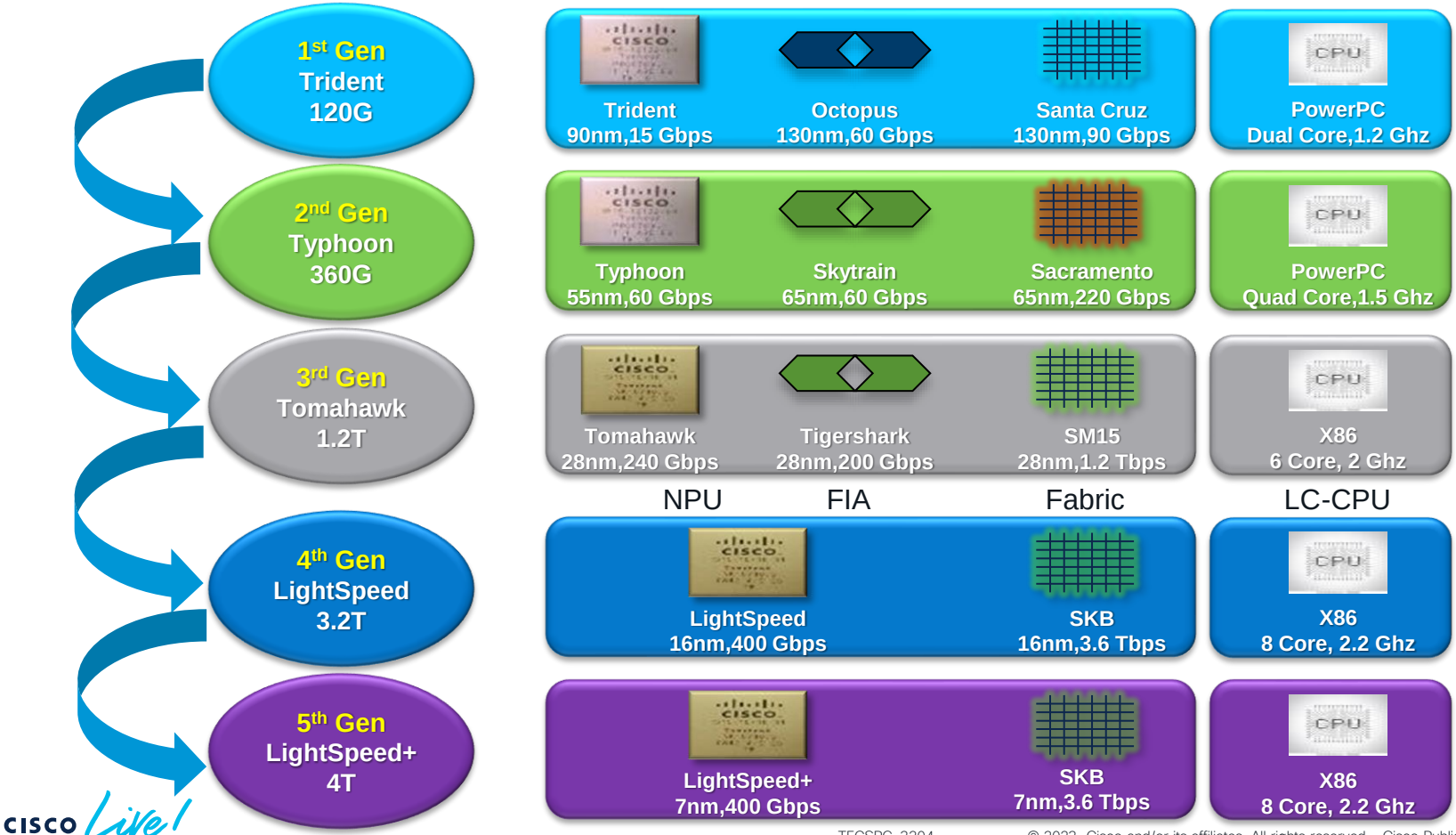
24-port Line card

P0	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	CLI to configure ports as shown
1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	12X1,12X10
10G	10G	10G	10G	1G	1G	1G	1G	1G	1G	1G	1G	10G	10G	10G	10G	10G	10G	10G	10G	1G	1G	1G	1G	4X10,8X1,8X10,4X1
10G	10G	10G	10G	10G	10G	10G	10G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	8X10,16X1
10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	1G	1G	1G	1G	10G	10G	10G	10G	16X10,4X1,4X10
10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	24X10

48-port Line card

P0	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25	P26	P27	P28	P29	P30	P31	P32	P33	P34	P35	P36	P37	P38	P39	P40	P41	P42	P43	P44	P45	P46	P47	CLI to configure ports as shown
1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	12X1,12X10,12X1,12X10
10G	10G	10G	10G	1G	1G	1G	1G	1G	1G	1G	1G	10G	10G	10G	10G	10G	10G	10G	10G	1G	1G	1G	1G	10G	10G	10G	10G	1G	1G	1G	1G	1G	1G	1G	1G	10G	10G	10G	10G	10G	10G	10G	10G	1G	1G	1G	1G	4X10,8X1,8X10, 4X1,4X10,8X1,8X10,4X1
10G	10G	10G	10G	10G	10G	10G	10G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	10G	10G	10G	10G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	24X10,24X1
10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	36X1,12X10
10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	48X10

ASR9000 Edge Linecard Silicon Evolution



ASR 9000 4/5th Gen NPU Architecture

Pluggable physical interfaces — PHY

- speeds: 10GE, 25G, 40GE, 100GE, 400GE
- colors: gray, CWDM, DWDM, Tunable

LightSpeed Plus NPU

Network Processor — NP

- forwarding and feature engine for the LC
- scales bandwidth via multiple NPs
  - up to 10 NPs/LC for performance vs. density options
- highly integrated silicon as opposed to multiple discrete components
  - shorter connections, faster communication channels
  - higher performance, density with lower power draw
  - simplified software development model

Fabric Interface ASIC — FIA

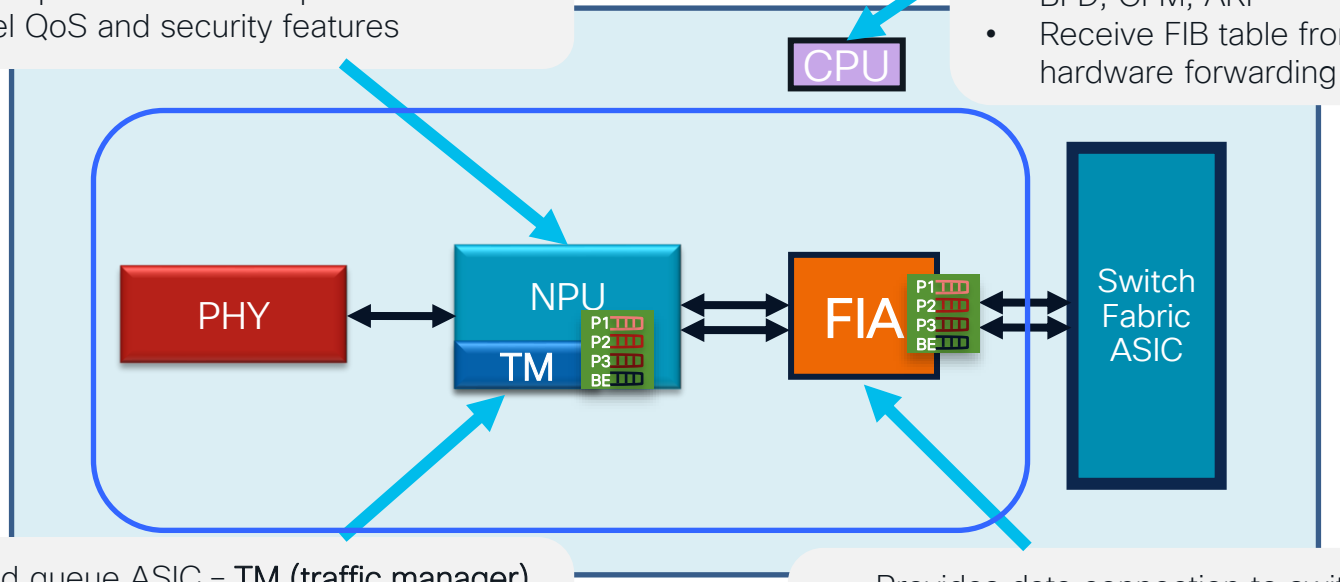
- interface between forwarding processor and system switch fabric
- arbitration, framing, accounting in HW
- provides buffering and virtual output queuing for the switch fabric
- QoS awareness for Hi/Lo and ucast/mcast
  - total flexibility regarding relative priority of unicast vs. multicast

- New, 22 billion transistors, in-house developed NPU
- 420G (bi-directional) bandwidth
- 300Mpps+ full duplex forwarding per NPU

ASR 9000 4/5th Gen Line Card Architecture

- Main forwarding engine for L2 and L3 lookups
- Multicast replication toward Optics
- User level QoS and security features

- Runs distributed control plane protocols for increased scale
- BFD, CFM, ARP
- Receive FIB table from RP and program hardware forwarding table



- Dedicated queue ASIC – **TM (traffic manager)** per NPU for QoS functions
- User configurable queue on TM
- Default Qort Queue always created

- Provides data connection to switch fabric
- Manage VoQ, Superframe and loadbalancing data traffic across switch fabric
- Mcast replication table for replication toward NPs

A99-32x100G-TR – Preparing for Zettabyte Era

Innovation

Cisco NPU 4 in 1 (16nm): Integrated NPU, PHY, FIA, Memory

Native support for 10/25/40/100/400G

Integrated 100GE FEC

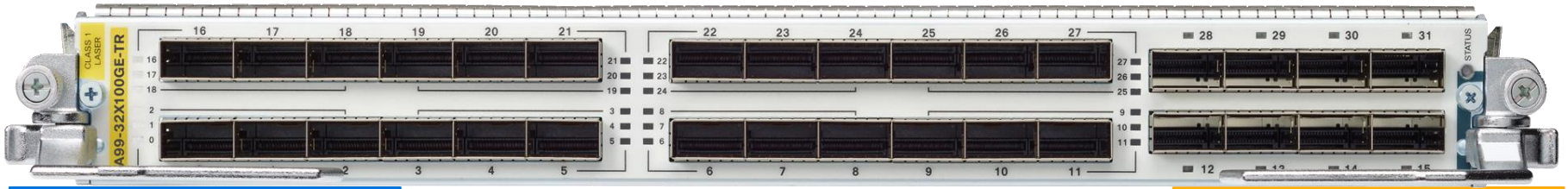
Leading the Market

4 x capacity increase per system

10GE, 40GE, 100GE and 400GE densities w/ rich features

Hitless FPD upgrade possible (no LC reload)

Sub Sec ISSU



Lower TCO

Low OPEX:

- Drastically lower power profile: ~ .5W/GE
- Improvement over Tomahawk w/ power down capability of the complete slice path including NP

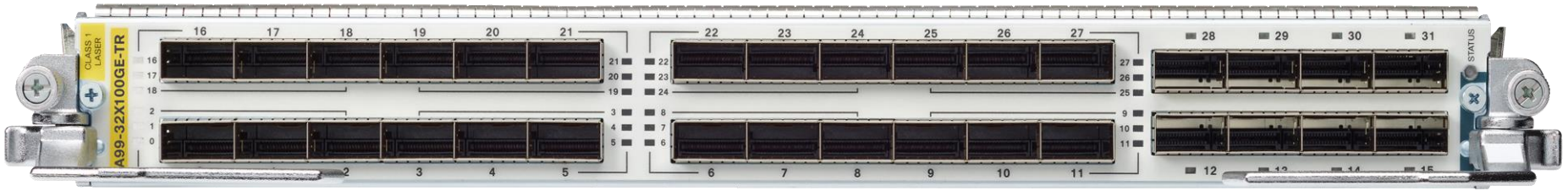
Low CAPEX - Vortex and PAYG

Scale

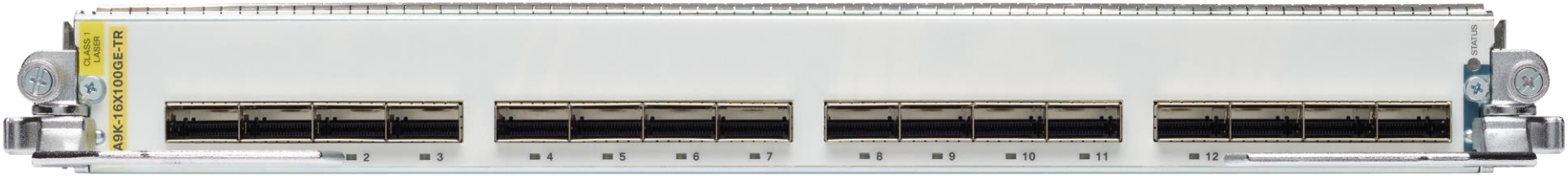
Ultra high control plane scale with eXR

HW acceleration for L2 classification, hashing, prefix lookup, ACL range compression, header re-write, flow ordering, statistics, policers, WRED

Line Cards: 4th Generation (Lightspeed)

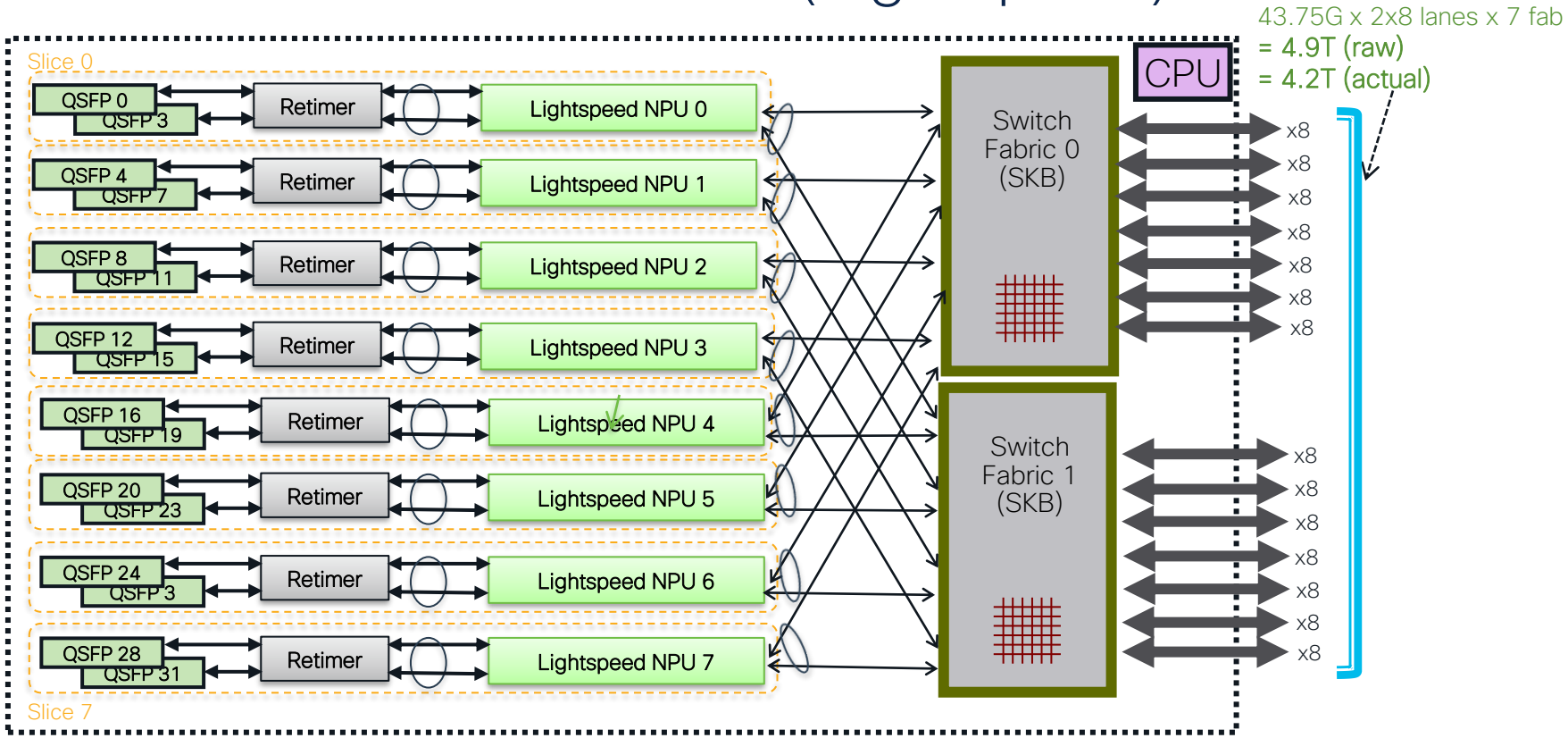


A99-32X100GE–TR



A99-16X100GE–TR

Line Cards: 4th Generation (Lightspeed)



ASR 9000 5th Gen Platform Overview

Cisco NPU 4 in 1 (7nm): Integrated NPU, PHY, FIA, Memory

Native support for 10/25/40/100/400G

Drastically lower power profile: ~ .3W/GE
- Offer power saving & power down capability per NPU

New Compact Systems

ASR 9902

ASR 9903

5th Generation Multi-rate 100GE, 400GE line cards

A9K/A99-4HG-FLEX-TR/SE

A9K-8HG-FLEX-TR/SE

A9K-20HG-FLEX-TR/SE

A99-32X100GE-X-SE/TR

A99-10X400GE-X-SE/TR

ASR 9000 5th Gen Line Cards Release

		PORTS	MACSEC	TIMING	RSP/RP/SFC	CHASSIS	
3.2T	 A99-32X100GE-X-SE/TR/FC	32x QSFP28	No	Class B SyncE	RSP5, RP3 SFC3	9904, 9906 9910, 9912 9922	7.1.15
2T	 A9K-20HG-FLEX-SE/TR/FC	5x QSFP-DD 15x QSFP28	Yes	Class C SyncE	RSP5, RP3 RP2, SFC3	9904, 9906, 9910 9912, 9922, 9006 9010	7.1.15
800G	 A9K-8HG-FLEX-SE/TR/FC	2 Ports QSFP-DD 6 Ports QSFP28	Yes	Class C SyncE	RSP5, RP3, SFC3 RSP880*, SFC2 RP2, RSP880-LT*	9904, 9906, 9910 9912, 9922, 9006 9010	7.1.15
4T	 A99-10X400GE-X-SE/TR/FC	10x QSFP-DD	Yes	Class B SyncE	RSP5, RP3 SFC3	9904, 9906 9910, 9912 9922	7.3.1
400G	 A9K-4HG-FLEX-SE/TR/FC A99-4HG-FLEX-SE/TR/FC	4x QSFP28 16x SFP28 24x SFP+	Yes	Class C SyncE	RSP5, RP3, SFC3 RSP880*, SFC2 RP2, RSP880-LT*	9904, 9906, 9910 9912, 9922, 9006 9010	7.4.0

ASR 9000 5th Gen LC Portfolio

400G to 4T

	Ports	Bandwidth	Combo Ports	MACSec	Timing	RSP / RP
 A99-10x400GE-SE/TR	10 Ports of QSFP-DD	4 Tbps	Yes	MACSec	Class B	RSP5, RP3
 A99-32X100GE-X-SE/TR	32 Ports of QSFP28	3.2 Tbps	No	No	Class B	RSP5, RP3
 A9K-20HG-FLEX-SE/TR	15 Ports QSFP28 5 Ports QSFP-DD	2 Tbps	Yes	MACSec	Class C	RSP5, RP3, RP2
 A9K-8HG-FLEX-SE/TR	6 Ports QSFP28 2 Ports QSFP-DD	800 Gbps	Yes	MACSec	Class C	RSP5, RSP880-LT, RSP880, RP3, RP2
 A9K-4HG-FLEX-SE/TR A99-4HG-FLEX-SE/TR	4 Ports QSFP28 16 Ports SFP28 (25G) 24 Ports SFP+	400 Gbps	Yes	MACSec	Class C	RSP5, RSP880-LT, RSP880, RP3, RP2

ASR 9000 5th Gen 100G Offers

Line Cards Systems	A99-10X400GE-X-SE A99-10X400GE-X-TR 	A99-32X100GE-X-SE A99-32X100GE-X-TR 	A9K-20HG-FLEX-SE A9K-20HG-FLEX-TR 	A9K-8HG-FLEX-SE A9K-8HG-FLEX-TR 	ASR 9903 + 2T PEC 	ASR 9902 
Max 100G Ports	40*	32	20	8	36	8
MACSec Support	Yes	No	Yes	Yes	Yes	Yes
PTP/Timing	Class-C	Class-C	Class-C	Class-C	Class-C	Class-C
100G Coherent***	Yes QSFP-DD ZR+	Yes QSFP-DD ZR+	Yes QSFP-DD ZR+	Yes QSFP-DD ZR+	Yes QSFP-DD ZR+	Yes QSFP-DD ZR+
Mixed Interface Rate	10x400G 40x100G*	32x100G 32x40G 84x10G*+1 1x100G	5x400G 20x100G 20x40G 80x10G*	2x400G 8x100G 8x40G 32x10G*	5x400G+1 6x100G 36x100G 36x40G 14x100G+ 20x10G**	8x100G 8x40G (16+16*)x25G (40+16*)x10G

*Breakout; **Fixed Board Only;***Roadmap

ASR 9000 5th Gen 25G/10G Options

Line Cards Systems	A99-32X100GE-X-SE A99-32X100GE-X-TR 	A9K-20HG-FLEX-SE A9K-20HG-FLEX-TR 	A9K-8HG-FLEX-SE A9K-8HG-FLEX-TR 	A9K-400G-X-SE A9K-400G-X-TR 	ASR9903 + 0.8T PEC 	ASR 9902 
Max 25G Ports	84*	80*	32*	16	32 + x*	16+16*
Max 10G Ports	84*	80*	32*	40	68	40+16*
MACSec Support	No	Yes	Yes	Yes	Yes	Yes
10G OTN	No	No	No	Yes**	Yes**	Yes**

*Breakout; **Roadmap

ASR 9000 5th Gen LC 1G Options

Native 1G Ports on THK LCs

MOD200/400 +20x1G or 32x1G MPA*

24/48x10G-1G

100FX/100BX/100ZX/ BX/EX/ZX/SX/LH/ DWDM/CWDM/Copper

*MACSec support

nV Satellite

ASR9000v-V2

44 1GE SFP access
4x10G SFP+ fabric links

100FX/100BX/BX/EX/ SX/LH/ZX/CWDM/Copper

NCS5001/5002

40-80 1/10GE SFP access
4x100G QSFP28 fabric-links

EX/SX/LH/ZX/ DWDM/CWDM/Copper

MACSec support

Smart Dual Rate optics*

Dual rate 1/10G SFP
SR/LR/ER

2.5 W/Optic

MACSec TBD

Smart 10x1G Breakout optics**

10x1G Breakout cable
SR/LR

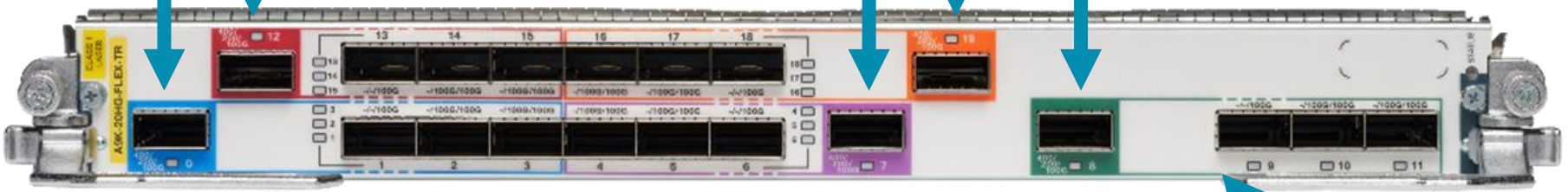
3 W/Optic

* CC'd

** Not committed

5th Gen Line Cards: A9K-20HG-FLEX-SE/TR

400G/200G/100G Multi-Rate Ports: 0/7/8/12/19



A9K-20HG-FLEX-SE/TR

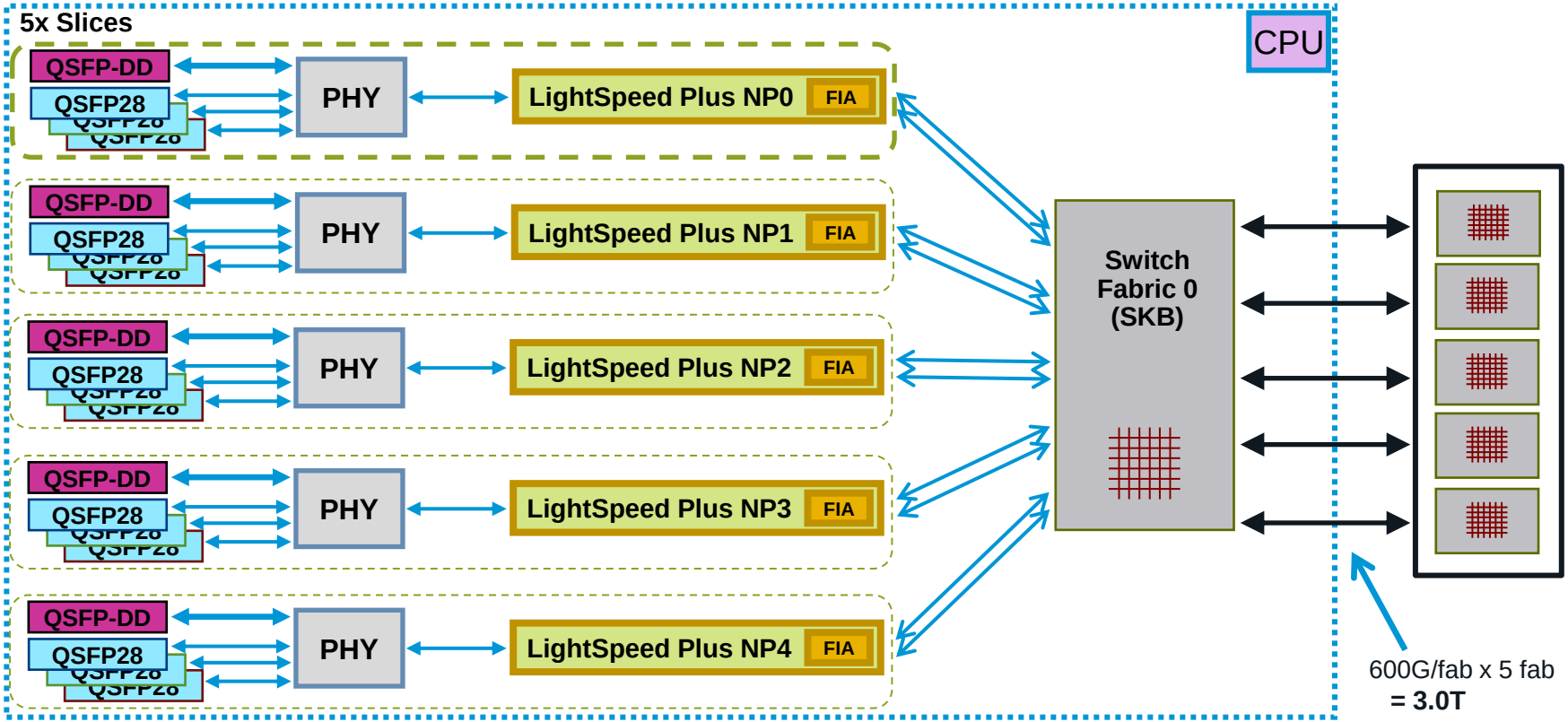
- 5 x 400G QSFP-DD ports
- 15 x 100G QSFP28 ports
- 400G/200G/100G/40G/25G/10G support
- Each slice can be independently configured as:
  - 1x400G –or-
  - 1x200G + 2x100G –or-
  - 4x100G
- Each 100G can break out into 4x25G or 4x10G

5 Slices:
1x 400G/200G/100G multi-rate port
& 3x100G ports per Slice

Uses Meta-DX1 PHY with MACsec and FlexE Support

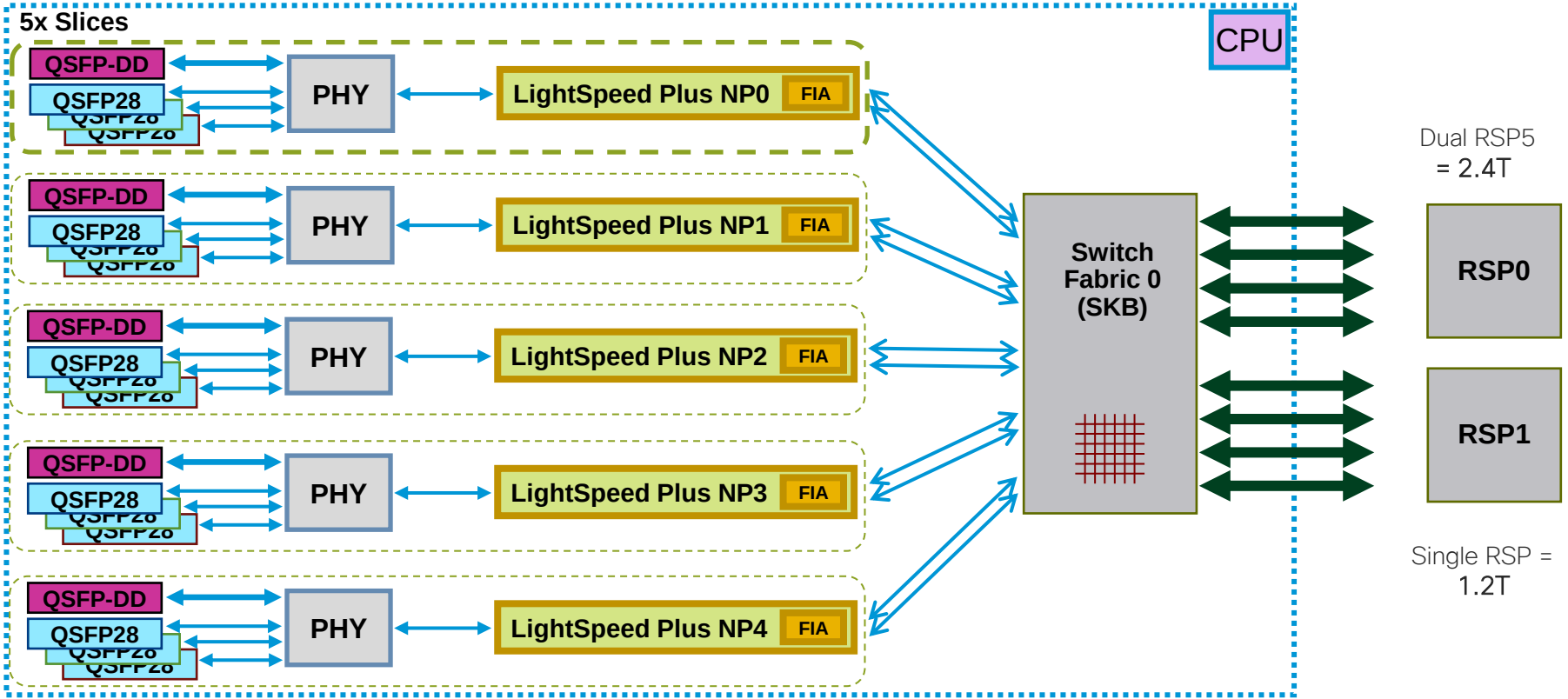
A9K-20HG-FLEX-SE/TR (5-fabric) LC Architecture

(when used in 9922, 9912, 9910 & 9906)



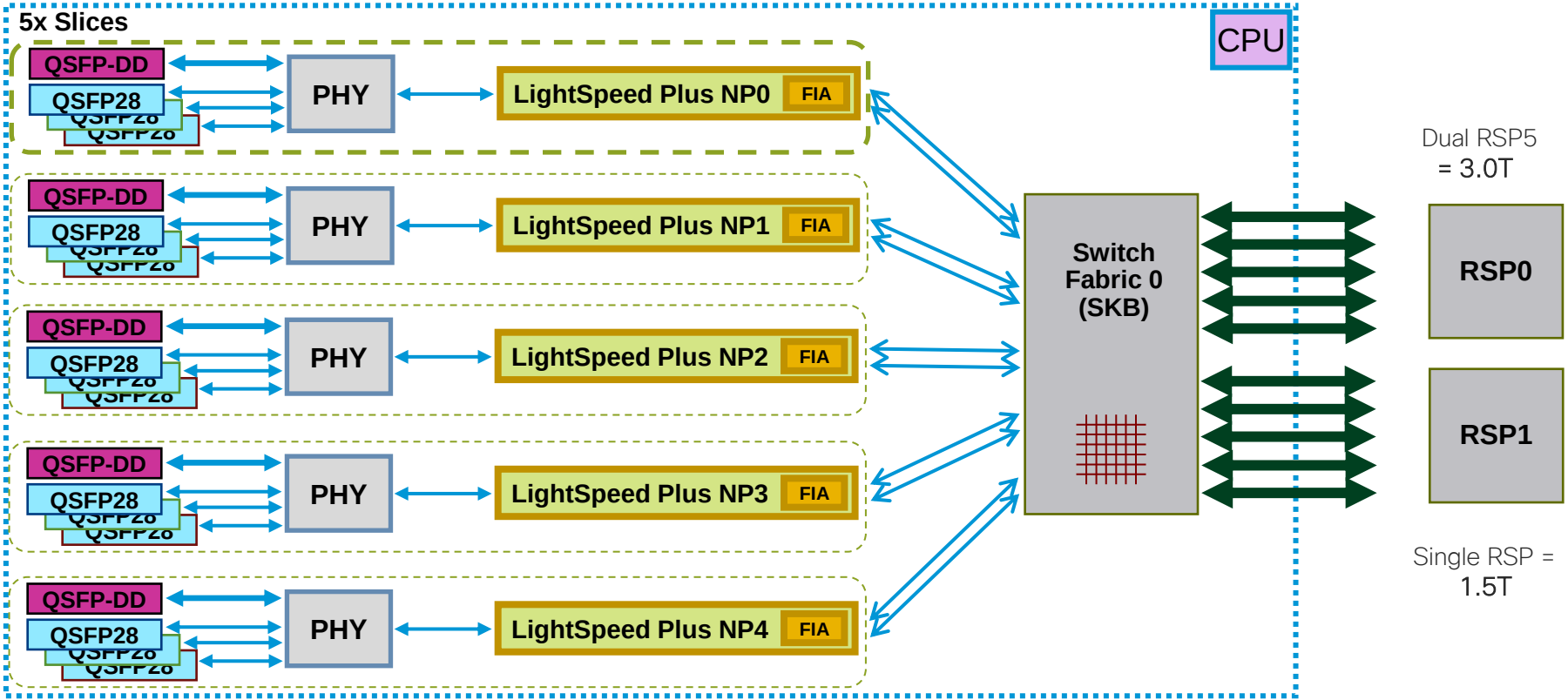
A9K-20HG-FLEX-SE/TR (5-fabric) LC Architecture

(when used in 9010 & 9006)



A9K-20HG-FLEX-SE/TR (5-fabric) LC Architecture

(when used in 9904)



ASR 9000 Line Card Backpressure

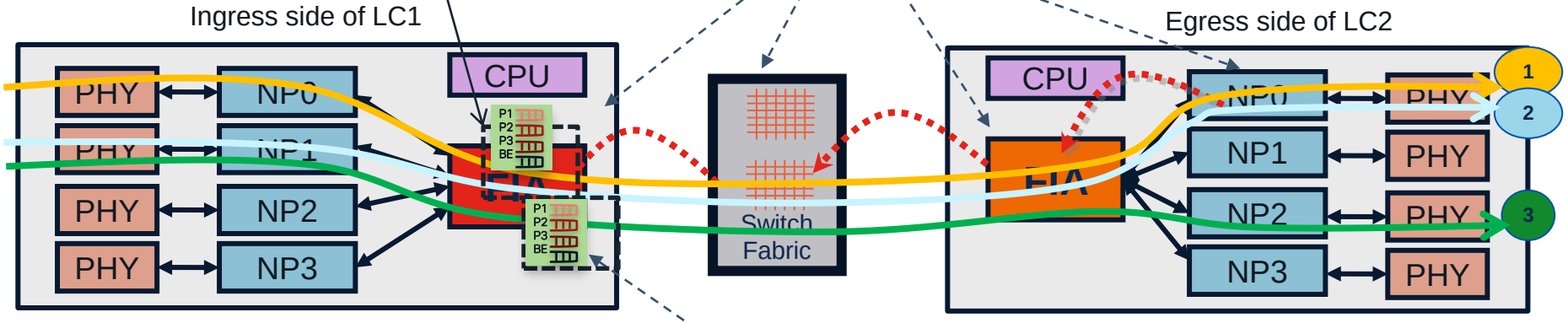
Egress NP congestion → backpressure to ingress FIA →

Packet is en-queued in the dedicated VoQ →

No impact of the packet going to different egress NP --- No head-of-line-block issue

One VoQ set (4 queues P1-3, BE) for each entity in the system

Backpressure: egress NP → egress FIA → fabric Arbiter → ingress FIA → VoQ



Packet going to different egress NP put into different VoQ set → Congestion on one NP won't block the packet going to different NP

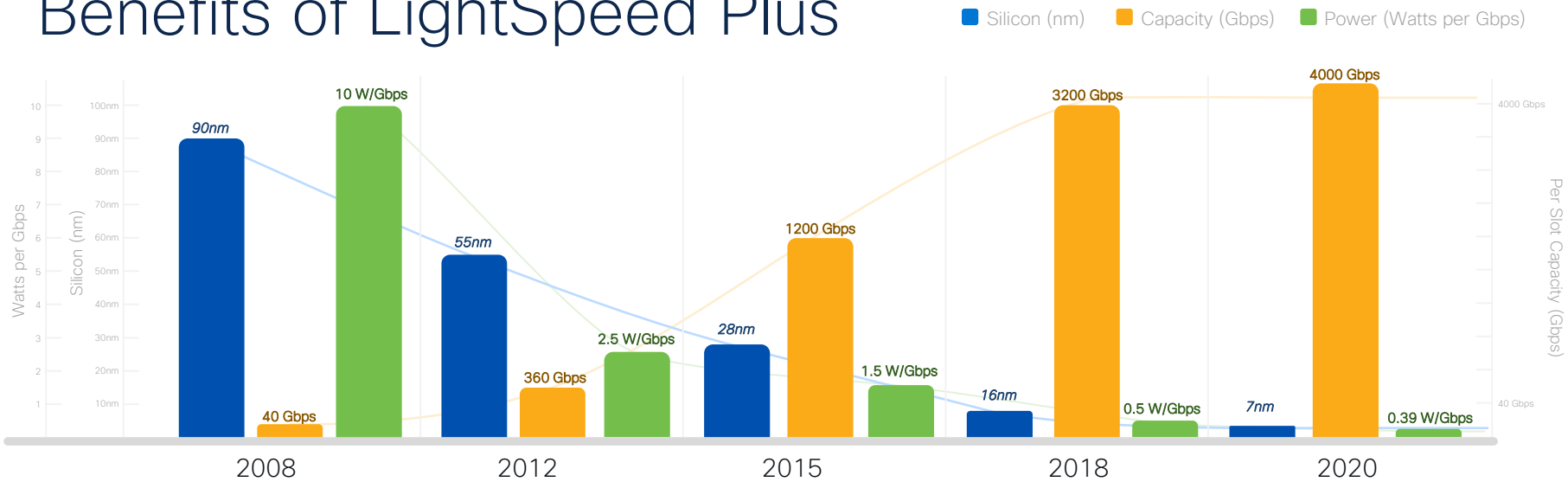
ASR 9000 5th Gen Line Card Compatibility

	LSP Line Card				
Support of:	A99-10X400GE-X-SE A99-10X400GE-X-TR (7-fabric)	A99-32X100GE-X-SE A99-32X100GE-X-TR (7-fabric)	A9K-20HG-FLEX-SE A9K-20HG-FLEX-TR (5-fabric)	A9K-8HG-FLEX-SE A9K-8HG-FLEX-TR (5-fabric)	A9K-4HG-FLEX-SE A9K-4HG-FLEX-TR A99-4HG-FLEX-SE A99-4HG-FLEX-TR (5 and 7-fabric)
ASR 9922, 9912, 9910, 9906 & 9904 chassis	✓	✓	✓	✓	✓ (5 & 7 fabric)
ASR 9010 & 9006 chassis	✗	✗	✓	✓	✓ (A9K-5 fabric)
RP3, RP3-X, SFC3, RSP5, RSP5-X	✓	✓	✓	✓	✓
RP2	✗	✗	✓	✓	✓
SFC2	✗	✗	✗	✓	✓
RSP880/RPS880-LT*	✗	✗	✗	✓*	✓

*Exception: 9010, 9006 & 9904 require RSP5 if fabric redundancy is required

NOTE: Do not forget that ASR 9922 requires the V3 Fan tray!

Investment Protection and Power Efficiency Benefits of LightSpeed Plus



Silicon	1St Generation Trident 90nm, 15 Gbps	2nd Generation Typhoon, 55nm, 60 Gbps	3rd Generation Tomahawk, 28nm, 240 Gbps	4th Generation LightSpeed, 16nm, 400 Gbps	5th Generation LightSpeed Plus 7nm, 400 Gbps
Fabric Interconnect ASIC	Octopus 130nm, 60 Gbps	Skytrain 65nm, 60 Gbps	Tigershark 28nm, 200 Gbps	-	-
CPU	PowerPC Dual Core,1.2 Ghz	PowerPC Quad Core,1.5 Ghz	X86 6 Core, 2 Ghz	X86 8 Core, 2.2 Ghz	X86 8 Core, 2.2 Ghz
	EoS	EoS	Shipping!!!	Shipping!!!	Shipping!!!

ASR 9000 5th Gen Line Card Power Saving

- ASR 9000 5th Gen LC does not require a specific Power System version!
  - Chassis must have enough power budget available to accommodate new line cards/commons

- ASR 9000 5th Gen LC support power saving

  (config)#hw-module location <slot#> slice <slice#> ?

```
ASR9K(config)#hw-module location 0/2/CPU0 slice 2 power-savings
```

```
ASR9K(config)#hw-module location 0/2/CPU0 slice 3 power-down
```

- Check available power budget using Cisco Power Calculator

  → http://tools.cisco.com/cpc

ASR9k Power Compatibility Matrix

A9K-AC-PEM-V2: 4 PSUs/PEM
A9K-AC-PEM-V3: 3 PSUs/PEM
A9K-DC-PEM-V2: 4 PSUs/PEM
A9K-DC-PEM-V3: 4 PSUs/PEM

			ASR 9922	ASR 9912	ASR 9910	ASR 9906	ASR 9904	ASR 9010	ASR 9006
	Number of Power Trays		4	3	2	1	1	2	1
AC Power	PEM	A9K-AC-PEM-V2	✓	✓			✓	✓	✓
		A9K-AC-PEM-V3	✓	✓	✓	✓		✓	
	Power Supply	PWR-3KW-AC-V2	✓	✓			✓	✓	✓
		PWR-6KW-AC-V3	✓	✓	✓	✓		✓	
DC Power	PEM	A9K-DC-PEM-V2	✓	✓			✓	✓	✓
		A9K-DC-PEM-V3	✓	✓	✓	✓		✓	
	Power Supply	PWR-2KW-DC-V2	✓	✓			✓	✓	✓
		PWR-4.4KW-DC-V3	✓	✓	✓	✓		✓	

Note: AC/DC & V2/V3 can not be mixed in a chassis!

LightSpeed Plus Power Consumption per Gbps

	Power Consumption per Bandwidth (Watts/Gbps) (without commons)		
	27C	40C	50C/55C
A99-10X400GE-X-SE/TR line card	0.31 W/Gbps	0.34 W/Gbps	0.36 W/Gbps
A99-32X100GE-X-SE/TR line card	0.29 W/Gbps	0.33 W/Gbps	0.34 W/Gbps
A9K-20HG-FLEX-SE/TR line card	0.42 W/Gbps	0.47 W/Gbps	0.49 W/Gbps
A9K-8HG-FLEX-SE/TR line card	0.52 W/Gbps	0.56 W/Gbps	0.64 W/Gbps

ASR9922/9912/9910/9906: Tomahawk to LightSpeed Plus Migration

Tomahawk LC can co-exist with LightSpeed LCs in the same chassis!

Trident / Typhoon cards are not supported with LightSpeed Plus!

1.5T per slot Tomahawk based System

Today: Tomahawk LCs

RSP880/RP2/SFC2

9906/9910: V3 PEMs, V1 FTs

9922: V2/V3 PEMs, V2 FTs

9912: V2/V3 PEMs, V1 FTs

Up to 4.2T per slot LSP based System

ASR9910/9906	ASR9912	ASR9922
New: 2x RSP5, 5xSFC3-(S/T)	New: 7x SFC3, RP3	New: 7xSFC3, RP3, V3 fan tray
Reuse: V1 FTs, V3 PEMs	Reuse: V1 FTs, V2/V3 PEMs	Reuse: V2/V3 PEMs

- A99-10X400GE-X-SE/TR LC
  - 4T BW per LC w/ 7xFabrics*
  - 3.6T BW per LC w/ 6xFabrics
- A99-32GE-X-SE/TR LC
  - 3.2T BW per LC w/ 7xFabrics
  - 3.2T BW per LC w/ 6xFabrics
- A9K-20HG-FLEX-SE/TR Combo LCs**
  - 2T BW per LC w/ 5 x Fabrics
  - 2T BW per LC w/ 4xFabrics
- A9K-8HG-FLEX-SE/TR Combo LC **
  - 800G BW per LC w/ 5xFabrics
  - 800G BW per LC w/ 4xFabrics
- A9K-4HG-FLEX-SE/TR ** or A99-4HG-FLEX-SE/TR Combo LC
  - 400G BW per LC w/ 5xFabrics
  - 400G BW per LC w/ 4xFabrics

* Fabric located on RSP3 or SFC3

** All 5x fabric LSP LCs require 5 x Fabrics in the system for redundancy (will be relaxed)

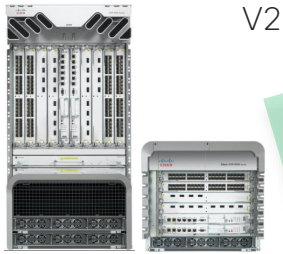
ASR9010/9006: Tomahawk to LightSpeed Plus Migration

Tomahawk LC can co-exist with LightSpeed LCs in the same chassis!

Trident / Typhoon cards are not supported with LightSpeed Plus!

860G per slot RSP880 based System

Today: Tomahawk LC
2x RSP880,
V2 PEMs, V2 FTs



Up to 2.4T per slot LSP based System

New: 2x RSP5*,
Reuse: V2 FTs, V2 PEMs**

- A9K-20HG-FLEX-SE/TR Combo LCs
  - 2T BW per LC w/ Dual RSP5
  - 1.2T BW per LC w/ Single RSP5
- A9K-8HG-FLEX-SE/TR*** Combo LC
  - 800G BW per LC w/ Dual RSP5
  - 800G BW per LC w/ Single RSP5
- A9K-4HG-FLEX-SE/TR**** Combo LC
  - 400G BW per LC w/ Dual RSP5
  - 400G BW per LC w/ Single RSP5

* It's allowed to have RSP5 and RSP880 in the same chassis during RSP migration.

** No specific requirements for Power System version; just check power budget using Power Calculator

*** A9K-8HG-FLEX-SE/TR can be used with Dual RSP880, but needed RSP5 to provide fabric redundancy

**** A9K-4HG-FLEX-SE/TR can be used with Dual RSP880

ASR9904: Tomahawk to LightSpeed Plus Migration

Tomahawk LC can co-exist with LightSpeed LCs in the same chassis!

Trident / Typhoon cards are not supported with LightSpeed Plus!

1.29T per slot RSP880 based System

Today: Tomahawk LC
2x RSP880,
V2 PEMs, V1 FTs

Up to 4.2T per slot LSP based System

New: 2x RSP5*,
Reuse: V1 FTs, V2 PEMs

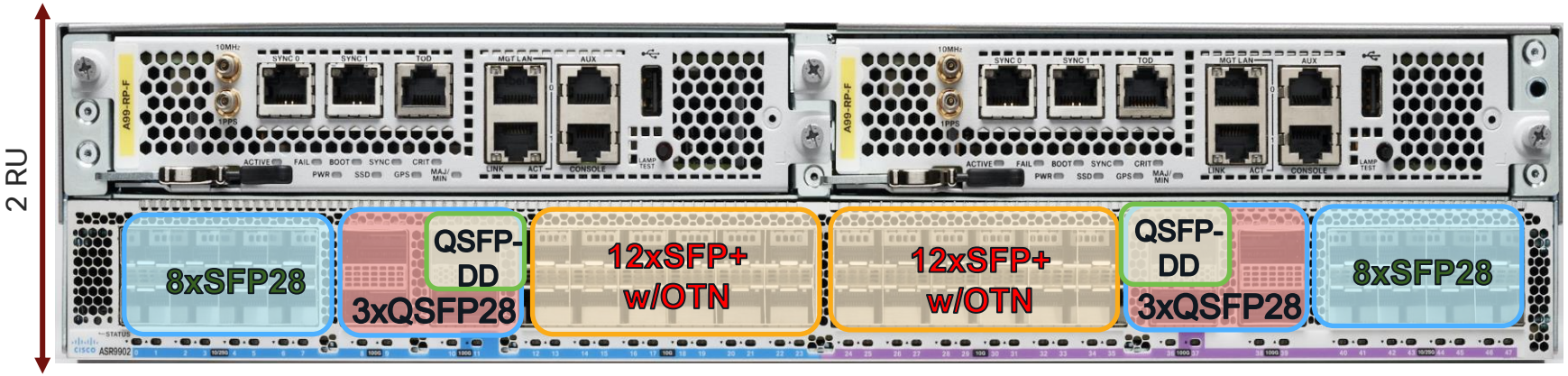
- A99-10X400GE-X-SE/TR LC
  - 4T BW per LC w/ Dual RSP5
  - 2.1T BW per LC w/ Single RSP5
- A99-32X100-X-SE/TR LC
  - 3.2T BW per LC w/ Dual RSP5
  - 1.8T BW per LC w/ Single RSP5
- A9K-20HG-FLEX-SE/TR Combo LCs
  - 2T BW per LC w/ Dual RSP5
  - 1.5T BW per LC w/ Single RSP5
- A9K-8HG-FLEX-SE/TR** Combo LC
  - 800G BW per LC w/ Dual RSP5
  - 800G BW per LC w/ Single RSP5
- A9K-4HG-FLEX-SE/TR*** or A99-4HG-FLEX-SE/TR Combo LC
  - 400G BW per LC w/ Dual RSP5
  - 400G BW per LC w/ Single RSP5

* It's allowed to have RSP5 and RSP880 in the same chassis during RSP migration.

** A9K-8HG-FLEX-SE/TR can be used with Dual RSP880, but needed RSP5 to provide fabric redundancy

*** A9K-4HG-FLEX-SE/TR and A99-4HG-FLEX-SE/TR can be used with Dual RSP880

ASR 9902 Compact System



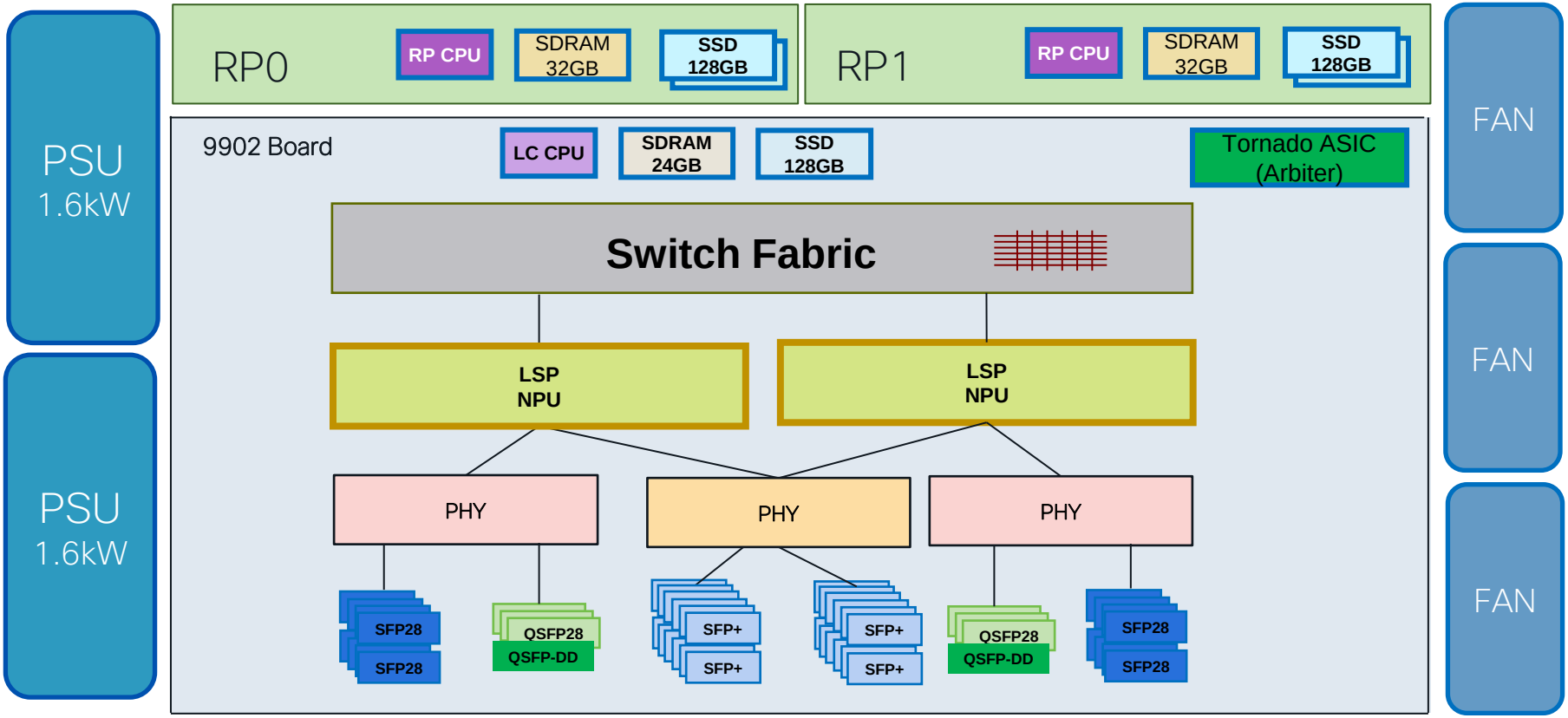
- 800G Throughput
- Compact size – 2RU
- **Redundant RPs**
- Redundant Power/Cooling
- Low Power Consumption (690W)

- 5th Generation NPU
- 10G OTN support
- MACSec
- **PTP and Class C Timing**
- 100G Coherent WDM (ZR+)

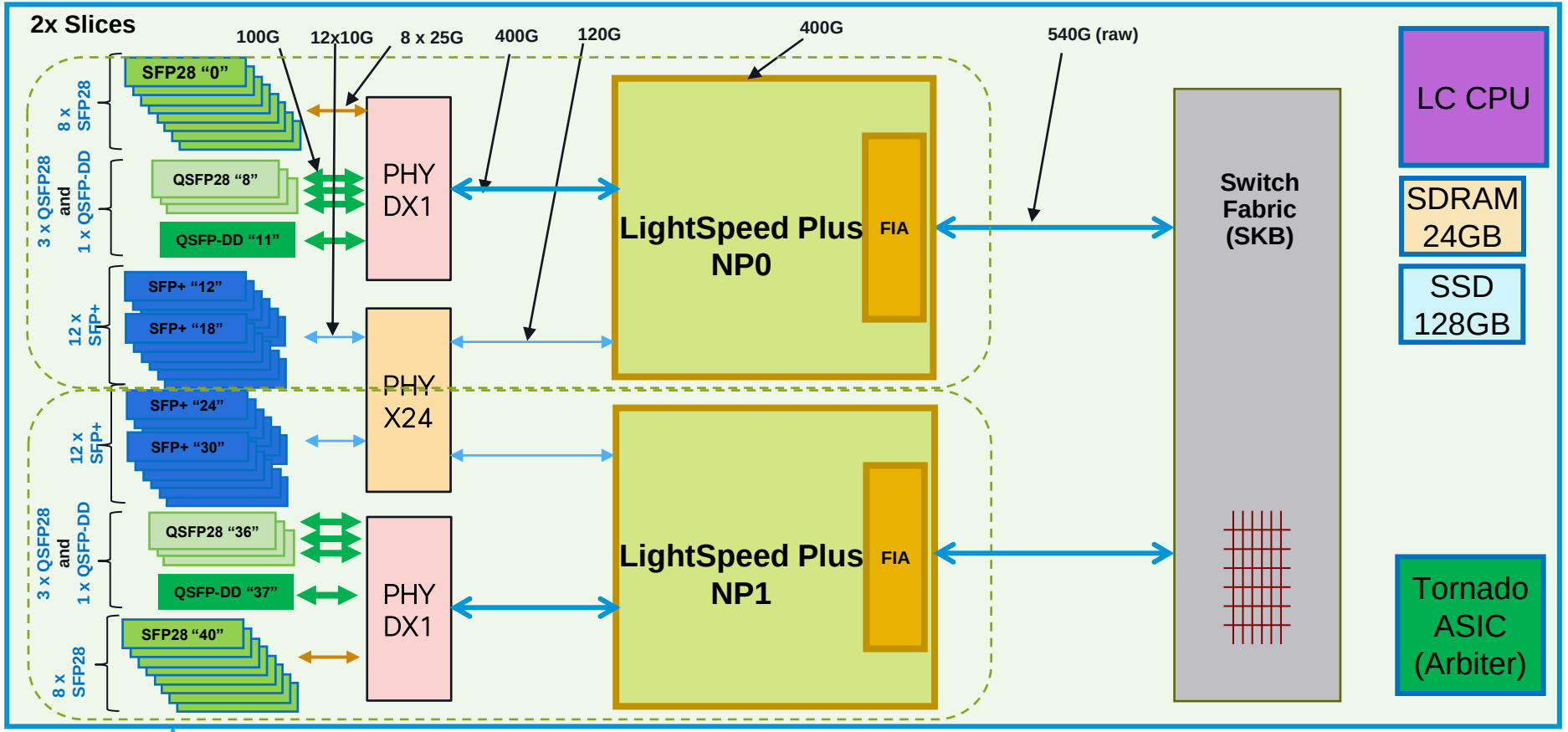
- Ethernet Ports*:
  - 6x100G QSFP28
  - 2x100G QSFP-DD (ZR+)
  - 16x10/25G SFP28
  - 24x10G SFP+ w/OTN

*1GE support via either option: nV satellite solution (available), Smart 10G/25G SFP or 10x1G breakout SFP (roadmap)

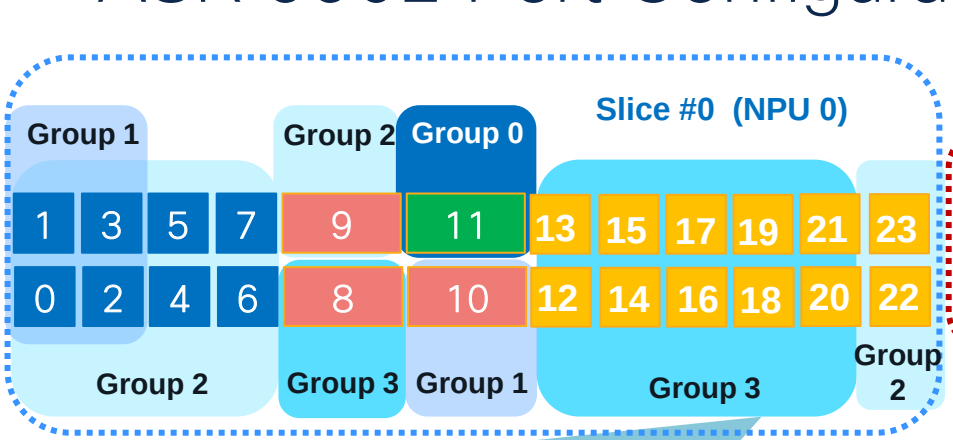
ASR 9902 Hardware Architecture

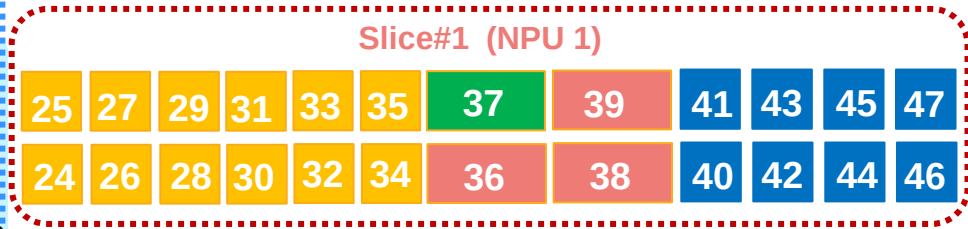


ASR 9902 Data Plane Architecture



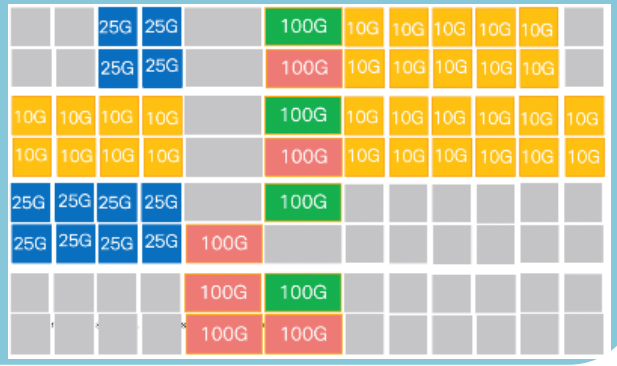
ASR 9902 Port Configuration Options (FCS)





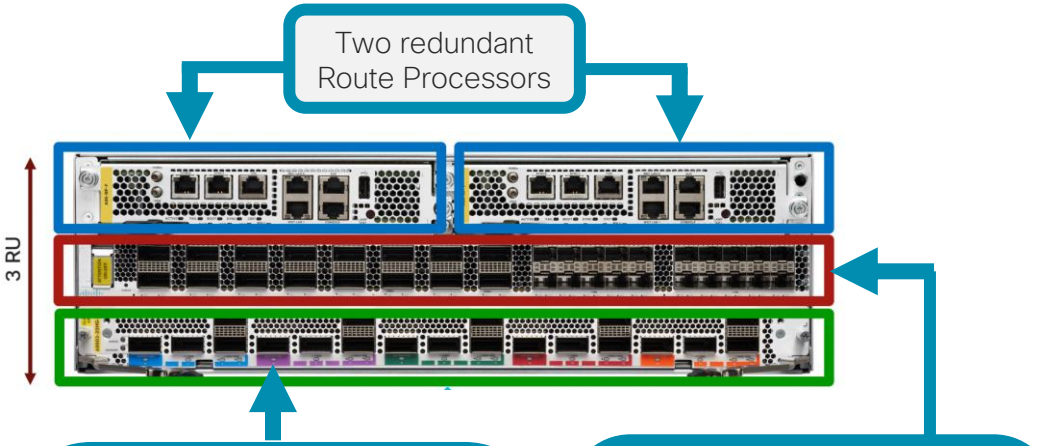
QSFP-DD
QSFP28
SFP28
SFP+
Not used

Option	Port Group 0	Port Group 1	Port Group 2	Port Group 3
1	Port 11 (100G)	Port 10 (100G)	Ports 4-7, (25G)	Ports 12-21 (10G)
2 (default)	Port 11 (100G)	Port 10 (100G)	Ports 0-7, 22-23 (10G)	Ports 12-21 (10G)
3	Port 11 (100G)	Ports 0-3 (25G)	Ports 4-7 (25G)	Port 8 (100G)
4	Port 11 (100G)	Ports 10 (100G)	Ports 9 (100G)	Ports 8 (100G)



Option 1: `hw-module location 0/0/CPU0 slicenum <0/1> config-mode 1x100GE,1x100GE,4x25GE,10x10GE`

ASR 9903 Compact System with PEC



PORT EXPANSION CARD (PEC)

- **2T PEC**: up to 3.6T per chassis

Maximum capacity, high-rate ports

- 5x QSFP-DD: 400/200/100/40/4X25/4X10GE
- 15x QSFP28: 100/40/4X25/4X10GE

- **800G PEC**: up to 2.4T per chassis

High density of low-rate ports

- 32x SFP+/SFP28: 25/10GE
- 16x SFP+: 10GE

FIXED BOARD

- **Integrated ports:**
  - 16x QSFP28: 100/40/4X25/4X10GE
  - 20x SFP+: 10GE
- **Integrated Switch Fabrics:**
  - 2x Fabric ASICs for redundancy

Up to 3.6T Bandwidth per box

Compact form-factor: 3RU, 600mm deep

Port Expansion Slot for extra capacity/flexibility

RP/Fabric/Power/Fan Redundancy

Front-to-Back Airflow

Wide range of supported interface speeds: 10G/25G/40G/100G/400G, including DCO

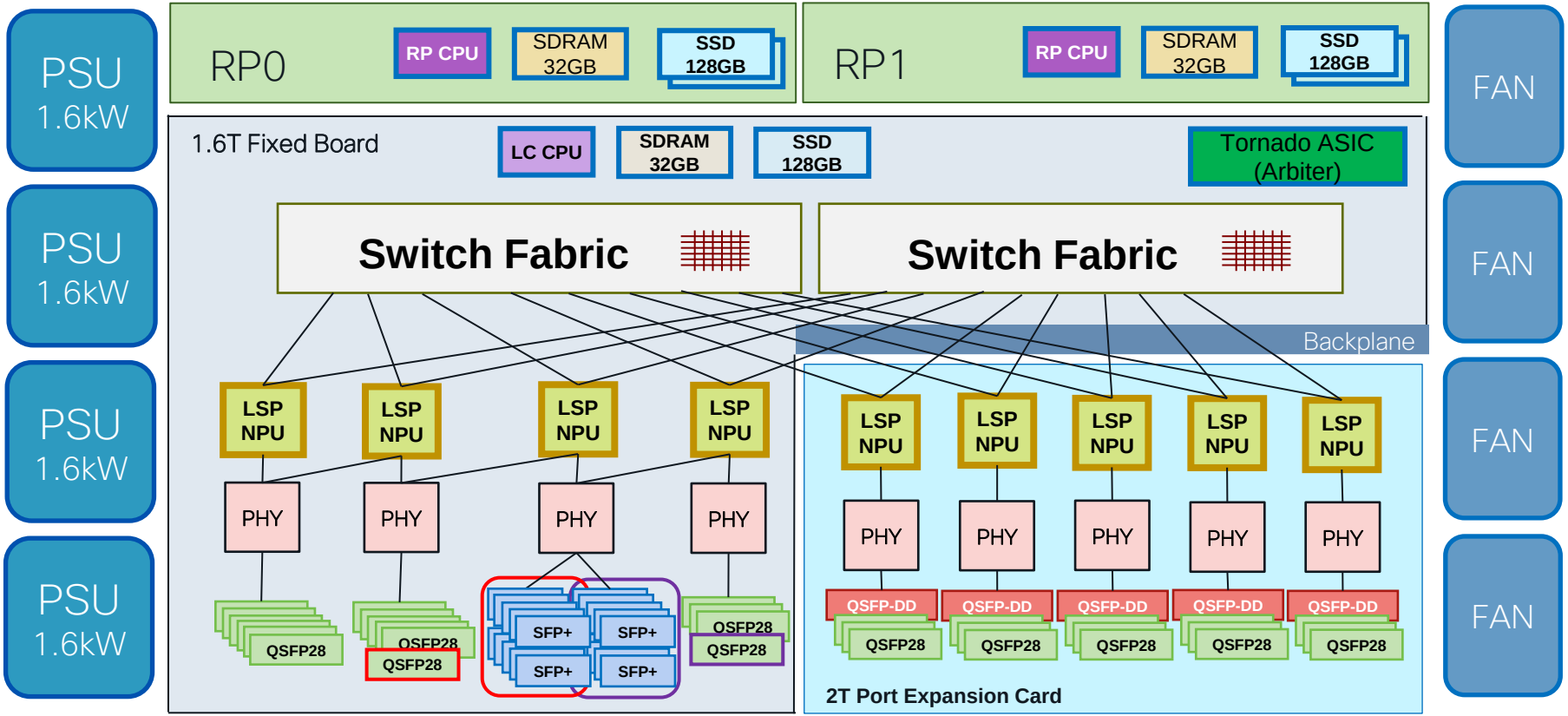
Full feature parity with ASR 9000 platform

Line-rate MACSec support

G.8273.2 Class C, Telecom Profiles, FlexE support for 5G Mobile Edge

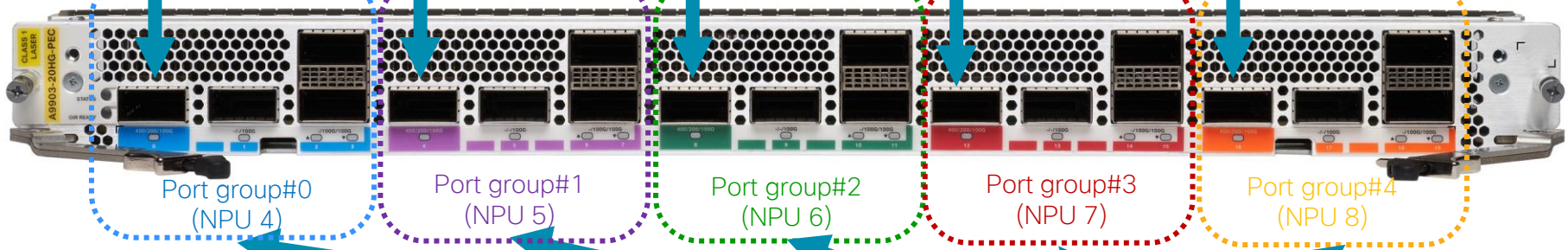
1GE support via either option: nV satellite solution, Smart SFP on 10G/25G ports (roadmap)

ASR 9903 Hardware Architecture



ASR 9903 – 2T Port Expansion Card

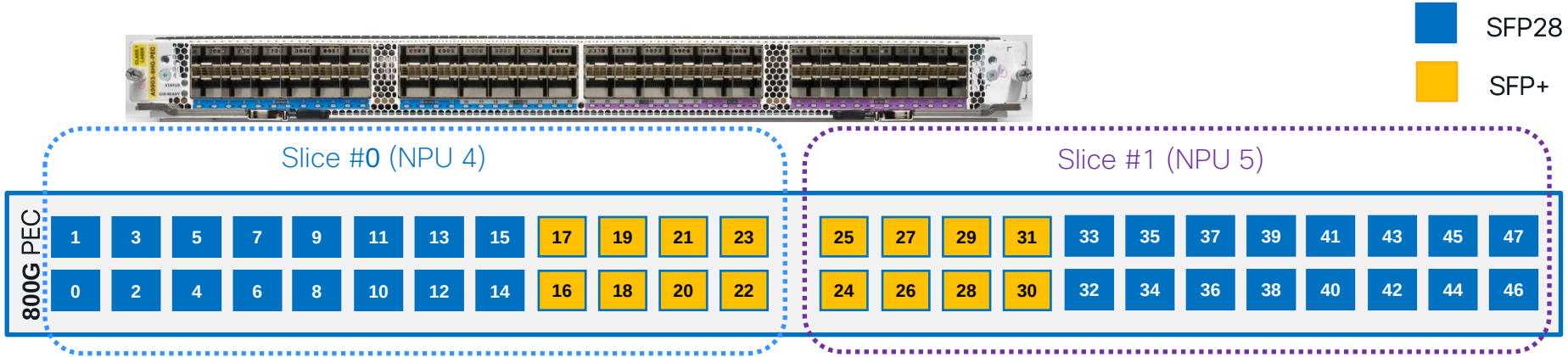
400G/200G/100G Multi-Rate
Ports: 0/4/8/12/16



- 5 x 400G QSFP-DD ports
- 15 x 100G QSFP28 ports
- 400G/200G/100G/40G/25G/10G support
- Each port group (slice) can be independently configured as:
  - 4x100G (default) –or-
  - 1x400G –or-
  - 1x200G + 2x100G
- Each 100G can break out into 4x25G or 4x10G
- OIR support*

5 Ports Groups (Slices):
1x 400G/200G/100G multi-rate port
& 3x100G ports per Slice

ASR 9903 800G Port Expansion Card



- 32 x 25G SFP28 ports and 16 x 10G SFP+ ports
- 25G ports are dual-rate: 25G (with SFP28)/10G (with SFP+)
- Slices #0 & #1 can be independently configured as:
  - 16 x 25G
  - -or-
  - 24 x 10G
- Ports in the slice are serviced by the one NPU
- Slice can either be in 10G mode or 25G mode only, no mix within 1 slice;
- OIR support*

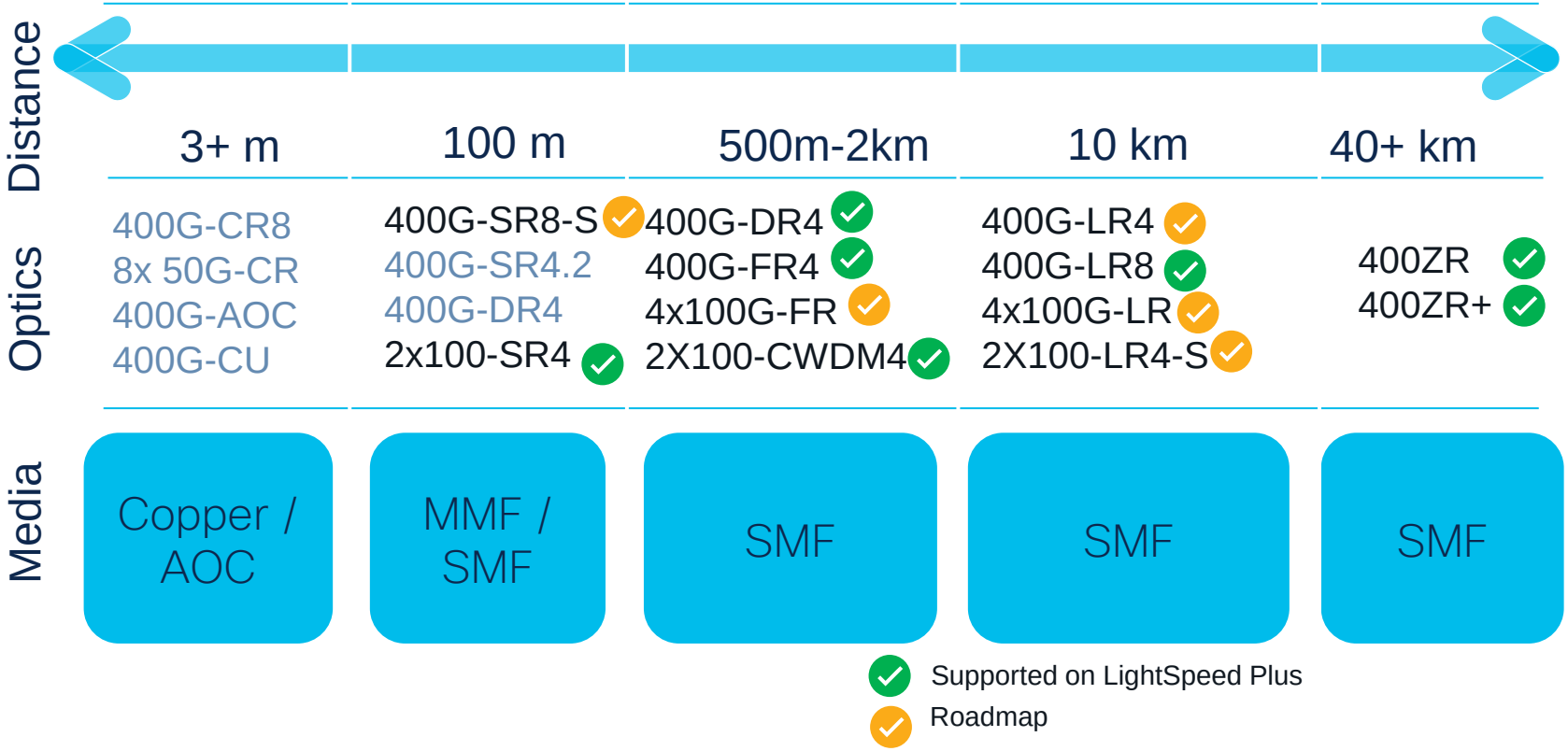
Slice #0 (as an example)

Option	SFP28 Ports (0-15)	SFP+ Ports (16-23)
1 (default)	Ports are enabled and work in 10GE mode (SFP+ optics are used)	Ports are enabled
2	Ports are enabled and work in 25GE mode (SFP28 optics are used)	Ports are disabled

ASR 9000 Compact Platforms Details

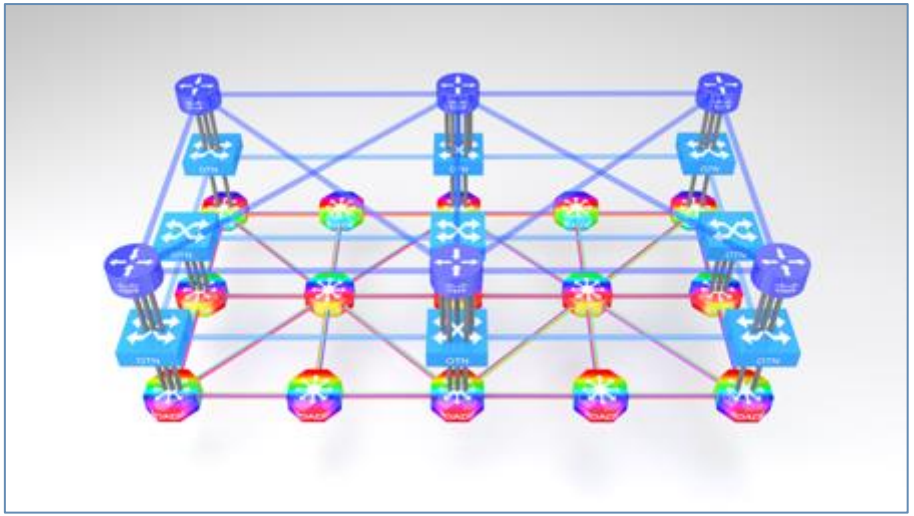
	ASR9901	ASR9902	ASR9903
RP	Single RP	Dual RP	Dual RP
Form-factor	2RU	2RU	3RU
Switch Fabric	Single Fabric	Single Fabric	Redundant Fabric (see Architecture section)
Air flow	Front-to-back	Front-to-back	Front-to-back
Throughput (# of NPU)	120G-456Gbps (2 Tomahawk)	800Gbps (2 LightSpeed Plus)	1.6Tbps / 2.4Tbps / 3.6Tbps (4-6-9 LightSpeed Plus)
Depth	23.62" (600mm)	19" (483mm)	23.62" (600mm)
Power Supply (# of PSU)	AC or DC (2)	AC or DC (2)	AC or DC (4)
QSFP-DD	N/A	Up to 2 (100GE mode)	Up to 5 (400GE / 200GE / 100GE mode)
Ports (1GE/10GE/25GE/40GE/100GE)	• 2x 100GE • 24x 10GE / 1GE • 16x 1GE	• 2x 100GE/40GE QSFP-DD / QSFP28 • 6x 100GE/40GE QSFP28 • 16x 25GE / 10GE • 24x 10GE • Mix of 10GE, 25GE, 40GE, 100GE • Aggregated to 800Gbps	• 16x 100GE + 20x 10GE fixed ports • 20x 100GE or 10x 200GE or 5x 400GE (using 2T PEC - Port Expansion Card) • 32x 25GE/10GE or 48x 10GE (using 800G PEC) • Mix of 10GE, 25GE, 100GE
MACSec support	All 100GE,10GE, 1GE ports	All 100GE, 40GE, 25GE, 10GE, 1GE ports	All 400G, 100GE, 40GE, 25GE, 10GE ports
FlexE	N/A	N/A	Yes
Timing support	Class B	Class C	Class C
1GE support	Up to 40x 1GE ports (SR & LR optics)	Smart SFP on 10GE / 25GE ports	Smart SFP on 10GE / 25GE ports

ASR 9000 Optics Use Cases and QSFP-DD



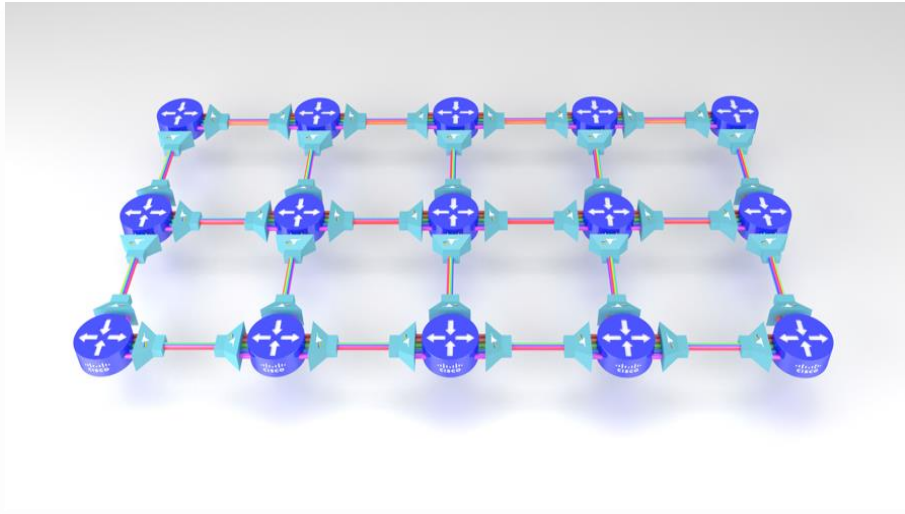
Challenges of traditional network architecture

- Multiple control planes, Each layer treated individually
  - IP/MPLS, GMPLS
  - Wavelength switched optical network (WSON)
  - Spectrum-Switched Optical Network (SSON)
- IP layer dynamically rely on transport layer
  - Adjustable Data Rate, Modulation, Baud Rate, Spectrum, etc…
- Operational life-cycle is complex
  - Planning, Feasibility Compute, Management, Optimization, etc
  - Optical / OTN switching adds complexity



Routed Optical Networking – Simplified Architecture

- Integrate routing, OTN, and Optical transport
  - Network Simplification
  - Reduce Carbon footprint
  - TCO Savings
- Converged IP+Optical network architecture
  - Hop-by-hop IPoEoF solution
  - Remove network complexity
  - Service Agility, fast deployment

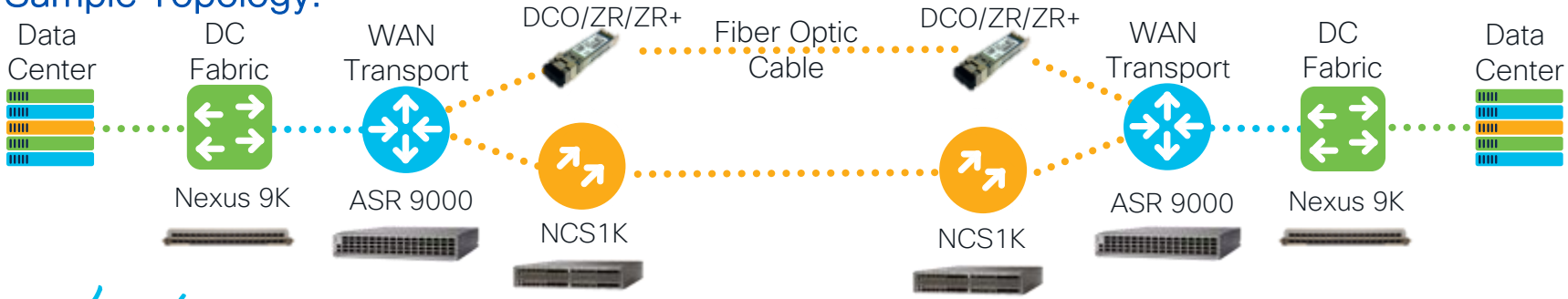


ASR 9000 DCI Opportunity with QSFP-DD

Cisco's Optical and WAN Transport DCI solutions are appropriate for any opportunity involving connecting multiple data centers, colocation or cloud sites together

New DCI Build	Typically consists of end-to-end DCI solution – over dark fiber, wavelength or carrier connectivity
Existing DCI Upgrade	Typically includes refresh of existing equipment or addition of new equipment to increase capacity or redundancy

Sample Topology:

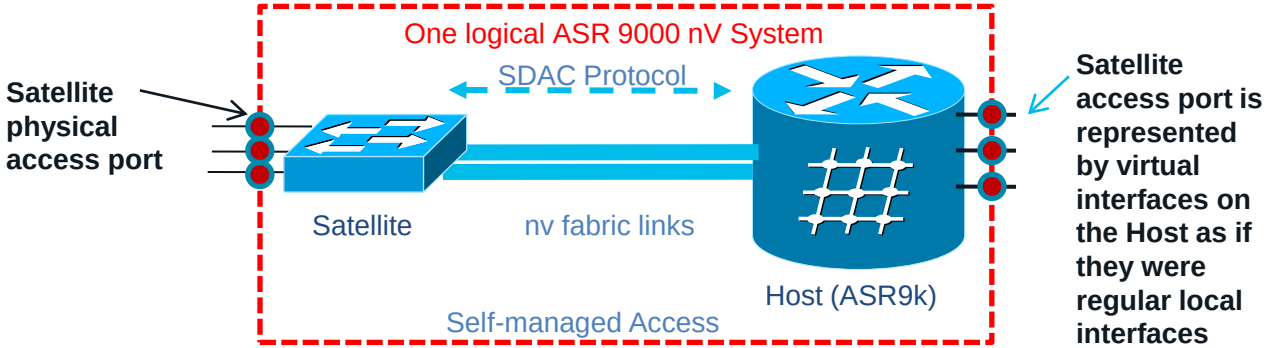


ASR 9000 Product Public Portals

- ASR 9000 Salesconect
  - https://salesconnect.cisco.com/c/r/salesconnect/index.html#/program/PAGE-10319
- ASR 9000 Sharepoints
  - https://cisco.sharepoint.com/sites/MIGRoutingPMTMECommunications/SitePages/ASR%209000-Platform/Cisco-ASR-9000-Routers.aspx
- ASR 9000 Optics Matrix
  - https://tmgmatrix.cisco.com/?si=ASR9000
- ASR 9000 Power Calculator
  - http://tools.cisco.com/cpc

nV Satellite

nV Satellite Overview
(Carrier class Virtual Chassis Fabric system)



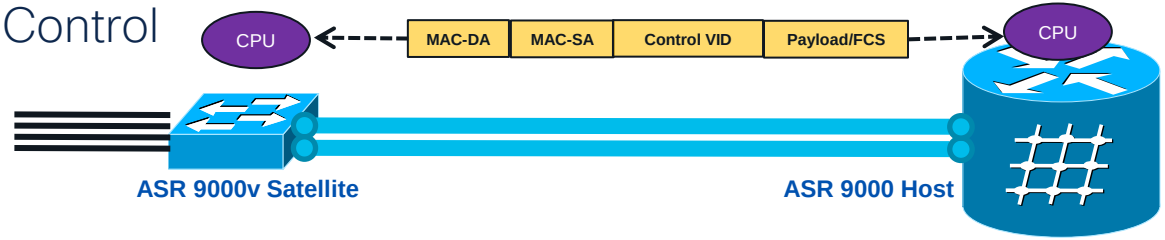
From end user point of view, satellite looks/feels/works like a ASR9K “remote or virtual” line card. The interfaces on the satellite looks/feels/works the same as the interfaces on the local ASR9K line cards

From end user point of view, Host and associated satellites is one virtual Router system.

Satellite is plug-n-play, zero touch configuration

ASR 9000v “Satellite”

Host-satellite operation: Control



- Discovery
  - Like CDP
- Heartbeat
  - One per second
- TCP control connection

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show nv satellite protocol discovery interface Bundle-
Ether 3 brief
Sat Dec 14 02:43:00.278 EST
Interface       Sat-ID  Status                          Discovered links
--------------  ------  ------------------------------  ----------------------
BE3             100     Satellite Ready                 Te0/1/0/3, Te0/1/1/3

RP/0/RSP0/CPU0:rasr9000-2w-a#show tcp brief | include 10.100.111.100
Sat Dec 14 02:47:59.152 EST
0x1002e004 0x6000000d    0      0  10.100.111.1:17514     10.100.111.100:13680   ESTAB

RP/0/RSP0/CPU0:rasr9000-2w-a#show nv satellite protocol control satellite 100 brief
Sat Dec 14 02:48:36.020 EST
Sat-ID  IP Address    Protocol state  Channels
------  ------------  --------------  ----------------------------------
100     10.100.111.100  Connected       Ctrl, If-Ext L1, If-Ext L2, X-link,
                                        VICL, Soft Reset, Inventory,
                                        EnvMon, Alarm, Platform

RP/0/RSP0/CPU0:rasr9000-2w-a#  show nv satellite status satellite 100 brief
Sat Dec 14 02:59:56.752 EST
Sat-ID  Type      IP Address    MAC address     State
------  --------  ------------  --------------  --------------------------------
100     asr9000v  10.100.111.100  8478.ac01.349c  Connected (Stable)
```

ASR 9000 nV Technology Overview

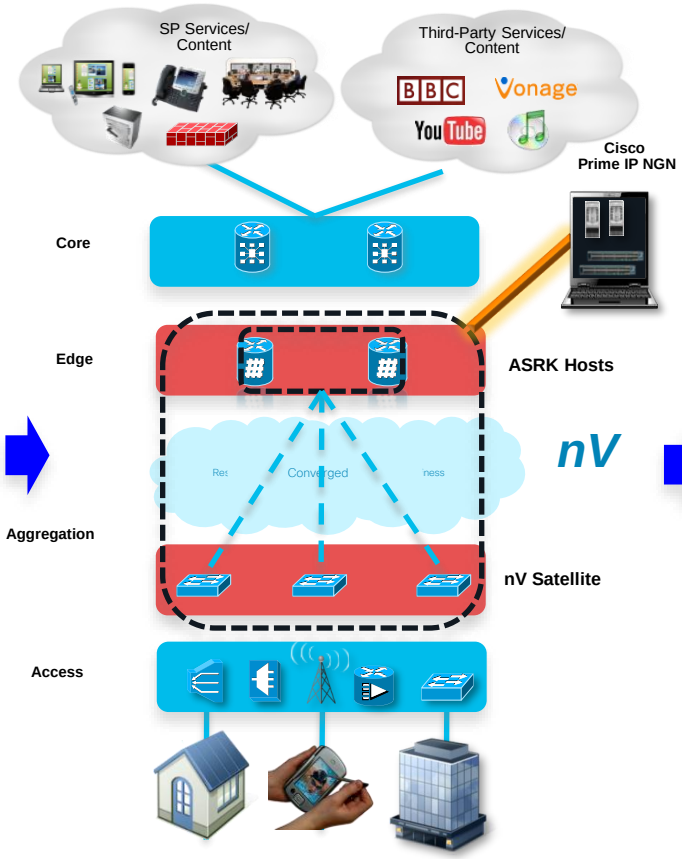
Before: nV Technology

Each device managed separately.

Inconsistent features between edge and aggregation.

Inconsistent service outages upon device failure.

Port scale limited to chassis.



After: nV Technology

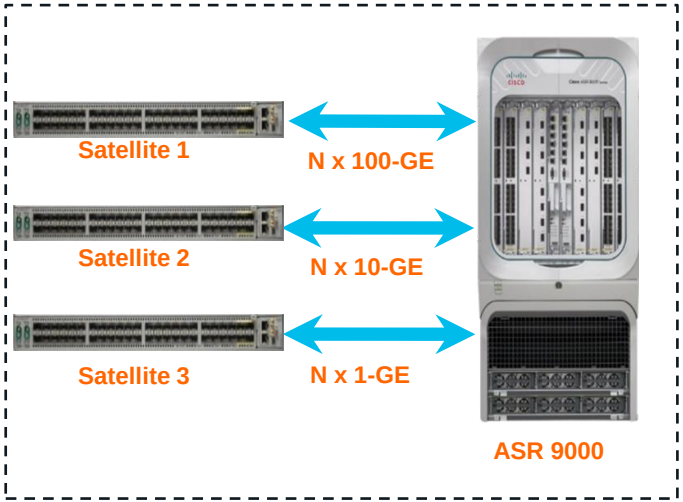
Edge and aggregation managed as one virtual system

Single release vehicle offering feature consistency.

Offers OPEX reduction

Reduced protocol complexity between edge and aggregation

The nV Satellite solution



A single logical switch / router built by interconnecting a **Host** router (ASR9K) and 1 or more smaller switches referred as "**satellites**"

Supported satellite HW: asr9000v and ncs5000

Port Density	• 400+ GigE and TenGE ports per slot, n x slot scale per chassis • Utilize Terabit scale LC capacity with flexible 1G/10G fan out and any choice of low/high power optics including DWDM and LR / ER optics
Opex Reduction	• Single point of chassis + services management - Fewer IP addresses, fewer routing adjacencies • Zero-touch provisioning, plug-n-play installation, network-wide image mgmt
Access network simplification	• Single auto-provisioned virtual switch / router system • Protocols run directly on the Hosts – Lesser protocol instances across the network
Feature consistency	•All current and upcoming IOS-XR features become immediately available on the access box at the same time. •Pre-qualified, pre-tested network level solution- Faster product qualification and faster time to revenue
New network capabilities	•New multi-chassis access redundancy •Better end to end QoS due to awareness of downstream congestion •Better NPU features from "on-loading" on the Host

ASR 9000v “Satellite”

Configuration view

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show running-config
.
vrf NV-MGMT
 description NV SATELLITE MANAGEMENT
 address-family ipv4 unicast
 !
!
.
control-plane
 management-plane
  inband
   interface Bundle-Ether3
    allow TFTP
   !
.
interface Bundle-Ether3
 vrf NV-MGMT
 ipv4 point-to-point
 ipv4 unnumbered Loopback100
 nv
  satellite-fabric-link satellite 100
   remote-ports GigabitEthernet 0/0/0-29
  !
.
interface Loopback100
 vrf NV-MGMT
 ipv4 address 10.100.111.1 255.255.255.0
!
.
```

Satellite control VRF [optional]

Allow TFTP for satellite upgrade

The host-satellite bundle (ICL)

Host control/TCP address [optional]

```
interface GigabitEthernet100/0/0/9
 mtu 4484
 service-policy input DSCP
 service-policy output DSCP
 vrf TRAFFIC
 ipv4 address 192.10.1.1 255.255.255.0
 negotiation auto
!
.
interface TenGigE0/1/0/3
 description ASR9000v
 bundle id 3 mode on
!
.
interface TenGigE0/1/1/3
 description ASR9000v
 bundle id 3 mode on
!
.
nv
 satellite 100
  type asr9000v
  serial-number CAT1702U20H
  description r9000v-1y-a
  ipv4 address 10.100.111.100
 !
!
.
```

Sample satellite interface config

Host-satellite bundle links

Satellite control configuration

nV Satellite: Monitoring & troubleshooting

cisco Live!

Monitoring: Basic status check

```
RP/0/RSP0/CPU0:TARDIS# show nv satellite status brief

Sat-ID  Type      IP Address    MAC address     State
------  --------  ------------  --------------  --------------------------------
100     ncs5002   10.0.100.1       c472.95a6.2003  Connected

RP/0/RSP0/CPU0:TARDIS# show nv satellite status satellite 100
-------------
Satellite 100
-------------
  Status: Connected (Stable)
  Type: ncs5002
  MAC address: c472.95a6.2003
  IPv4 address: 10.0.100.1 (auto, VRF: **nVSatellite)
  Serial Number: FOC1920R0V7
  Remote version: Compatibility Unknown (no local version)
    FPGA: 1.0
    XR: 600.1
  Received candidate fabric ports:
    None (channel down)
  Configured satellite fabric links:
    HundredGigE0/1/0/0
    ------------------
      Status: Satellite Ready
      Remote ports: TenGigE0/0/4-79
```

Monitoring: Discovery protocol status check

```
RP/0/RSP0/CPU0:TARDIS# show nv satellite protocol discovery brief
Interface       Sat-ID  Status                          Discovered links
--------------  ------  ------------------------------  ----------------------
Hu0/1/0/0       100     Satellite ready                 Hu0/1/0/0
Hu0/1/0/1       100     Satellite ready                 Hu0/1/0/1
```

```
RP/0/RSP0/CPU0:TARDIS#show nv satellite protocol discovery interface HundredGigE 0/1/0/0
Interface HundredGigE0/1/0/0
----------------------------
  Interface Status: Probing for satellites

  Satellite ID: 100
  Status: Satellite ready
  Host IPv4 Address: 100.100.100.101
  Satellite IPv4 Address: 100.100.100.100
  Vendor: 1.3.6.1.4.1.9.12.3.1.3.1705,
  Serial Id: FOC1920R0V7
  Remote ID: 50331907
  Remote MAC address: c472.95a6.2056
  Chassis MAC address: c472.95a6.2003
```

Monitoring: Control protocol status check

```
RP/0/RSP0/CPU0:TARDIS#sh nv satellite protocol control brief
Sat-ID  IP Address    Protocol state  Channels
------  ------------  --------------  ----------------------------------------
100     100.100.100.100  Connected    Ctrl, If-Ext L1, If-Ext L2, X-link,
                                      VICL, DevMgmt, Inventory, EnvMon,
                                      Alarm, Password, Topology,
RP/0/RSP0/CPU0:TARDIS#sh nv satellite protocol control
Satellite 100
-------------
  Status: Connected since 2015/10/28 16:11:35.930
  IP address: 100.100.100.100 (VRF: default)
  Channels:
    Control (0)
    -----------
      Channel status: Open
      Messages sent: 15 (15 control), received: 14 (14 control)
      Version: 0

    Interface Extension Layer 1 (1)
    -------------------------------
      Channel status: Open
      Messages sent: 7 (5 control), received: 459 (3 control)
      Version: 0

    Interface Extension Layer 2 (2)
    -------------------------------
      Channel status: Open
      Messages sent: 15 (5 control), received: 615 (3 control)
      Version: 0
```

Monitoring: Check Satellite Inventory

```
RP/0/RSP0/CPU0:TARDIS(admin)#show inventory
...

NAME: "fantray SAT100/FT0/SP", DESCR: "Cisco NCS 5002 Series Router Fan Back"
PID: NCS-5002-FN-BK, VID: N/A, SN: N/A

NAME: "fantray SAT100/FT1/SP", DESCR: "Cisco NCS 5002 Series Router Fan Back"
PID: NCS-5002-FN-BK, VID: N/A, SN: N/A

NAME: "power-module SAT100/PM0/SP", DESCR: "Cisco NCS 5000 Series Router power AC 650W Back"
PID: NC5K-PAC-650W-BK=, VID: V01, SN: LIT1919198Z

NAME: "power-module SAT100/PM1/SP", DESCR: "Cisco NCS 5000 Series Router power AC 650W Back"
PID: NC5K-PAC-650W-BK=, VID: V01, SN: LIT1919199H

NAME: "Satellite Chassis NCS5002 ID 100", DESCR: "80-Port 10 GE + 4-Port 100GE NCS5002 Chassis"
PID: NCS-5002, VID: V00, SN: FOC1920R0V7
```

<- Satellite chassis, fan tray, power module, optics included within normal ASR9K inventory reports

```
RP/0/RSP0/CPU0:TARDIS(admin)#show inventory  rack
  Rack                Chassis PID           S/N
  ----                ------------      ----------
  0                   ASR-9904-AC       FOX1739G94Y
  100                 NCS-5002          FOC1920R0V7
```

<-

Each satellite appears as a new rack within the ASR9K Inventory (rack # -> satellite ID)

Debugging on the satellite (When all else fails … telnet into the satellite)

In rare cases, when all else fails, you may need to telnet in-band into the satellite for debug. Note that, if discovery and IP connectivity is the issue under debug, then a direct console may be needed to the NCS 5002 device. Once in, the satellite will show normal XR console. Only "basic" show commands should be used here to avoid conflicts with nV host driven state and XR config mode is blocked.

```
RP/0/RSP0/CPU0:TARDIS#telnet vrf **nVSatellite 10.0.100.1
Trying 10.0.100.1...
Connected to 10.0.100.1.
Escape sequence is '^^q'.

User Access Verification

Username: root
Password:
RP/0/RP0/CPU0:Satellite#
```

Direct in-band telnet from ASR9K
IOS-XR prompt to satellite
assigned IP address
("`telnet satellite <n>`" also supported)

Now use "XR" show
Commands on the satellite

```
RP/0/RP0/CPU0:Satellite#show ipv4 interface brief | i Hundred
Thu Oct 29 03:52:47.798 UTC
HundredGigE0/0/1/0             unassigned      Down                Down
HundredGigE0/0/1/1             unassigned      Down                Down
HundredGigE0/0/1/2             10.0.100.1      Up                  Up
```

ICL on the satellite side with the IP inherited from unnumbered association to a loopback interface

Always fetch "show tech-support sdac" first before the traces wrap!

Debug scenarios – On NCS5k satellite

```
RP/0/RP0/CPU0:Satellite#show sdac protocol discovery
Thu Oct 29 04:15:33.383 UTC
Interface HundredGigE0/0/1/2
--------------------------------
  Status:                 Discovered
  Satellite IPv4 Address: 10.0.100.1
  Host IPv4 Address:      10.0.0.1
  Vendor:                 Cisco System Inc
  Remote ID:              1216
  Remote MAC address:     001d.e5e9.2a4c
  Chassis MAC address:    e4c7.223f.0ba6

RP/0/RP0/CPU0:Satellite#show sdac protocol control
Thu Oct 29 04:16:54.775 UTC
Host: e4c7.223f.0ba6
--------------------
  Status: Connected since 03:42:02.513 UTC Thu Oct 29 2015
  IP address: 10.0.0.1
  Channels:
    Interface Extension Layer 1 (1)
    -------------------------------
      Channel status: Open
      Messages sent: 55293 (199 control), received: 2169 (300 control).
      .........
```

Check the satellite's view of SDAC discovery and control State machines

Debug scenarios – On NCS5k satellite

```
RP/0/RP0/CPU0:Satellite#show sdac internal intf-mode
Thu Oct 29 04:19:32.987 UTC
Interface Modes:
Interface  | Poss Acc Port | Act Acc Port | Poss ICL | Act ICL
-----------+---------------+--------------+----------+--------
Hu0/0/1/0  | N             | N            | Y        | N
Hu0/0/1/1  | N             | N            | Y        | N
Hu0/0/1/2  | N             | N            | Y        | Y
Hu0/0/1/3  | N             | N            | Y        | N
Te0/0/0/0  | Y             | N            | N        | N
Te0/0/0/1  | Y             | N            | N        | N
Te0/0/0/10 | Y             | Y            | N        | N
```

Show mapping mode per interface
Indicates which is Active ICL , active access port etc

Show mapping from satellite Front ports to ICL port

```
RP/0/RP0/CPU0:Satellite#show l2vpn xconnect
Thu Oct 29 04:25:53.018 UTC
Legend: ST = State, UP = Up, DN = Down, AD = Admin Down, UR = Unresolved,
        SB = Standby, SR = Standby Ready, (PP) = Partially Programmed

XConnect                    Segment 1                       Segment 2
Group      Name       ST    Description            ST       Description            ST
-------------------------   ------------------------------  ------------------------------
2_6        2_6        UP    Te0/0/0/4              UP       Hu0/0/1/2.6            UP
-------------------------------------------------------------------------------------------
```

Here access port is Te0/0/0/4 which is mapped to ICL Hu0/0/1/2. 6 is the internal sat vlan that is added to switch the packets between satellite and host.

Agenda

- ✓ System Architecture: System anatomy & health
- ➢ Operating System & Configuration: IOS-XR & configuration models
- ➢ Control, Management, Security: Processing of control & exceptions
- ➢ Transit Packet/Frame Journey: Life of L3/L2 unicast/multicast
- ➢ MPLS Operation: Processing, forwarding & L3/L2 service operation
- ➢ Troubleshooting: Diagnostics, counters, drops, and packet capture

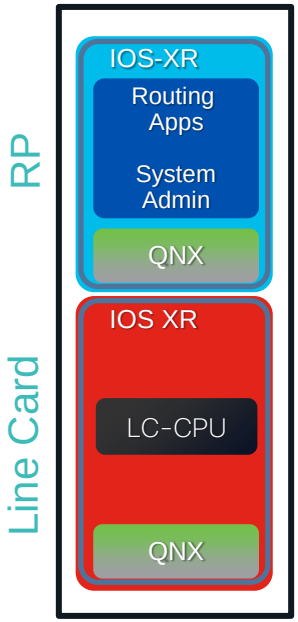
2 Operating System & Configuration

IOS-XR Architecture

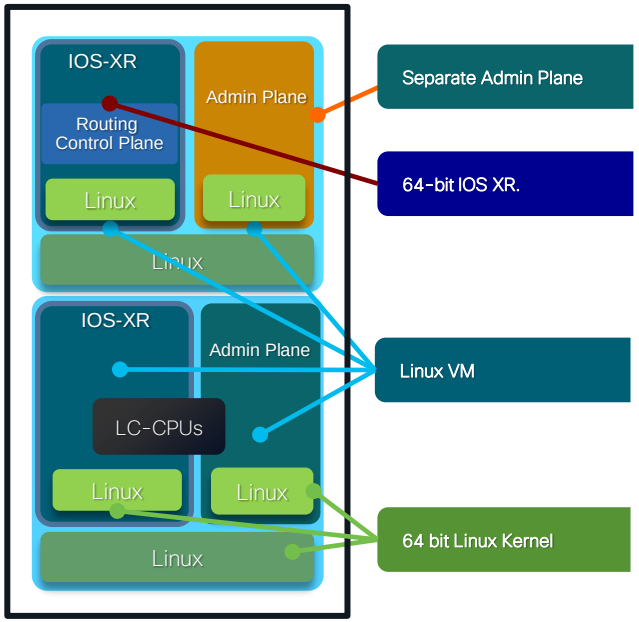
IOS XR Architecture Overview

cXR-Classis IOSXR with 32 Bit; eXR-extended IOSXR with 64 Bit

- IOS XR Exists in two flavors
- 32-bit in XR12k, CRS, ASR9000
  - QNX-based
  - No virtualization
- 64-bit in ASR9000, NCS 5500, NCS 5000, NCS 1000 and in NCS 6000
  - Linux based
  - Larger addressable memory
  - Separation Networking OS and Admin Plane
  - Virtualization: VM or Container
  - ASR9000 Running with VMs



"Classic" IOS XR 32 Bit

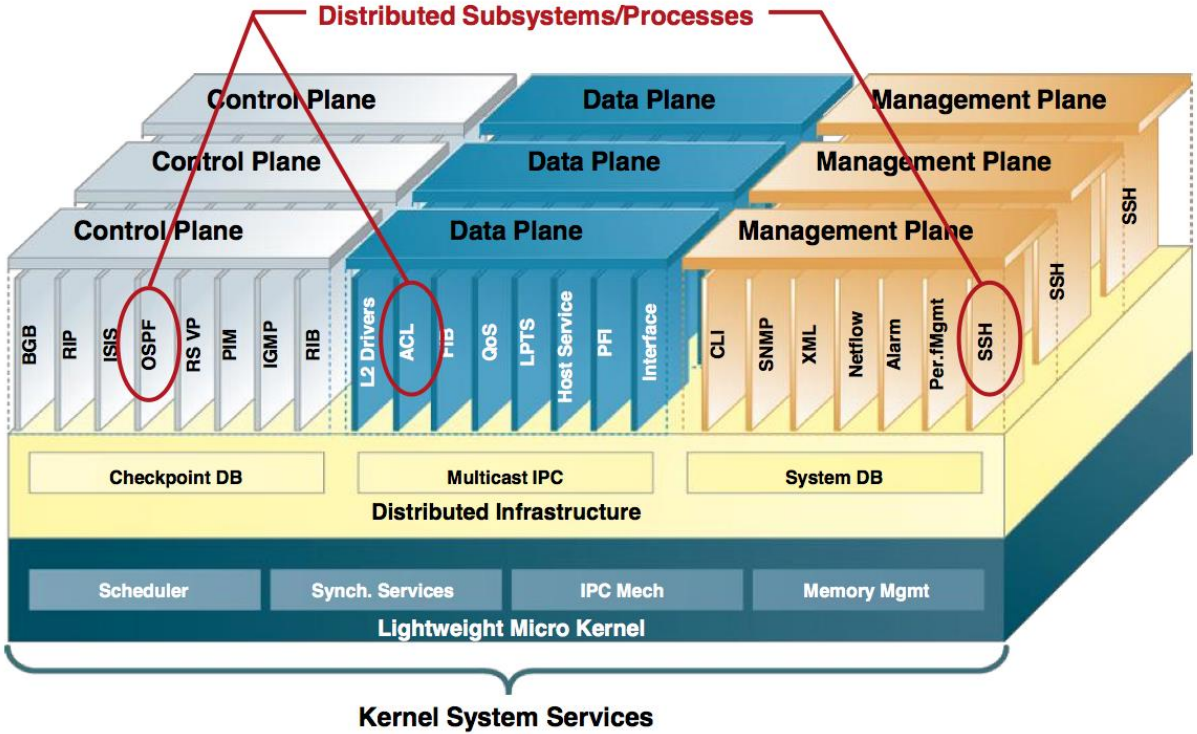


Linux-based Virtualized IOS XR 64 Bit

IOS-XR 32 Bit Operating System

32-bit Kernel and distributed processes, addressable memory 4GB

- Micro-kernel
  - QNX kernel
- Restartable processes
  - A process may start/terminate based on configuration
  - Scheduler keeps track of process starts/spawning/priority/path
  - A process can crash/restart/patched
- Distributed processing
  - Processes run on RP and LC CPU's



IOS-XR 64 bit Architecture Overview

Cisco developed packages for core network functions (BGP, MPLS, etc.)

Yocto packages for standard Linux tools and libraries (bash, python, tcpdump, etc.).

XR VM
(Control Plane)

Admin VM
(Admin Plane)

Runs processes responsible to perform system diags, monitor env. variables, and manage hardware components

First VM to be booted by the Hypervisor, responsible for the start and maintenance of the Control Plane VM

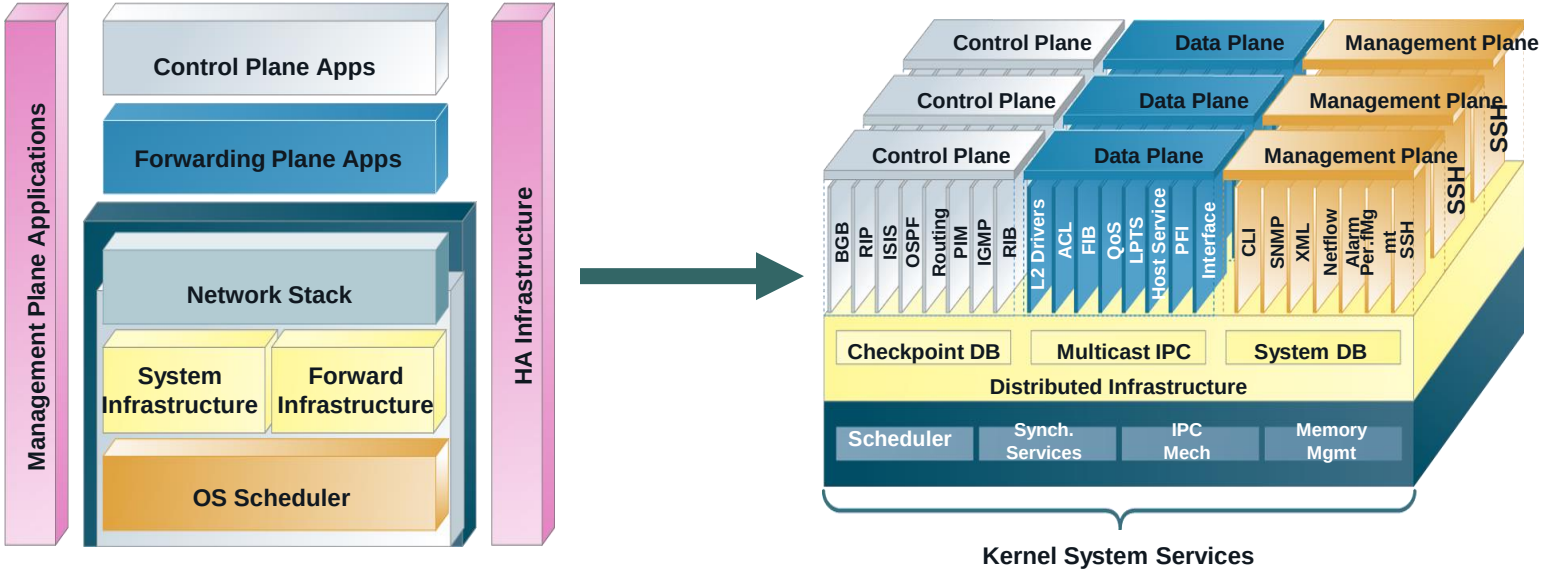
Hypervisor

64-bit Host OS

Routing Processor

From IOS to IOS-XR

From monolithic to micro-kernel



- Monolithic Kernel
- Centralized Infrastructure
- Centralized applications

- Micro Kernel
- Distributed Infrastructure
- Distributed applications

IOS-XR Operating System

Kernel and distributed processes

- Process instances
  - On RSP (e.g. BGP), LC (e.g. ARP, IGMP), or both
  - Mandatory or configuration-dependent (e.g. OSPF)
- Restartable
  - Means also patchable!

```
RP/0/RSP0/CPU0:rasr9000-2w-b#run uname -a
Thu Nov 21 12:07:01.637 EST
QNX node0_RSP0_CPU0 6.4.0 2012/07/25-07:54:16PDT asr9k-rsp x86

RP/0/RSP0/CPU0:rasr9000-2w-b#show processes location 0/RSP0/CPU0
Thu Nov 21 12:13:37.588 EST
JID   TID CPU Stack pri state        TimeInState    HR:MM:SS:MSEC   NAME
1       1   0    0K   0  Running      0:00:00:0000   25:25:11:0433 procnto-smp-instr
1       2   1    0K   0  Running      0:00:00:0000   25:22:47:0535 procnto-smp-instr
1       3   2    0K   0  Running      0:00:00:0000   25:19:47:0529 procnto-smp-instr
1       4   3    0K   0  Ready        0:00:00:0000   25:18:23:0806 procnto-smp-instr
1       5   2    0K  10 Receive      0:03:18:0905    0:00:00:0473 procnto-smp-instr
.
RP/0/RSP0/CPU0:rasr9000-2w-b#show processes location 0/RSP0/CPU0 | include bgp
Thu Nov 21 12:27:52.017 EST
143     1   1   36K  10 Receive     25:35:35:0474    0:00:00:0035 bgp_policy_reg_agent
143     2   1   36K  10 Receive     25:43:46:0465    0:00:00:0000 bgp_policy_reg_agent
143     3   0   36K  10 Receive     25:43:46:0345    0:00:00:0000 bgp_policy_reg_agent
1048    1   2  388K  10 Receive      0:00:25:0434    0:00:00:0253 bgp
.
RP/0/RSP0/CPU0:rasr9000-2w-b#show processes location 0/0/CPU0 | include arp
Thu Nov 21 12:28:07.105 EST
116     1   3   56K  10 Receive      0:00:14:0000    0:00:00:0336 arp
116     2   1   56K  10 Receive     25:31:48:0919    0:00:00:0020 arp
116     3   0   56K  10 Receive     25:53:18:0643    0:00:00:0000 arp
116     4   3   56K  10 Receive     25:53:18:0637    0:00:00:0001 arp
116     5   1   56K  10 Receive      0:00:00:0802    0:00:00:0290 arp
327     1   2   16K  10 Receive      0:00:19:0397    0:00:00:0066 slarp_lite
```

Processes

On an 8-core CPU [RSP880]

```
RP/0/RSP1/CPU0:ASR9006-2w-a.PE2_CE1#admin show processes location 0/RSP1/CPU0
Thu Feb 16 16:27:59.815 EST
  JID    TID CPU Stack pri state        TimeInState    HR:MM:SS:MSEC   NAME
  1        1   0    0K   0  Running        0:00:00:0000 1003:07:39:0085 procnto-smp-instr
  1        2   1    0K   0  Running        0:00:00:0000 1003:07:00:0840 procnto-smp-instr
  1        3   2    0K   0  Running        0:00:00:0000 1002:11:09:0034 procnto-smp-instr
  1        4   3    0K   0  Running        0:00:00:0000 1000:35:38:0237 procnto-smp-instr
  1        5   4    0K   0  Running        0:00:00:0000  994:28:11:0049 procnto-smp-instr
  1        6   5    0K   0  Running        0:00:00:0000  989:59:09:0679 procnto-smp-instr
  1        7   6    0K   0  Ready          0:00:00:0000  980:21:04:0983 procnto-smp-instr
  1        8   7    0K   0  Running        0:00:00:0000  978:19:01:0459 procnto-smp-instr
  1        9   6    0K  10  Receive        0:00:00:0079    0:17:56:0824 procnto-smp-instr
  1       10   6    0K  10  Receive        0:02:59:0387    0:21:13:0432 procnto-smp-instr
  1       11   6    0K  10  Receive     1004:36:29:0567    0:00:00:0000 procnto-smp-instr
  1       12   6    0K  10  Receive     1004:39:39:0705    0:00:00:0000 procnto-smp-instr
  1       13   6    0K  10  Running        0:00:00:0000    0:28:36:0391 procnto-smp-instr
  1       15   6    0K  10  Receive        0:00:00:0072    0:00:00:0033 procnto-smp-instr
  1       17   3    0K  10  Receive     1004:36:29:0567    0:00:00:0000 procnto-smp-instr
  1       19   6    0K  10  Receive     1004:36:29:0567    0:00:00:0000 procnto-smp-instr
  1       20   4    0K  10  Receive     1004:39:39:0705    0:00:00:0000 procnto-smp-instr
  1       22   6    0K  10  Receive        0:00:00:0000    0:28:57:0786 procnto-smp-instr
  1       26   5    0K  10  Receive     1004:41:13:0004    0:00:00:0000 procnto-smp-instr
 97        1   6   96K  10  Receive     1004:42:27:0059    0:00:00:0017 wd-critical-mon
 97        3   3   96K  63  Nanosleep      0:00:00:0240    0:00:00:0007 wd-critical-mon
.
```

Distributed Processing

Process example

```
RP/0/RSP0/CPU0:rasr9k-1y#show processes location 0/0/cpu0
Wed Mar 13 12:52:30.904 UTC
JID    TID CPU Stack pri state        TimeInState    HR:MM:SS:MSEC   NAME
.
250      1   2   24K  10 Receive        0:01:03:0295    0:00:00:0099 l2snoop
250      2   0   24K  10 Receive      485:14:51:0408    0:00:00:0000 l2snoop
250      3   0   24K  10 Sigwaitinfo  485:14:51:0407    0:00:00:0000 l2snoop
.
RP/0/RSP0/CPU0:rasr9k-1y#show processes location 0/rsp0/cpu0
Wed Mar 13 13:05:04.550 UTC
JID    TID CPU Stack pri state        TimeInState    HR:MM:SS:MSEC   NAME
.
306      1   2   24K  10 Receive        0:01:51:0885    0:00:00:0046 l2snoop
306      2   2   24K  10 Receive      485:31:56:0557    0:00:00:0000 l2snoop
306      3   1   24K  10 Sigwaitinfo  485:31:56:0557    0:00:00:0000 l2snoop
.
RP/0/RSP0/CPU0:rasr9k-1y#show processes location 0/rsp1/cpu0 | include snoop
Wed Mar 13 13:11:50.386 UTC
306      1   2   24K  10 Receive        0:00:37:0754    0:00:00:0043 l2snoop
306      2   0   24K  10 Receive      485:36:33:0246    0:00:00:0000 l2snoop
306      3   3   24K  10 Sigwaitinfo  485:36:33:0245    0:00:00:0000 l2snoop

RP/0/RSP0/CPU0:rasr9k-1y#show processes distribution l2snoop
Wed Mar 13 13:13:18.810 UTC
3 processes found
NODE              PID           JID   #THR  TYPE  PROGRAM
0/RSP0/CPU0       274643        306      3  RP    l2snoop
0/RSP1/CPU0       266401        306      3  RP    l2snoop
0/0/CPU0          233611        250      3  LC    l2snoop
```

Running the Needed Processes

Process example

```
RP/0/RSP0/CPU0:rasr9k-1y#show processes location 0/rsp0/cpu0 | include eigrp
Wed Mar 13 17:04:38.971 UTC
203    1    3    36K  10 Receive      489:30:31:0725    0:00:00:0029 eigrp_policy_reg_agent
203    2    3    36K  10 Receive      489:30:32:0243    0:00:00:0000 eigrp_policy_reg_agent
203    3    2    36K  10 Receive      489:30:32:0193    0:00:00:0000 eigrp_policy_reg_agent
RP/0/RSP0/CPU0:rasr9k-1y#configure
Wed Mar 13 17:04:43.082 UTC
RP/0/RSP0/CPU0:rasr9k-1y(config)#router eigrp 100
RP/0/RSP0/CPU0:rasr9k-1y(config-eigrp)#commit
Wed Mar 13 17:04:50.633 UTC
RP/0/RSP0/CPU0:Mar 13 17:04:50.681 : rmf_svr[386]: %HA-REDCON-1-STANDBY_NOT_READY : standby card is NOT
ready
RP/0/RSP0/CPU0:Mar 13 17:04:50.682 : rmf_svr[386]: %HA-REDCON-1-BACKUP_NOT_READY : backup process groups
between 0/RSP0/CPU0 and 0/RSP1/CPU0 are NOT ready
RP/0/RSP1/CPU0:Mar 13 17:04:50.681 : rmf_svr[386]: %HA-REDCON-1-STANDBY_NOT_READY : standby card is NOT
ready
RP/0/RSP0/CPU0:Mar 13 17:04:50.684 : rmf_svr[386]: %HA-REDCON-1-BACKUP_READY : backup process groups
between 0/RSP0/CPU0 and 0/RSP1/CPU0 are ready
RP/0/RSP0/CPU0:Mar 13 17:04:50.684 : rmf_svr[386]: %HA-REDCON-1-BACKUP_NOT_READY : backup process groups
between 0/RSP0/CPU0 and 0/RSP1/CPU0 are NOT ready
RP/0/RSP0/CPU0:Mar 13 17:04:50.687 : rmf_svr[386]: %HA-REDCON-1-BACKUP_READY : backup process groups
between 0/RSP0/CPU0 and 0/RSP1/CPU0 are ready
RP/0/RSP0/CPU0:Mar 13 17:05:00.682 : rmf_svr[386]: %HA-REDCON-1-STANDBY_READY : standby card is ready
```

Running the Needed Processes

Process example - continued

```
RP/0/RSP0/CPU0:rasr9k-1y(config-eigrp)#do show processes location 0/rsp0/cpu0 | include eigrp
Wed Mar 13 17:05:32.977 UTC
203    1    3    36K  10 Receive      489:31:25:0732    0:00:00:0029 eigrp_policy_reg_agent
203    2    3    36K  10 Receive      489:31:26:0249    0:00:00:0000 eigrp_policy_reg_agent
203    3    2    36K  10 Receive      489:31:26:0199    0:00:00:0000 eigrp_policy_reg_agent
1002   1    2   124K  10 Receive        0:00:40:0441    0:00:00:0147 eigrp
1002   2    1   124K  10 Receive        0:00:42:0400    0:00:00:0000 eigrp
1002   3    2   124K  10 Receive        0:00:42:0387    0:00:00:0000 eigrp
1002   4    3   124K  10 Sigwaitinfo    0:00:42:0104    0:00:00:0000 eigrp
1002   5    2   124K  10 Receive        0:00:42:0169    0:00:00:0000 eigrp
1002   6    2   124K  10 Receive        0:00:42:0047    0:00:00:0003 eigrp
1002   7    1   124K  10 Receive        0:00:42:0031    0:00:00:0000 eigrp
1002   8    0   124K  10 Receive        0:00:42:0030    0:00:00:0000 eigrp
1002   9    2   124K  10 Receive        0:00:42:0030    0:00:00:0000 eigrp
1002   10   0   124K  10 Receive        0:00:42:0030    0:00:00:0000 eigrp
1002   11   2   124K  10 Receive        0:00:42:0029    0:00:00:0000 eigrp
RP/0/RSP0/CPU0:rasr9k-1y(config-eigrp)#no router eigrp 100
RP/0/RSP0/CPU0:rasr9k-1y(config)#commit
Wed Mar 13 17:05:46.305 UTC
RP/0/RSP0/CPU0:rasr9k-1y(config)#do show processes location 0/rsp0/cpu0 | include eigrp
Wed Mar 13 17:05:50.441 UTC
203    1    3    36K  10 Receive      489:31:43:0186    0:00:00:0029 eigrp_policy_reg_agent
203    2    3    36K  10 Receive      489:31:43:0704    0:00:00:0000 eigrp_policy_reg_agent
203    3    2    36K  10 Receive      489:31:43:0654    0:00:00:0000 eigrp_policy_reg_agent
```

IOS-XR Operating System

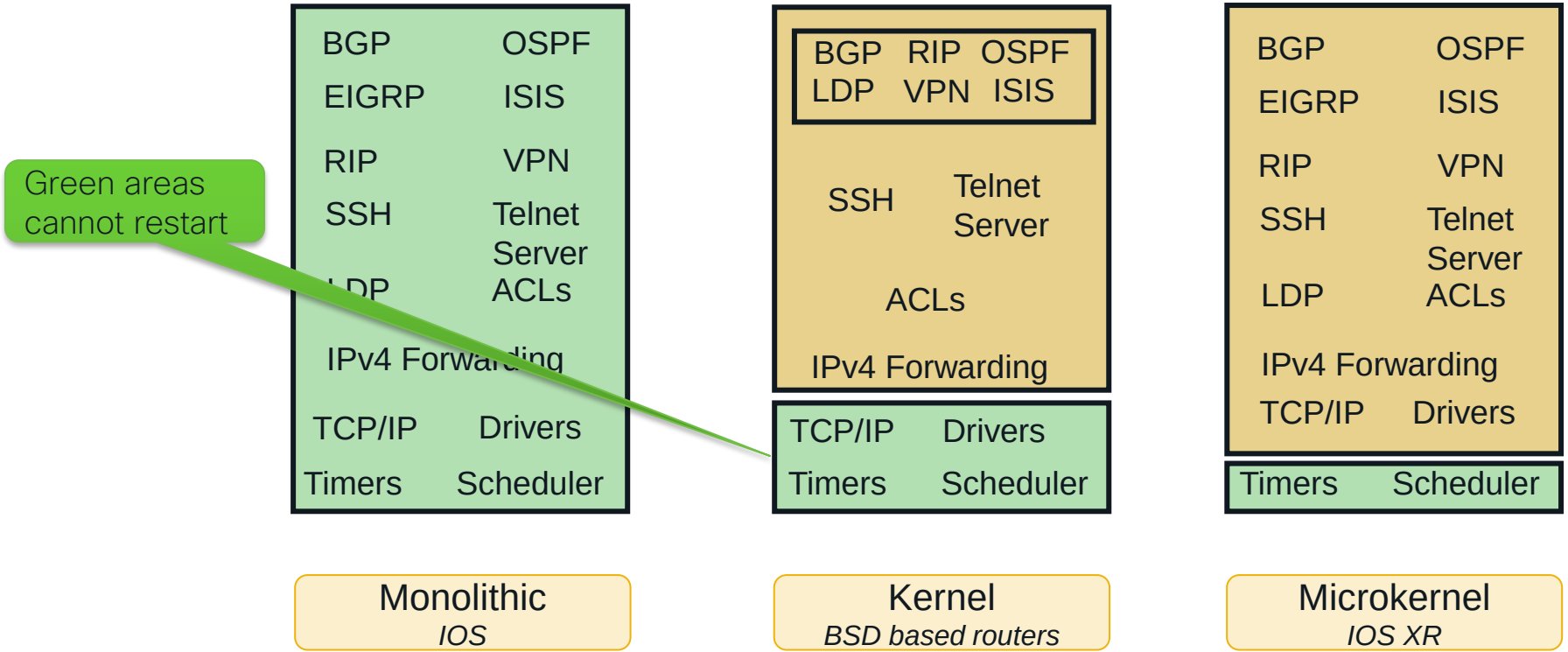
RSP and LC CPU's

- To monitor: a CPU on every card
- Instances of processes running on RSP and LC CPU's

```
RP/0/RSP0/CPU0:rasr9k-1y#show processes cpu location 0/RSP0/CPU0 |
exclude " 0%      0%       0%"
Wed Nov 28 01:36:52.203 UTC

CPU utilization for one minute: 26%; five minutes: 25%; fifteen
minutes: 22%

PID    1Min    5Min    15Min Process
94243    3%      3%       3% spp
254074  23%     22%      19% netio

RP/0/RSP0/CPU0:rasr9k-1y#show processes cpu location 0/0/CPU0 |
exclude " 0%      0%       0%"
Wed Nov 28 01:28:52.281 UTC

CPU utilization for one minute: 46%; five minutes: 48%; fifteen
minutes: 39%

PID    1Min    5Min    15Min Process
45085   22%     23%      22% spp
180316  23%     23%      23% netio
```

Process Restart

Most processes are re-startable



Demonstrating Process Restart

Same Job ID, New Process ID

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show processes bgp
Wed Jan 15 17:58:34.050 EST
                  Job Id: 1048
                     PID: 287056
          Executable path: /disk0/iosxr-routing-
4.2.3.CSCuh52959-1.0.0/0x100305/bin/bgp
              Instance #: 1
              Version ID: 00.00.0000
                 Respawn: ON
           Respawn count: 1
  Max. spawns per minute: 12
            Last started: Thu Jan  2 09:11:18 2014
           Process state: Run
           Package state: Normal
       Started on config: default
            Feature name: ON
                     Tag : default
           Process group: v4-routing
                    core: MAINMEM
                Max. core: 0
                Placement: Placeable
           startup_path: /pkg/startup/bgp.startup
                   Ready: 0.700s
               Available: 85.082s
         Process cpu time: 21.760 user, 2.619 kernel,
24.379 total
JID   TID CPU Stack pri state        TimeInState
1048   1    2  384K  10 Receive        0:00:03:0395
.
```

```
RP/0/RSP0/CPU0:rasr9000-2w-a#process restart bgp
Wed Jan 15 18:03:24.836 EST
RP/0/RSP0/CPU0:Jan 15 18:03:24.874 :
sysmgr_control[65784]: %OS-SYSMGR-4-PROC_RESTART_NAME
: User cisco (con0_RSP0_CPU0) requested a restart of
process bgp at 0/RSP0/CPU0
RP/0/RSP0/CPU0:rasr9000-2w-a#show processes bgp
Wed Jan 15 18:03:28.726 EST
                  Job Id: 1048
                     PID: 3182840
          Executable path: /disk0/iosxr-routing-
4.2.3.CSCuh52959-1.0.0/0x100305/bin/bgp
              Instance #: 1
              Version ID: 00.00.0000
                 Respawn: ON
           Respawn count: 2
  Max. spawns per minute: 12
            Last started: Wed Jan 15 18:03:24 2014
           Process state: Run (last exit due to SIGTERM)
           Package state: Normal
       Started on config: default
            Feature name: ON
                     Tag : default
           Process group: v4-routing
                    core: MAINMEM
                Max. core: 0
                Placement: Placeable
           startup_path: /pkg/startup/bgp.startup
                   Ready: 0.225s
.
```

Process Dumps

Where?

```
RP/0/RSP0/CPU0:rasr9000-2w-a#admin show exception
Sat Jan  4 00:15:27.885 EST
 Exception path for choice 1 is not configured or removed
 Exception path for choice 2 is not configured or removed
 Exception path for choice 3 is not configured or removed
 Choice fallback one path = dumper_harddisk:/dumper compress = on filename = <process_name>
 Choice fallback two path = dumper_disk1a:/dumper compress = on filename = <process_name>
 Choice fallback three path = dumper_disk0a:/dumper compress = on filename = <process_name>
 Kernel dump not configured
 Tftp route for kernel core dump not configured
 No config for pakmem tuple
 No config for sparse tuple
 No config for sprsize tuple
 No config for coresize tuple
 No config for memory-threshold tuple
 No config for core-verification tuple

RP/0/RSP0/CPU0:rasr9000-2w-a#dir harddisk:/dumper
Sat Jan  4 00:16:10.138 EST
Directory of harddisk:/dumper
24922       -rw-  216304651   Sat Jan  4 00:16:08 2014  ce_switch.log
24665       -rw-  42408       Tue Nov  5 19:06:35 2013  crashinfo.by.kernel.19070930-173606
24694       -rw-  1586390     Tue Nov  5 19:06:36 2013  kernel_core.by.kernel.19070930-173606.Z
24695       -rw-  1044480     Tue Nov  5 19:06:36 2013  pcds_dump.19070930-173606
24697       -rw-  4813080     Fri Nov  8 17:03:11 2013  first.mpls_lsd_338.node0_RSP0_CPU0.x86.Z
.
```

Process Dumps

On line cards

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show filesystem location 0/1/CPU0 | include lc
Sat Jan  4 00:26:39.269 EST
  1644150784  1635583488                 rw  lcdisk0:
   411041792   410915840                 rw  lcdisk0a:

RP/0/RSP0/CPU0:rasr9000-2w-a#dir lcdisk0:/dumper location 0/0/CPU0
Sat Jan  4 00:29:22.703 EST

Directory of net/node0_0_CPU0/lcdisk0:/dumper

6361        -rw-  3205840     Fri Nov  8 23:00:02 2013  first.cluster_dlm_lc_143.sparse.node0_0_CPU0.ppc.Z
6362        -rw-  7289        Fri Nov  8 23:00:02 2013
first.cluster_dlm_lc_143.sparse.node0_0_CPU0.ppc.txt
6363        -rw-  495281      Fri Nov  8 23:00:08 2013
first.cluster_dlm_lc_143.sparse.node0_0_CPU0.ppc.cpu_info.Z
6370        -rw-  2293471     Fri Nov  8 23:29:30 2013
first.eth_server_57.by.wdsysmon.sparse.node0_0_CPU0.ppc.Z
6365        -rw-  13722       Fri Nov  8 23:00:33 2013  l2fib_mgr_247.sparse.node0_0_CPU0.ppc.txt
6366        -rw-  515048      Fri Nov  8 23:00:36 2013  l2fib_mgr_247.sparse.node0_0_CPU0.ppc.cpu_info.Z
6367        -rw-  5632747     Fri Nov  8 23:00:40 2013  pm_294.sparse.node0_0_CPU0.ppc.Z
6368        -rw-  10818       Fri Nov  8 23:00:41 2013  pm_294.sparse.node0_0_CPU0.ppc.txt
6369        -rw-  543561      Fri Nov  8 23:00:41 2013  pm_294.sparse.node0_0_CPU0.ppc.cpu_info.Z
6371        -rw-  4429        Fri Nov  8 23:29:30 2013
first.eth_server_57.by.wdsysmon.sparse.node0_0_CPU0.ppc.txt
6372        -rw-  311267      Fri Nov  8 23:29:30 2013
first.eth_server_57.by.wdsysmon.sparse.node0_0_CPU0.ppc.cpu_info.Z
```

IOS-XR Components

eXR for Tomahawk, cXR for any

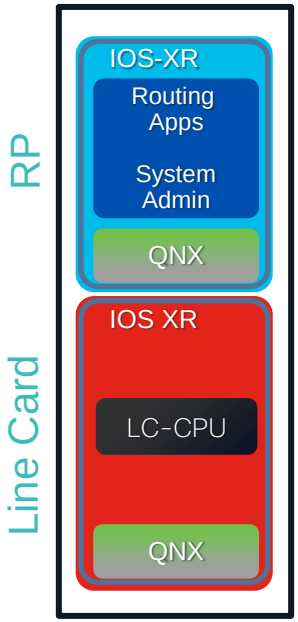
- Starting 6.1.2: eXR is 64-bit, cXR is 32-bit
- eXR only if all hardware supports. cXR for any mix
- eXR: Support for 3rd party app hosting ☺
- Wind River Linux OS

```
RP/0/RP0/CPU0:PE137# run uname -a
Wed Feb  1 08:28:57.332 EST
Linux xr-vm_node0_RP0_CPU0 3.14.23-WR7.0.0.2_standard #1 SMP Fri Jun 17 17:51:29 PDT 2016 x86_64 x86_64 x86_64 GNU/Linux
```

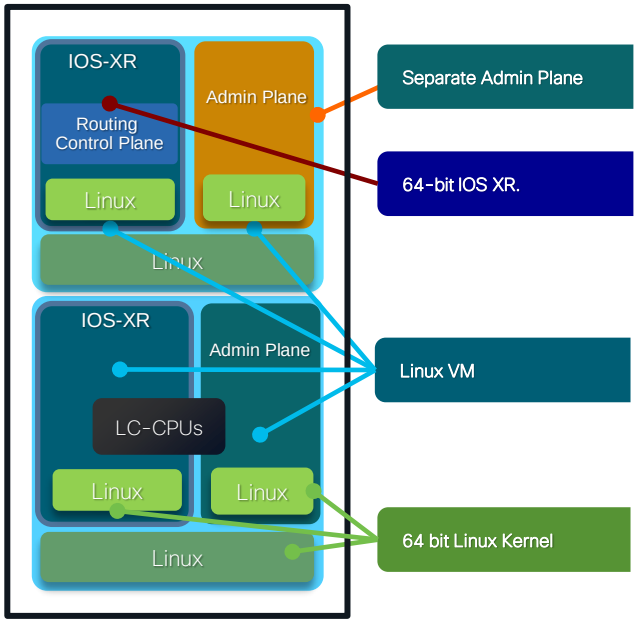
IOS XR 64 Bit: Architecture Overview

cXR-Classis XR-32 Bit eXR-extended XR Release-64 Bit

- IOS XR Exists in two flavors
- 32-bit in XR12k, CRS, ASR9000
  - QNX-based
  - No virtualization
- 64-bit in ASR9000, NCS 5500, NCS 5000, NCS 1000 and in NCS 6000
  - Linux based
  - Larger addressable memory
  - Separation Networking OS and Admin Plane
  - Virtualization: VM or Container
  - ASR9000 Running with VMs

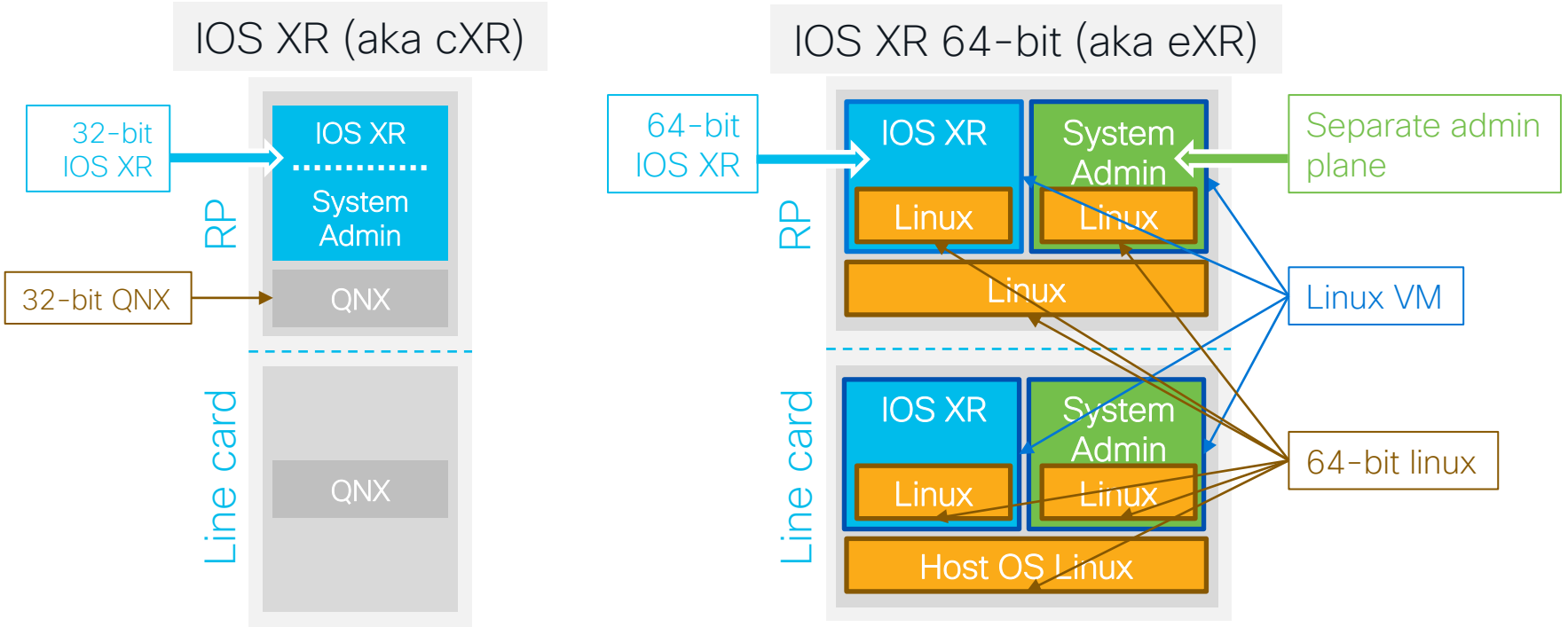


"Classic" IOS XR 32 Bit



Linux-based Virtualized IOS XR 64 Bit

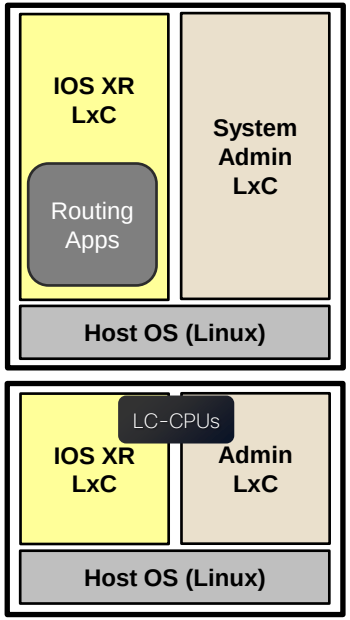
IOS XR Key Components



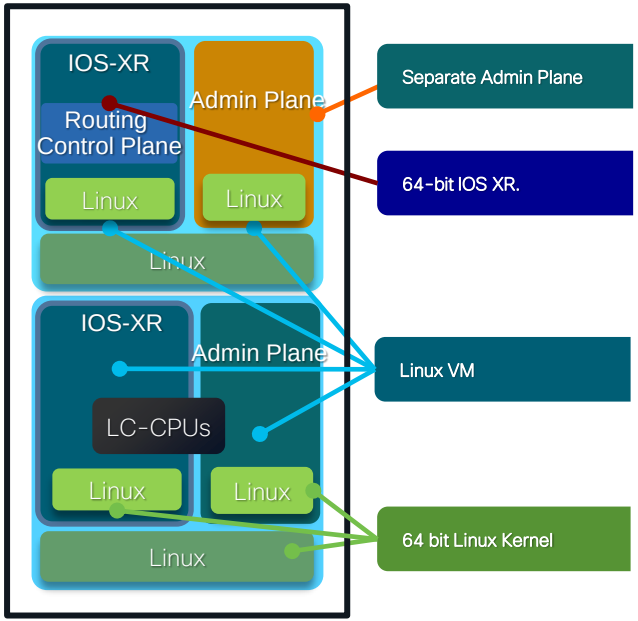
IOS XR 64 Bit

Container versus VM

- IOS XR 64 Bit Can Run with Container or VM
- VM Based 64 Bit XR:
  - In ASR9000, NCS6000
  - VM OS Completely Separated from Host OS
- Container Based 64 Bit XR:
  - In NCS 5500, NCS5000
  - Container OS share the same kernel as Host OS



Container Based
IOS XR 64 Bit
NCS5500, NCS5000



Virtual Machine Based
IOS XR 64 Bit
ASR9000, NCS6000

ASR 9000 Network OS Difference Summary

32 BIT versus 64 BIT

Category	IOS XR 32-bit (Classic XR)	IOS XR 64-bit (Enhanced XR)
Kernel	QNX (32 bit)	Yocto WR Linux (64 bit)
Virtualization	All applications running as different processes. No virtualization.	Two VMs: Admin VM and XR VM on RP/LC CPU
Software Packaging	• PIE based packages. • Special VM image for fresh installation (Turboboot)	• ISO/RPM based packages. • ISO image for bootup and fresh installation. • Flexible Golden ISO image for customer • Offline RPM package management.
Boot Facility	ROMMON: • CLI based • TFTP Network boot	iPXE: • Menu Based • Enables Zero-Touch-Provisioning (ZTP) • TFTP/FTP/SFTP/HTTP/HTTPs

Software Packaging Terminology

Package types

PIE?

Mini?

Package?

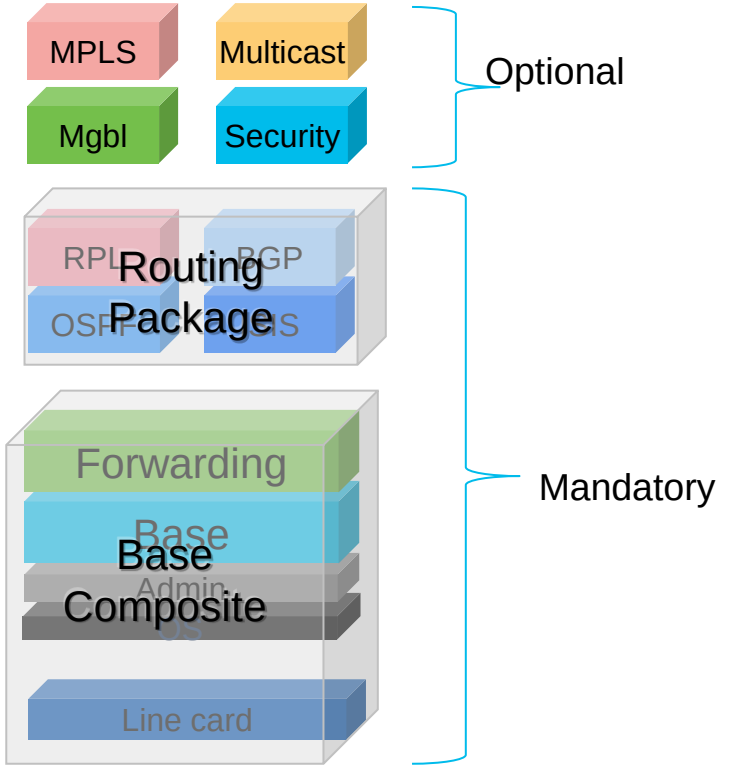
SMU?

32-bit only
64-bit only

Package Terminology

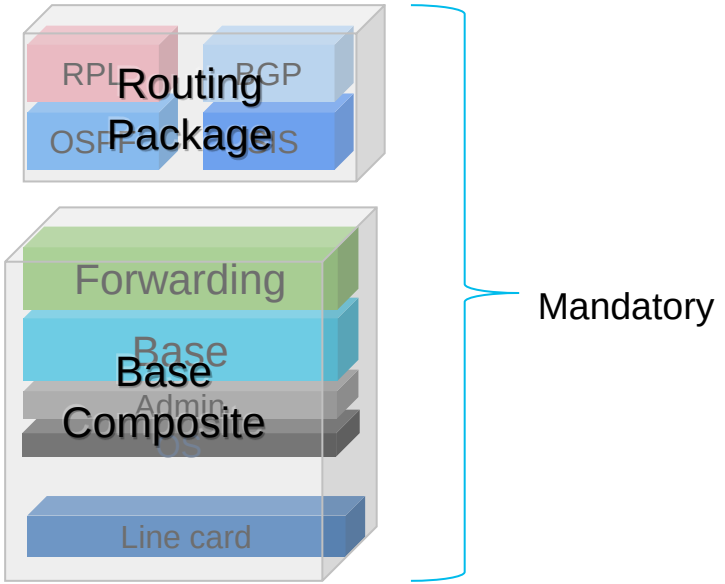
- PIE – Package Installation Envelopes (PIEs) are nonbootable files that contain a single package or a set of packages (called a composite package or bundle). Because the files are nonbootable, they are used to add software package files to a running router.
- RPM – Linux RPM
- SMU – When a PIE file contains software for a specific bug fix, it is called a software maintenance upgrade (SMU).
- Umbrella SMU – When A SMU contains multiple bug fixes
- Service Pack – All SMU's posted to date combined into in a single package, reload required
- SP-GISO – All SMU's posted to date + all base packages combined into a single bootable ISO
- Turboboot – The initial installation of IOS XR SW to disk
- Install Upgrade/Replace – Performing an IOS XR SW Upgrade via the install process
- FPD – Field Programmable Devices

Packages: Mandatory and Optional



- Packages are unique sets of components and represent *potential* units of delivery
- Packages are visible in the code base – "build" infrastructure prevents illegal dependencies between packages
- Packages can be grouped into composites for ease of delivery
- Code base files are organized into components – these are versioned and visible to the development engineer
- Packages can be downloaded from Cisco.com via "tar" files

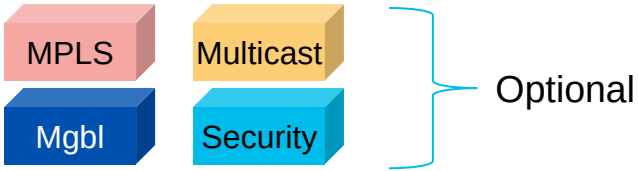
Packages: Mandatory



- The **Cisco IOS XR Unicast Routing Core Bundle** is a Mandatory composite package containing the following packages:

- Forwarding
- Administration
- Base
- Operating system (OS)
- Routing
- Line card drivers

Packages: Optional



Optional packages provide additional features:

- **Manageability** – Support for HTTP, XML, SNMP and other management tools
- **Multicast** – Support for multicast protocols
- **MPLS** – Support for Multiprotocol Label Switching (MPLS)
- **Security** – Support for Secure Sockets Layer (SSL), certificates and other security tools

Note: These are just some of the optional packages available on the ASR9K, there are additional optional SW packages

PIE – Package Installation Envelope

- PIEs are a delivery mechanism for packages
  - Used to deliver
    - Releases – (5.1.3, 5.3.4, 6.1.2)
    - SMUs – Fix for a specific bug
- Mini is a bundle of the mandatory packages
- Includes authentication info
- Installed from admin mode

Reading Installed Packages

Example

```
RP/0/RSP0/CPU0:rasr9000-2w-b#show install active
detail
Sun Jan  4 23:43:14.325 EST
Secure Domain Router: Owner
  Node 0/RSP0/CPU0 [RP] [SDR: Owner]
    Boot Device: disk0:
    Boot Image: /disk0/asr9k-os-mbi-
5.1.3/0x100305/mbiasr9k-rsp3.vm
    Active Packages:
      disk0:asr9k-mpls-px-5.1.3
          disk0:iosxr-mpls-5.1.3
      disk0:asr9k-mgbl-px-5.1.3
          disk0:asr9k-mgbl-supp-5.1.3
          disk0:iosxr-mgbl-5.1.3
      disk0:asr9k-optic-px-5.1.3
          disk0:asr9k-optics-supp-5.1.3
      disk0:asr9k-k9sec-px-5.1.3
          disk0:iosxr-security-5.1.3
          disk0:asr9k-k9sec-supp-5.1.3
      disk0:asr9k-doc-px-5.1.3
          disk0:asr9k-doc-supp-5.1.3
      disk0:asr9k-fpd-px-5.1.3
          disk0:asr9k-fpd-5.1.3
      disk0:asr9k-mini-px-5.1.3
          disk0:asr9k-scfclient-5.1.3
          disk0:asr9k-os-mbi-5.1.3
           disk0:asr9k-cpp-5.1.3
          disk0:asr9k-ce-5.1.3
          disk0:iosxr-ce-5.1.3
          disk0:asr9k-diags-supp-5.1.3
          disk0:iosxr-diags-5.1.3
          disk0:asr9k-fwding-5.1.3
          disk0:iosxr-fwding-5.1.3
          disk0:iosxr-routing-5.1.3
          disk0:iosxr-infra-5.1.3
          disk0:asr9k-base-5.1.3
      disk0:asr9k-mcast-px-5.1.3
          disk0:asr9k-mcast-supp-5.1.3
          disk0:iosxr-mcast-5.1.3

.
.
  Node 0/1/CPU0 [LC] [SDR: Owner]
    Boot Device: mem:
    Boot Image: /disk0/asr9k-os-mbi-5.1.3/lc/mbiasr9k-
lc.vm
    Active Packages:
      disk0:asr9k-mpls-px-5.1.3
          disk0:iosxr-mpls-5.1.3
      disk0:asr9k-optic-px-5.1.3
          disk0:asr9k-optics-supp-5.1.3
      disk0:asr9k-mini-px-5.1.3
          disk0:asr9k-scfclient-5.1.3
          disk0:asr9k-os-mbi-5.1.3
          disk0:asr9k-cpp-5.1.3
          disk0:asr9k-ce-5.1.3
          disk0:iosxr-ce-5.1.3
          disk0:asr9k-diags-supp-5.1.3
          disk0:iosxr-diags-5.1.3
          .
```

SMU and Service Pack Delivery

- SMU is named by release and bugid
  - asr9k-px-5.3.4.CSCvb41169.tar
  - asr9k-px-5.1.3.CSCvc42959.tar

- Service Pack is named by release and SP number
  - asr9k-px-5.1.3.sp10.tar

- Service Pack Golden ISO is named by release and SP number
  - ncs5500-goldenk9-x-7.5.2-SPGISO0001

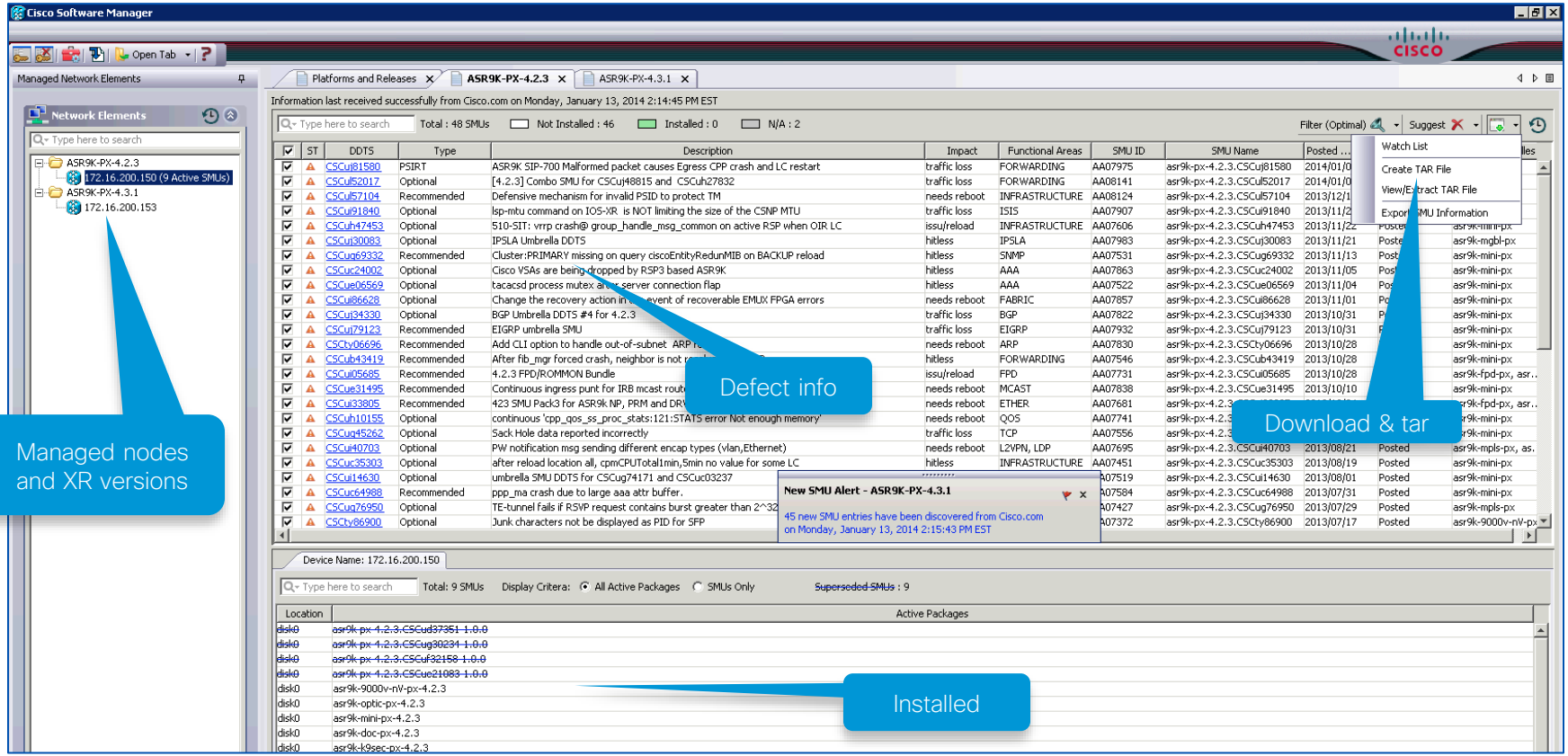
IOS XR ASR 9000 64 Bit Packaging

eXR ISO , Packages & SMU's

Bootable Images		
Minimum Image	asr9k-mini-x64-6.1.2.iso	Core packages: OS, Admin, Forwarding, Modular Services Card, Basic Routing, SNMP, Alarm Correlation
Golden ISO (GISO)	Customized ISO image includes mini ISO + required packages + SMUs + XR config	

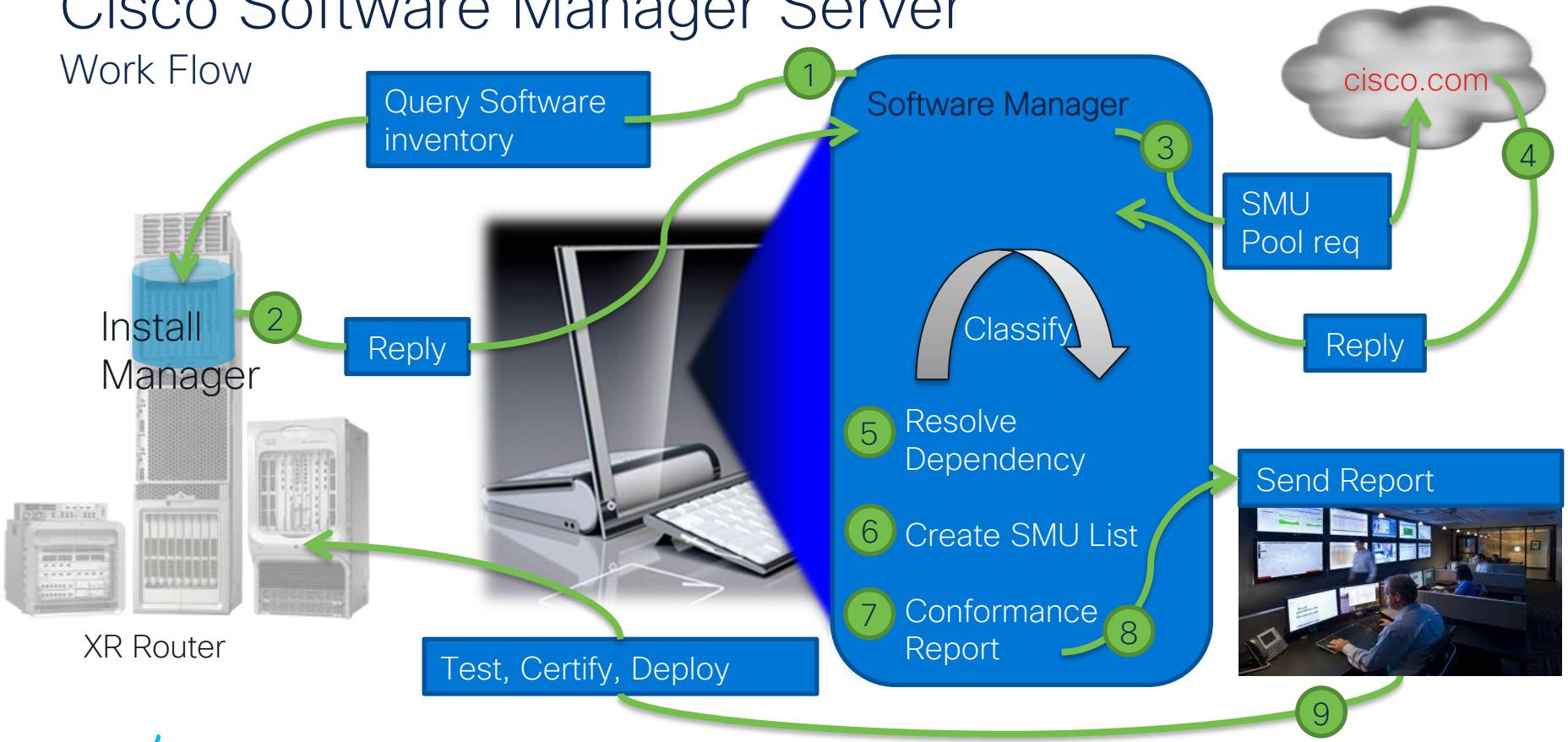
Optional Feature Packages	
asr9k-eigrp-x64-1.0.0.0-r612.x86_64.rpm asr9k-isis-x64-1.1.0.0-r612.x86_64.rpm asr9k-ospf-x64-1.1.0.0-r612.x86_64.rpm asr9k-m2m-x64-2.0.0.0-r612.x86_64.rpm asr9k-mgbl-x64-3.0.0.0-r612.x86_64.rpm asr9k-mpls-te-rsvp-x64-1.2.0.0-r612.x86_64.rpm	asr9k-mpls-x64-2.1.0.0-r612.x86_64.rpm asr9k-mcast-x64-2.0.0.0-r612.x86_64.rpm asr9k-optic-x64-1.0.0.0-r612.x86_64.rpm asr9k-li-x64-1.1.0.0-r612.x86_64.rpm asr9k-k9sec-x64-3.1.0.0-r612.x86_64.rpm

Managing SMU: Cisco Software Manager App



Cisco Software Manager Server

Work Flow



What CSM Does

- Software Management:
  - Automated and Simplified image (releases and SMUs) retrieval, reporting and alerts
  - Pushes image to one or many devices
  - End to end SW management
  - Patch recommendation, and conformance reporting
  - Migration from 32-bit XR to 64-bit XR
- Operations Simplification:
  - Auto-updates: you can schedule installation, pre- and post- installation verifications
  - Easier access to image and patch details (documentation)
  - Multi-platform and multi-OS support
- Inventory Management:
  - Visibility into hardware, cards, slots, S/N, optic types

Solves For:

- Time consuming, manual, laborious, repetitive, error-prone SW installation
- Complicated patch dependencies
- High costs

Big Wins:

- Huge time and resource savings
- Up to 90% time savings on SW upgrades

CSM Server – Supported Cisco Tool

How to download	Download CSM: https://software.cisco.com/download/release.html?mdfid=282423206&softwareid=284777134&release=3.5&relind=AVAILABLE&rellifecycle=&reltype=latest
How to use	**CSM Server Documentation:** https://supportforums.cisco.com/document/13154846/cisco-software-manager-33-overview-documentation **CSM Server Videos:** Introduction to CSM Server: https://youtu.be/isxN08x-mr4 Getting Started with CSM Server: https://www.youtube.com/watch?v=omdpr3uP_b4 ASR9K IOS XR 32 bit to 64 bit Migration using CSM Server: https://youtu.be/RVgR0TdbpVw **CSM Application Video:** https://www.youtube.com/watch?v=PYO2Om-nUKQ
Support forum	https://community.cisco.com/t5/service-providers/ct-p/4441-service-providers

Supported on:

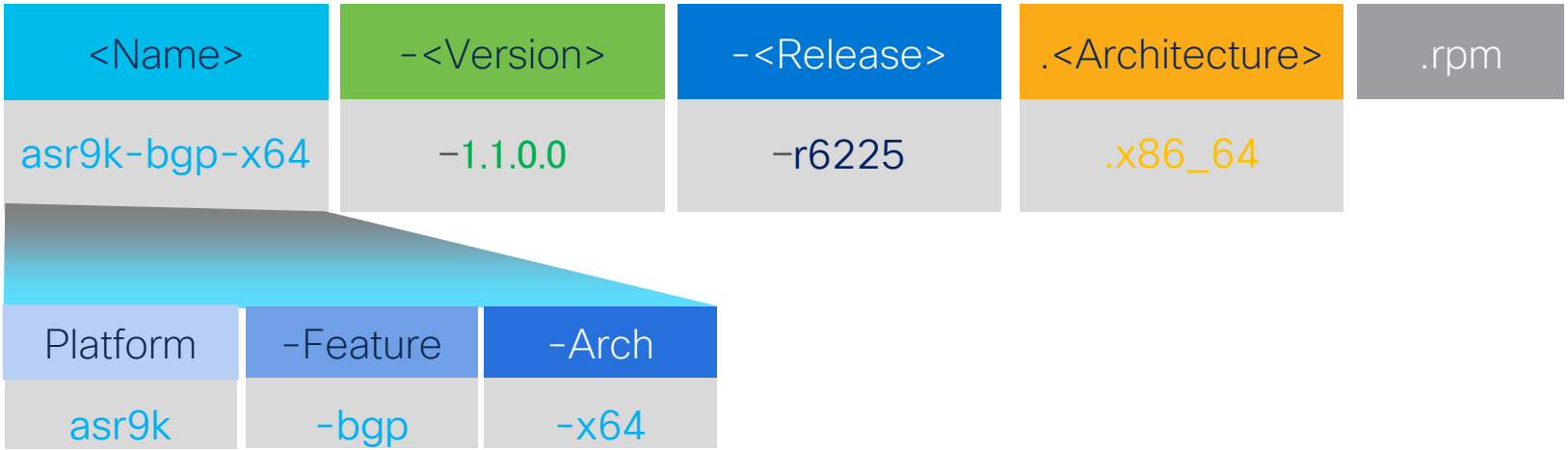
- IOS XR:
  - CRS
  - NCS
  - ASR 9000
- IOS XE:
  - asr9xx
- More coming soon!

IOS XR 64 Bit Package Naming

Typical RPM package name :

asr9k-bgp-x64-1.1.0.0-r6225.x86_64.rpm

Name Convention:



IOS XR 64 Bit Package Naming

Typical XR SMU name :

asr9k-mcast-x64-2.1.0.3-r6225I.CSCvf43058.x86_64.rpm

Typical Admin SMU name :

asr9k-sysadmin-mgbl-6.2.25.13-r6225I.CSCvf09972.x86_64.rpm

Typical ISO name:

asr9k-mini-x64-6.2.25.iso

Typical SP-GISO name:

asr9k-x64-goldenk9-x-7.5.2-SPGISO0001

Golden ISO Solution Overview (GISO)

64 Bit IOS XR

- Golden ISO is a customized ISO which customers can build offline out of the mini ISO by using the Cisco Released Golden ISO build script
- When the System is booted up with the Golden ISO, additional SMU's & Optional packages present in the Golden ISO will be auto installed
- The XR config if present in the Golden ISO will be auto applied (ASR 9000 only).

- Golden ISO holds additional files apart from files present in MINI ISO.

    - Third-party(Yocto pkg/SMU) for host, calvados and XR
    - Cisco SMU's for Host,
    - Cisco SMU's for Calvados
    - Cisco pkg/smu for XR.
    - XR configuration. (ASR9K-eXR)

Golden ISO Use Cases

ASR 9000 64 Bit IOS XR

- cXR(Classic XR) to eXR Migration (ASR9K):
  - Golden ISO allows customers to migrate their router software from Classic XR to eXR
  - customer can build their own Golden ISO with relevant XR configs and boot the system
  - Golden ISO will make the migration simpler
  - Retains the desired XR configurations and all the applicable SMU/Optional packages installed

- First time router deployment with latest software.
- Software Disaster recovery.
- System Upgrade (SU).

Golden ISO supported platforms (eXR):

- ASR9K-X64
- NCS1K
- NCS5K
- NCS5500

ASR9922 cXR to eXR Migration GISO script (Golden ISO)

ASR 9000 64 Bit IOS XR

```
root@virl:/home/6225# ./gisobuild.py -i /home/dpothier/6225/asr9k-mini-x64-6.2.25.iso -r /home/dpothier/6225/ -l v1 –m
Golden ISO build process starting...

System requirements check [PASS]
/dev/loop1
Platform: asr9k Version: 6.2.25
Scanning repository [/home/6225]...

Building RPM Database...
Total 5 RPM(s) present in the repository path provided in CLI
/dev/loop1
Following XR x86_64 rpm(s) will be used for building Golden ISO:

    (+) asr9k-isis-x64-1.2.0.0-r6225.x86_64.rpm
    (+) asr9k-mgbl-x64-2.0.0.0-r6225.x86_64.rpm
    (+) asr9k-k9sec-x64-2.2.0.0-r6225.x86_64.rpm
    (+) asr9k-mpls-x64-2.0.0.0-r6225.x86_64.rpm
    (+) asr9k-mpls-te-rsvp-x64-1.3.0.0-r6225.x86_64.rpm
/dev/loop1

    ...RPM compatibility check [PASS]

Building Golden ISO...
Summary .....
XR rpms:
    asr9k-isis-x64-1.2.0.0-r6225.x86_64.rpm
    asr9k-mgbl-x64-2.0.0.0-r6225.x86_64.rpm
    asr9k-k9sec-x64-2.2.0.0-r6225.x86_64.rpm
    asr9k-mpls-x64-2.0.0.0-r6225.x86_64.rpm
    asr9k-mpls-te-rsvp-x64-1.3.0.0-r6225.x86_64.rpm

    ...Golden ISO creation SUCCESS.

Golden ISO Image Location: /home/6225/asr9k-goldenk9-x64.iso-6.2.25.v1
Detail logs: /home/6225/Giso_build.log-2019-01-16:18:52:45.421533
```

execute script on external server (has requirements)

	gisobuild.py options
-i	path to 64-bit mini-x iso
-r	path to rpm's + smu's
-l	label version
-m	build the migration tar from 32-bit cXR to 64-bit eXR

directory with rpm's + any smu's included in golden iso tar file

```
root@virl:/home/dpothier/6225#  tar -tf asr9k-goldenk9-x64-migrate_to_eXR.tar-6.2.25.v1
boot/
boot/certs/
boot/certs/crl.der
boot/certs/Root_Certificate_Store.bin
boot/certs/CertFile
boot/initrd.img
boot/signature.initrd.img
boot/bzImage
EFI/boot/
EFI/boot/grub.cfg
EFI/boot/grub.efi
asr9k-goldenk9-x64.iso-6.2.25.v1
```

creates Golden ISO Image: asr9k-goldenk9-x64.iso-6.2.25.v1

creates Migration tar: asr9k-goldenk9-x64-migrate_to_eXR.tar-6.2.25.v1

New install commands

64 Bit IOS XR (eXR)

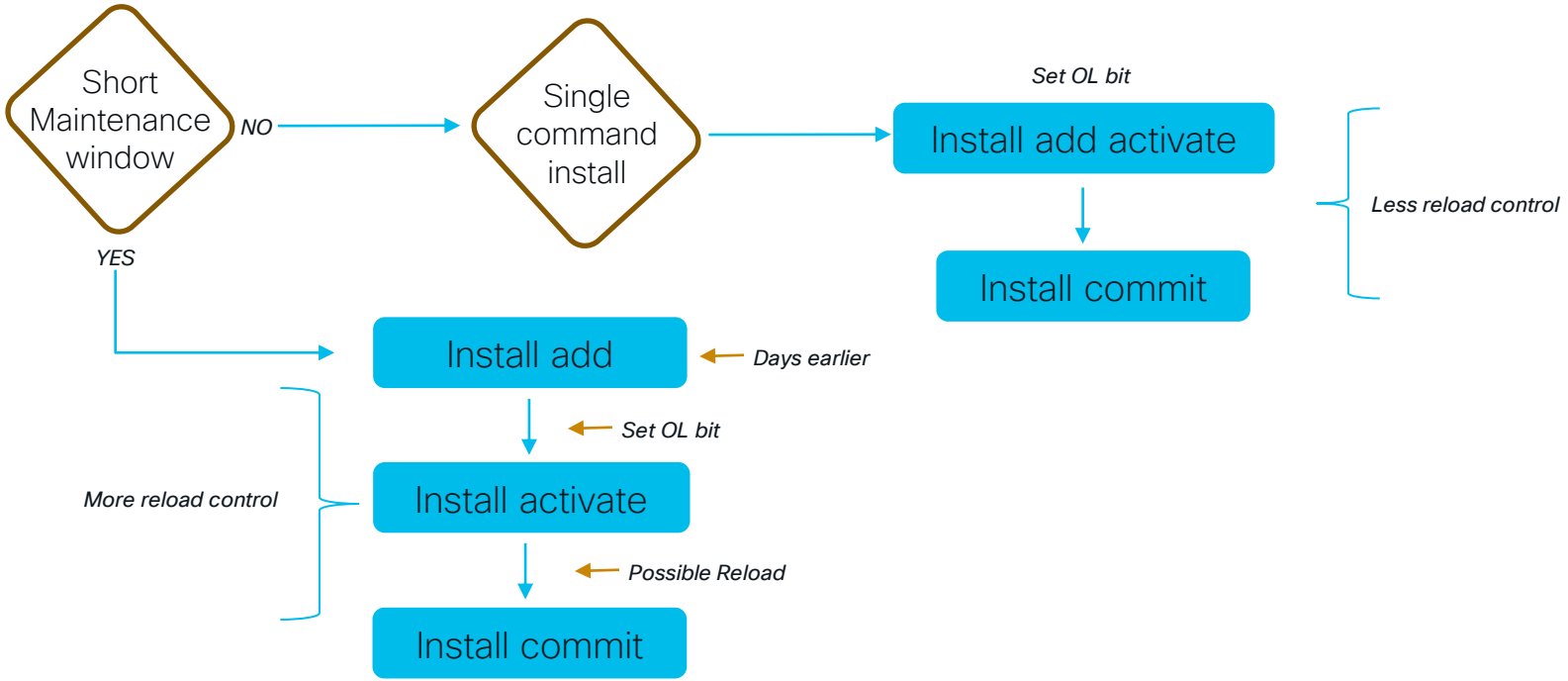
- 6.2.25
  - show install active summary (the summary keyword was added)
  - show install inactive summary
  - show install commit summary

  - show install supersede (to view superseded SMU's)
  - requires SMU CSCvj64412

- 6.3.3
  - install update source harddisk:/sw/633 asr-goldenk9-x-6.3.3-v1.iso replace noprompt (the "replace" keyword is added)
  - install upgrade (the "upgrade" keyword will be deprecated in later releases)

IOS-XR Install & Upgrade

cXR install workflow



Turboboot

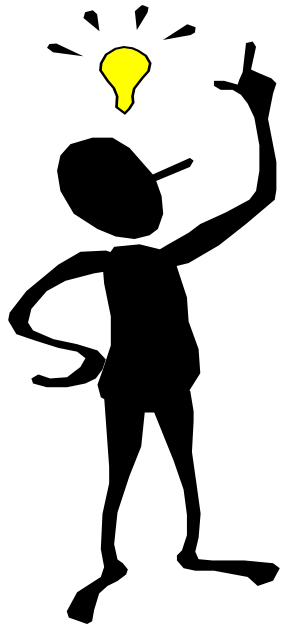
A “fresh” start!

```
Rommon1>unset BOOT
Rommon2>confreg 0x102
Rommon3>sync
Rommon4>reset


Rommon1>IP_ADDRESS=<a.b.c.d>
Rommon2>IP_SUBNET_MASK=<mask>
Rommon3>TFTP_SERVER=<a.b.c.d>
Rommon4>DEFAULT_GATEWAY=<a.b.c.d>
Rommon5>TFTP_RETRY_COUNT=4
Rommon6>TFTP_TIMEOUT=60
Rommon7>TFTP_CHECKSUM=1
Rommon8>priv
Rommon9>diswd
Rommon10>unset BOOT
Rommon11>TURBOBOOT=on,disk0,format
Rommon12>sync
Rommon13>boot tftp://a.b.c.d/path/asr9k-
mini-px.vm-4.2.3
```

- Previous disk data is wiped
- A clean “re-image” as a new router
- Potential recovery mechanism
- Not an “upgrade” method

PIE Installation Concepts

- PIE install used once system is operational w/ XR
  - .vm files can be used if booting from rommon is required
- Packages can be added or upgraded
- Install from Exec or Admin Mode
- 3 phase install
  - Add – Copy package and unpack
  - Activate – Restart processes/nodes with new code
  - Commit – Lock activated packages through reset



Install Add

Copy package to disk

```
RP/0/RSP0/CPU0:Nevada# install add tftp://7.1.1.1/asr9k-mpls-p.pie-4.3.2 sync
Install operation 3 'install add /tftp://7.1.1.1/asr9k-mpls-p.pie-4.3.2
synchronous' started by user 'ww' on SDR Owner via CLI at 22:21:54 EDT Sun Jul
26 2009.
Warning:  This add operation will add the specified package to SDR: Owner only.
Warning:  Any further operations on this package will only be allowed from SDR:
Warning:  Owner.
Warning:  If the package is meant to be added to the entire router, then please
Warning:  stop this operation and perform the operation from the admin level.
Continue with the operation? [confirm]
Info:     The following package is now available to be activated:
Info:
Info:         disk0:asr9k-mpls-4.3.2
Info:
Info:     The package can be activated on the following SDR:
Info:
Info:         Owner
Info:
Install operation 3 completed successfully at 22:22:14 EDT Sun Jul 26 2009.
```

- Copy pie from tftp, harddisk, flash, or other source
- Unpack pie into destination directory in disk
- Does not restart processes or trigger any changes to functionality
- Displays package name to activate in next step

Install Activate

Activate (install) packages

```
RP/0/RSP0/CPU0:Nevada# install activate disk0:asr9k-mpls-4.3.2 sync
Install operation 4 'install activate disk0:asr9k-mpls-4.3.2 synchronous'
started by user 'ww' on SDR Owner via CLI at 22:24:50 EDT Sun Jul 26 2009.
Info:     Install Method: Parallel Process Restart
Info:     The changes made to software configurations will not be persistent
Info:     across system reloads. Use the command '(admin) install commit' to
Info:     make changes persistent.
Info:     Please verify that the system is consistent following the software
Info:     change using the following commands:
Info:         show system verify
Info:         install verify packages
Install operation 4 completed successfully at 22:25:38 EDT Sun Jul 26 2009.
```

- Restart appropriate processes or nodes with new software
- May trigger reboot, depending on software being activated
- New functions will be available once activation completes

Install Commit

Commit the active packages

```
RP/0/RSP0/CPU0:rasr9000-2w-a#admin install commit
Sun Jan 26 16:59:22.742 EST
Install operation 36 '(admin) install commit' started by user 'cisco' via CLI
at 16:59:22 EST Sun Jan 26 2014.
Install operation 36 completed successfully at 16:59:24 EST Sun Jan 26 2014.
RP/0/RSP0/CPU0:rasr9000-2w-a#admin show install committed summary
Sun Jan 26 16:59:44.478 EST
Default Profile:
  SDRs:
    Owner
  Committed Packages:
    disk0:asr9k-px-4.2.3.CSCud37351-1.0.0
    disk0:asr9k-px-4.2.3.CSCug30234-1.0.0
    disk0:asr9k-px-4.2.3.CSCuf32158-1.0.0
    disk0:asr9k-px-4.2.3.CSCue21083-1.0.0
    disk0:asr9k-9000v-nV-px-4.2.3
    disk0:asr9k-optic-px-4.2.3
    disk0:asr9k-mini-px-4.2.3
    disk0:asr9k-doc-px-4.2.3
    disk0:asr9k-k9sec-px-4.2.3
    .
```

- The active packages become committed for future system loads

Install Deactivate

Removing a package from “running”

```
RP/0/RSP0/CPU0:Nevada# show install active
  Node 0/RSP0/CPU0 [RP] [SDR: Owner]
    Boot Device: disk0:
    Boot Image: /bootflash/disk0/asr9k-os-mbi-4.3.2/mbiasr9k-rp.vm
    Active Packages:
      disk0:asr9k-mpls-4.3.2
      disk0:asr9k-base-4.3.2.CSCsy23972-1.0.0
      disk0:comp-asr9k-mini-4.3.2

RP/0/RSP0/CPU0:Nevada# install deactivate disk0:asr9k-mpls-4.3.2 sync
Install operation 6 'install deactivate disk0:asr9k-mpls-4.3.2 synchronous'
started by user 'ww' on SDR Owner via CLI at 22:28:55 EDT Sun Jul 26 2009.
Info:     Install Method: Parallel Process Restart
Info:     The changes made to software configurations will not be persistent
Info:     across system reloads. Use the command '(admin) install commit' to
Info:     make changes persistent.
Info:     Please verify that the system is consistent following the software
Info:     change using the following commands:
Info:         show system verify
Info:         install verify packages
Install operation 6 completed successfully at 22:29:49 EDT Sun Jul 26 2009.
```

- Package’s features are no longer available
- Package is still “added” and on disk
- Package can be reactivated

Install Remove

Deleting packages from disk

```
RP/0/RSP0/CPU0:Nevada# install remove disk0:asr9k-mpls-4.3.2 sync
Install operation 9 'install remove disk0:asr9k-mpls-4.3.2 synchronous' started
by user 'ww' on SDR Owner via CLI at 22:35:08 EDT Sun Jul 26 2009.
Info:     This operation will remove the following package:
Info:         disk0:asr9k-mpls-4.3.2
Info:     After this install remove the following install rollback point will
Info:     no longer be reachable, as the required packages will not be present:
Info:         4
Proceed with removing these packages? [confirm]
Install operation 9 completed successfully at 22:35:12 EDT Sun Jul 26 2009.
```

- Package/PIE is completely removed
- **`install remove inactive`** available for cleaning up disk

Installation Log

Example

```
RP/0/RSP0/CPU0:rasr9000-2w-b#admin show install log reverse
Thu Jan  9 23:09:25.511 EST
Install operation 42 started by user 'cisco' via CLI at 21:27:00 EST Thu Jan 09 2014.
    (admin) install commit
    Install operation 42 completed successfully at 21:27:01 EST Thu Jan 09 2014.
--------------------------------------------------------------------------------
Install operation 41 started by user 'cisco' via CLI at 21:11:27 EST Thu Jan 09 2014.
    (admin) install rollback to 39
    Install operation 41 completed successfully at 21:12:22 EST Thu Jan 09 2014.
--------------------------------------------------------------------------------
Install operation 40 started at 14:11:08 UTC Thu Jan 02 2014.
    Reload of router to the committed software.
--------------------------------------------------------------------------------
Install operation 39 started by user 'cisco' via CLI at 17:18:32 EST Thu Dec 19 2013.
    (admin) install activate disk0:*4.3.1*
    Install operation 39 completed successfully at 17:27:56 EST Thu Dec 19 2013.
--------------------------------------------------------------------------------
Install operation 38 started by user 'cisco' via CLI at 17:17:24 EST Thu Dec 19 2013.
    (admin) install activate disk0:*4.3.1* test
    Install operation 38 completed successfully at 17:18:06 EST Thu Dec 19 2013.
--------------------------------------------------------------------------------
```

Installation Entries

Example

```
RP/0/RSP0/CPU0:Nevada# show install log 6 detail

Install operation 6 started by user 'ww' on SDR Owner via CLI at 22:28:55 EDT
Sun Jul 26 2009.
install deactivate disk0:asr9k-mpls-4.3.2 synchronous
Install operation 6 completed successfully at 22:29:49 EDT Sun Jul 26 2009.
Install logs:
    Install operation 6 'install deactivate disk0:asr9k-mpls-4.3.2 synchronous'
    started by user 'ww' on SDR Owner via CLI at 22:28:55 EDT Sun Jul 26 2009.
    Info:     Install Method: Parallel Process Restart
    Info:     The changes made to software configurations will not be
    Info:     persistent across system reloads. Use the command '(admin)
    Info:     install commit' to make changes persistent.
    Info:     Please verify that the system is consistent following the
    Info:     software change using the following commands:
    Info:         show system verify
    Info:         install verify packages
    Install operation 6 completed successfully at 22:29:49 EDT Sun Jul 26 2009.
Summary:
    Sub-operation 1:
    Install method: Parallel Process Restart
    Summary of changes on node 0/RSP0/CPU0:
        Deactivated:  asr9k-mpls-4.3.2
            6 asr9k-mpls processes affected (0 updated, 0 added, 6 removed, 0 impacted)
    Summary of changes on node 0/0/CPU0:
        Deactivated:  asr9k-mpls-4.3.2
            1 asr9k-mpls processes affected (0 updated, 0 added, 1 removed, 0 impacted)
```

Cards' FPD

Verify/upgrade FPD version

```
RP/0/RSP0/CPU0:rasr9000-2w-b#admin show hw-module fpd location all
================================== ========================================
                                   Existing Field Programmable Devices
                                   ========================================
                                     HW                          Current SW Upg/
Location     Card Type               Version Type Subtype Inst   Version   Dng?
============ ======================= ======= ==== ======= ==== =========== ====
0/RSP0/CPU0  A9K-RSP440-SE             1.0   lc   cbc     0      16.115    No
                                             lc   fpga1   0       0.09     No
                                             lc   fpga2   0       1.06     No
                                             lc   fpga3   0       4.09     No
                                             lc   rommon  0       0.62     No
-------------------------------------------------------------------------------
0/RSP0/CPU0  ASR-9006-FAN              1.0   lc   cbc     2       5.02     No
-------------------------------------------------------------------------------
0/0/CPU0     A9K-24x10GE-SE            1.0   lc   cbc     0      19.110    No
                                             lc   fpga2   0       1.02     No
                                             lc   fpga3   0       1.01     No
                                             lc   fpga4   0       1.05     No
                                             lc   rommon  0       1.28     No
-------------------------------------------------------------------------------
0/1/CPU0     A9K-MOD80-SE              1.0   lc   cbc     0      20.116    No
                                             lc   fpga2   0       1.01     No
                                             lc   fpga4   0       1.05     No
                                             lc   rommon  0       1.28     No
RP/0/RSP0/CPU0:rasr9000-2w-a#admin upgrade hw-module fpd all location 0/1/CPU0
```

No == good. No change needed

If LC1 needs upgrade

Cards' FPD

Verify/upgrade FPD version

```
RP/0/RSP0/CPU0:rasr9000-2w-a#admin show running-config
.
fpd auto-upgrade
.
RP/0/RSP0/CPU0:rasr9000-2w-a#admin upgrade hw-module fpd all location all
Wed Jan 15 18:57:04.683 EST

***** UPGRADE WARNING MESSAGE: *****
  *  This upgrade operation has a maximum timout of 160 minutes.  *
  *  If you are executing the cmd for one specific location and  *
  *  card in that location reloads or goes down for some reason  *
  *  you can press CTRL-C to get back the RP's prompt.           *
  *  If you are executing the cmd for _all_ locations and a node *
  *  reloads or is down please allow other nodes to finish the   *
  *  upgrade process before pressing CTRL-C.                     *

% RELOAD REMINDER:
  - The upgrade operation of the target module will not interrupt its normal
    operation. However, for the changes to take effect, the target module
    will need to be manually reloaded after the upgrade operation. This can
    be accomplished with the use of "hw-module <target> reload" command.
  - If automatic reload operation is desired after the upgrade, please use
    the "reload" option at the end of the upgrade command.
  - The output of "show hw-module fpd location" command will not display
    correct version information after the upgrade if the target module is
    not reloaded.
NOTE: Chassis CLI will not be accessible while upgrade is in progress.
Continue? [confirm]
```

Auto FPD upgrade configuration

Manual FPD upgrade

Satellite OS and Firmware

Verify/upgrade satellites

```
RP/0/RSP0/CPU0:rasr9000-2w-a#admin show install committed summary | include nV
Sat Dec 14 21:28:38.192 EST
    disk0:asr9k-9000v-nV-px-4.2.3
RP/0/RSP0/CPU0:rasr9000-2w-a#show nv satellite status satellite 100
Sat Dec 14 20:55:00.733 EST
Satellite 100
-------------
  State: Connected (Stable)
  Type: asr9000v
  Description: r9000v-1y-a
  MAC address: 8478.ac01.349c
  IPv4 address: 10.100.111.100
  Configured Serial Number: CAT1702U20H
  Received Serial Number: CAT1702U20H
  Remote version: Compatible (not latest version)
    ROMMON: 124.0 (Available: 125.0)
    FPGA: 1.13 (Latest)
    IOS: 200.5 (Available: 210.0)
  Configured satellite fabric links:
    Bundle-Ether3
    -------------
      State: Satellite Ready
      Port range: GigabitEthernet0/0/0-29
      Discovered satellite fabric links:
        TenGigE0/1/1/3: Satellite Ready; No conflict
        TenGigE0/1/0/3: Satellite Ready; No conflict
```

Host has a newer version

Satellite Upgrade

Prepare

```
RP/0/RSP0/CPU0:rasr9000-2w-a#admin show install committed summary | include nV
Sat Dec 14 21:28:38.192 EST
    disk0:asr9k-9000v-nV-px-4.2.3

RP/0/RSP0/CPU0:rasr9000-2w-a#show running-config control-plane management-plane inband
Sat Dec 14 21:29:28.655 EST
control-plane
 management-plane
  inband
   interface Bundle-Ether3
    allow TFTP
   !
  !
 !
!

RP/0/RSP0/CPU0:rasr9000-2w-a#install nv satellite 100 transfer
Sat Dec 14 21:24:28.826 EST
Install operation initiated successfully.

RP/0/RSP0/CPU0:rasr9000-2w-a#show logging | inc transfer
Sat Dec 14 21:24:31.977 EST
RP/0/RSP0/CPU0:Dec 14 21:23:19.083 : icpe_gco[1149]: %PKT_INFRA-ICPE_GCO-6-TRANSFER_DONE : Image transfer
 completed on Satellite 100
```

Satellite Upgrade

Upgrade

```
RP/0/RSP0/CPU0:rasr9000-2w-a#install nv satellite 100 activate
Sat Dec 14 21:25:26.000 EST
WARNING: This install operation will reload the requested satellite(s)
Do you wish to continue?[confirm(y/n)]
y
Install operation initiated successfully.
RP/0/RSP0/CPU0:rasr9000-2w-a#LC/0/1/CPU0:Dec 14 21:25:28.992 : ifmgr[201]: %PKT_INFRA-LINK-3-UPDOWN : Interface TenGigE0/1/1/3,
changed state to Down
LC/0/1/CPU0:Dec 14 21:25:28.995 : ifmgr[201]: %PKT_INFRA-LINK-3-UPDOWN : Interface TenGigE0/1/0/3, changed state to Down
RP/0/RSP0/CPU0:Dec 14 21:25:28.999 : ifmgr[247]: %PKT_INFRA-LINK-3-UPDOWN : Interface GigabitEthernet100/0/0/11, changed state to Down
RP/0/RSP0/CPU0:Dec 14 21:25:28.999 : ifmgr[247]: %PKT_INFRA-LINK-3-UPDOWN : Interface GigabitEthernet100/0/0/9, changed state to
Down
LC/0/1/CPU0:Dec 14 21:25:42.929 : pfm_node_lc[291]: %PLATFORM-XFP-2-LOW_RX_POWER_ALARM : Set|envmon_lc[172118]|XFP(0x102a003)|Port_0/03
LC/0/1/CPU0:Dec 14 21:25:42.929 : pfm_node_lc[291]: %PLATFORM-XFP-2-LOW_RX_POWER_ALARM : Set|envmon_lc[172118]|XFP(0x102a017)|Port_1/03
LC/0/1/CPU0:Dec 14 21:27:04.331 : ifmgr[201]: %PKT_INFRA-LINK-3-UPDOWN : Interface TenGigE0/1/0/3, changed state to Up
LC/0/1/CPU0:Dec 14 21:27:04.332 : ifmgr[201]: %PKT_INFRA-LINK-3-UPDOWN : Interface TenGigE0/1/1/3, changed state to Up
LC/0/1/CPU0:Dec 14 21:27:04.334 : ifmgr[201]: %PKT_INFRA-LINK-3-UPDOWN : Interface TenGigE0/1/0/3, changed state to Down
LC/0/1/CPU0:Dec 14 21:27:04.339 : ifmgr[201]: %PKT_INFRA-LINK-3-UPDOWN : Interface TenGigE0/1/1/3, changed state to Down
LC/0/1/CPU0:Dec 14 21:27:05.337 : ifmgr[201]: %PKT_INFRA-LINK-3-UPDOWN : Interface TenGigE0/1/0/3, changed state to Up
LC/0/1/CPU0:Dec 14 21:27:05.338 : ifmgr[201]: %PKT_INFRA-LINK-3-UPDOWN : Interface TenGigE0/1/1/3, changed state to Up
RP/0/RSP0/CPU0:Dec 14 21:27:08.659 : eem_ed_oir[191]: Messge received content : Event 0 NodeId: 0xffffffff
RP/0/RSP0/CPU0:Dec 14 21:27:08.660 : eem_ed_oir[191]: Messge received content : Event 0 NodeId: 0xffffffff
RP/0/RSP0/CPU0:Dec 14 21:27:09.210 : envmon[206]: %PLATFORM-ENVMON-2-SAT_FAN_MFGMEM : Fan MFGMEM failure - cleared on Satellite 100
RP/0/RSP0/CPU0:Dec 14 21:27:09.210 : envmon[206]: %PLATFORM-ENVMON-2-SAT_FAN_MISSING : Fan unit missing - cleared on Satellite 100
RP/0/RSP0/CPU0:Dec 14 21:27:09.210 : eem_ed_oir[191]: Messge received content : Event 0 NodeId: 0xffffffff
RP/0/RSP0/CPU0:Dec 14 21:27:09.210 : envmon[206]: %PLATFORM-ENVMON-2-SAT_BAT_FAL_A : Battery Failure A - cleared on Satellite 100
RP/0/RSP0/CPU0:Dec 14 21:27:09.210 : envmon[206]: %PLATFORM-ENVMON-2-SAT_BAT_FAL_B : Battery Failure B - cleared on Satellite 100
RP/0/RSP0/CPU0:Dec 14 21:27:09.244 : ifmgr[247]: %PKT_INFRA-LINK-3-UPDOWN : Interface GigabitEthernet100/0/0/9, changed state to Up
RP/0/RSP0/CPU0:Dec 14 21:27:09.612 : ifmgr[247]: %PKT_INFRA-LINK-3-UPDOWN : Interface GigabitEthernet100/0/0/11, changed state to Up
```

Satellite Upgrade

Verify

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show nv satellite status satellite 100
Sat Dec 14 21:27:24.660 EST
Satellite 100
-------------
  State: Connected (Stable)
  Type: asr9000v
  Description: r9000v-1y-a
  MAC address: 8478.ac01.349c
  IPv4 address: 10.100.111.100
  Configured Serial Number: CAT1702U20H
  Received Serial Number: CAT1702U20H
  Remote version: Compatible (latest version)
    ROMMON: 125.0 (Latest)
    FPGA: 1.13 (Latest)
    IOS: 210.0 (Latest)
  Configured satellite fabric links:
    Bundle-Ether3
    -------------
      State: Satellite Ready
      Port range: GigabitEthernet0/0/0-29
      Discovered satellite fabric links:
        TenGigE0/1/0/3: Satellite Ready; No conflict
        TenGigE0/1/1/3: Satellite Ready; No conflict
```

Based on host package
asr9k-9000v-nV-px-4.2.3

Package Installation Process Summary

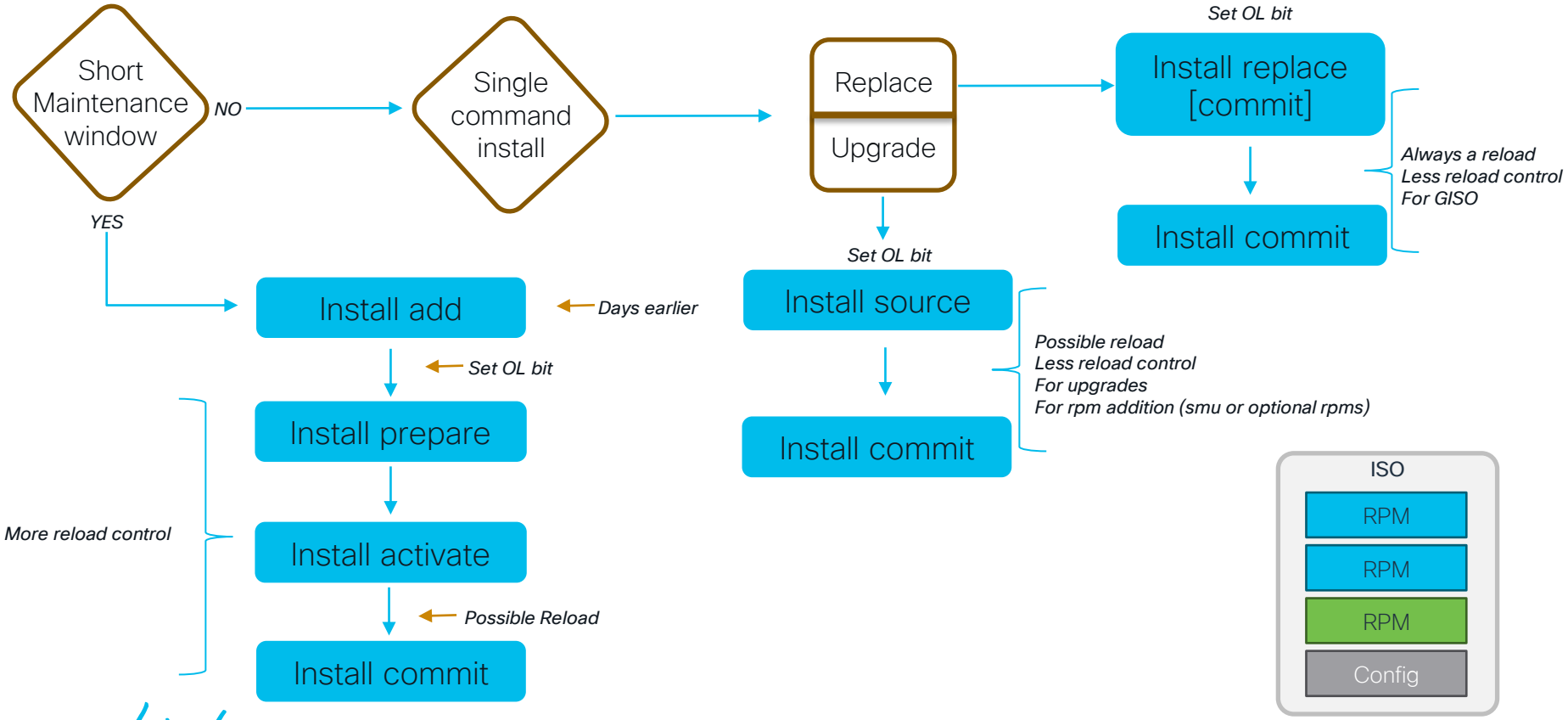
- PIE install used once system is operational
- Packages can be added or upgraded
  - Versions of all packages must be consistent
- Install from Exec or Admin Mode
- 3 phase install
  - Add – Copy package and unpack
  - Activate – Start new code
  - Commit – Lock down installed code
- FPD check – Satellite upgrade – Remove old

System Reloading Operation Difference

XR 64 Bit	cXR
• 2 VMs on each of RSP/RP/LC CPU. Reloading can happen at VM (Admin/XR), hardware module and whole chassis level • “reload” command from XR CLI only reload XR VM • “reload” command from Admin CLI reload VMs (Admin VM, XR VM or all VMs) only • “hw-module location $location reload” from Admin CLI to reload specific module or the whole chassis.	• No VMs. Reloading only happens at hardware module(each RSP/RP/LC) or whole chassis level. • “reload” command from XR CLI only reload the corresponding RSP/RP node • “reload” command from Admin CLI reload the specified hardware module. “Reload location all” reloads the whole chassis • “hw-module location $location reload” from Admin CLI to reload specific module

FPD upgrading requires Module/Chassis Reload. VM level reloading does not make new FPD effective

eXR install workflow



XR 64 Bit Operational Differences

Same XR Control Plane: XR Feature Configurations Stay Unchanged

Different Boot Process

Classic XR: ROMMON. CLI based. Support TFTP/USB boot. No ZTP capability

XR 64 Bit: Open-source iPXE. Menu/Scripts process control.IPv4/IPv6 support. TFTP/FTP/HTTP/USB boot. ZTP capable. Directly boot from ISO instead of Turboboot.

Different CLI to Reload Chassis

Classic XR: “Reload location all” from ADMIN exec prompt to reload whole chassis

XR 64 Bit: “hw-module location all reload” from ADMIN exec prompt to reload the whole chassis. “reload location all” reload the VMs only.

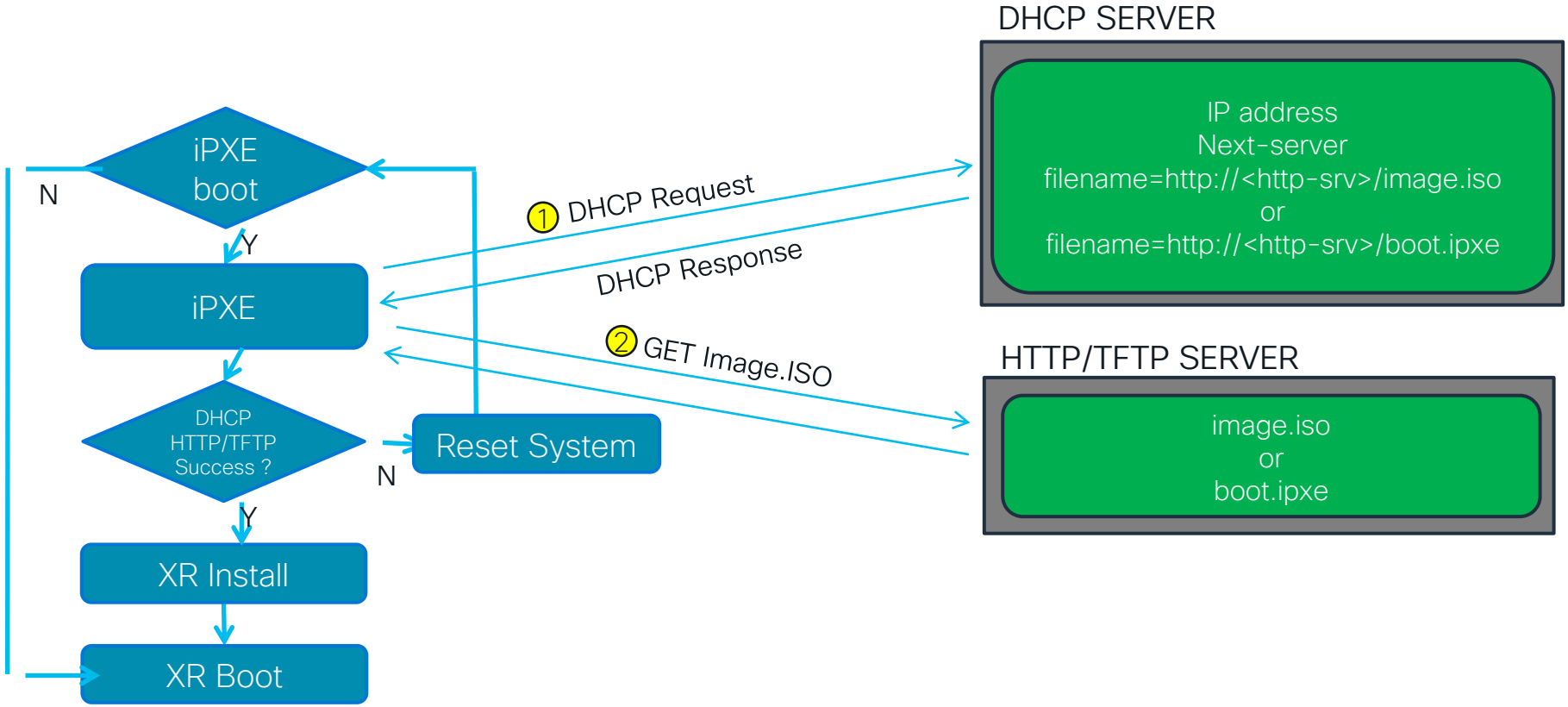
XR 64 Bit image upgrade: Admin VM → XR VM

XR 64 Bit: Added Install commands (update/upgrade) for flexible package management

Admin CLI changes: Configuration, Exec and Show commands

XR Exec and Show command CLI change. No major XR configuration CLI change

IOS-XR Boot Process with iPXE



ASR 9000 IOS XR-64 Bit Boot Menu

- admin hw-module location all reload
  - Ctrl-C to drop into Boot Menu
- admin hw-module location all bootmedia network reload

```
Booting IOS-XR 64 bit Boot previously installed image - Press Ctrl-c to stop

Please select the operating system and the boot device:
     1) IOS-XR (32 bit Classic XR)
     2) IOS-XR 64 bit Boot previously installed image
     3) IOS-XR 64 bit Mgmt Network boot using DHCP server
     4) IOS-XR 64 bit Mgmt Network boot using local settings (iPXE)
     (Press 'p' for more option)
Selection [1/2/3/4]: p
Please select the operating system and the boot device:
     1) IOS-XR (32 bit Classic XR)
     2) IOS-XR 64 bit Boot previously installed image
     3) IOS-XR 64 bit Mgmt Network boot using DHCP server
     4) IOS-XR 64 bit Mgmt Network boot using local settings (iPXE)
     5) IOS-XR 64 bit Internal network boot from RSP/RP
     6) IOS-XR 64 bit Local boot using embedded USB media
     7) IOS-XR 64 bit Local boot using front panel USB media
Selection [1/2/3/4/5/6/7]:
```

ASR 9000 XR 64 Bit iPXE Prompt

- Select option 3 from boot menu
- Right after iPXE 1.0.0+, Ctrl-B to drop into iPXE prompt

```
Please select the operating system and the boot device:
     1) IOS-XR (32 bit Classic XR)
     2) IOS-XR 64 bit Boot previously installed image
     3) IOS-XR 64 bit Mgmt Network boot using DHCP server
     4) IOS-XR 64 bit Mgmt Network boot using local settings (iPXE)
     (Press 'p' for more option)
Selection [1/2/3/4]: 3
Selected IOS-XR 64 bit Mgmt Network boot using DHCP server, Continue ? Y/N: y

Serial ATA Port 4 : SMART iSATA SHSLM32GEBCITHD02
Serial ATA Port 5 : SMART iSATA SHSLM32GEBCITHD02
Selected external PXE

CISCO iPXE initialising devices...ok

iPXE 1.0.0+ (b6461) -- Open Source Network Boot Firmware -- http://ipxe.org
Features: DNS HTTP TFTP EFI ISO9660 Menu
IPXE embed script.......
IPXE preference ...0
Press Ctrl-B for the iPXE command line...
iPXE>
```

iPXE CLI Operation Example

- Booting Image Manually

Only Supported Method for Password Recovery

```
iPXE> set net0/ip:ipv4 172.30.0.101
iPXE> set net0/netmask 255.255.255.0
iPXE> set net0/gateway 172.30.0.1
iPXE> ifopen net0
iPXE> route
net0: 172.30.0.101/255.255.255.0 gw 172.30.0.1
net0: fe80::d66d:50ff:fe18:1a50/64
net0: fd:30::d66d:50ff:fe18:1a50/64 gw fe80::fa72:eaff:fe8b:ce80
net1: fe80::d66d:50ff:fe18:1a51/64 (inaccessible)
net2: fe80::d66d:50ff:fe18:1a52/64 (inaccessible)
net3: fe80::d66d:50ff:fe18:1a53/64 (inaccessible)
iPXE> ping -c 2 172.30.0.22
64 bytes from 172.30.0.22: seq=1
64 bytes from 172.30.0.22: seq=2
iPXE> boot http://172.30.0.22/asr9k/6.4.2/asr9k-mini-x64-6.4.2.iso
http://172.30.0.22/asr9k/6.4.2/asr9k-mini-x64-6.4.2.iso... ok
Memory required for image[asr9k-mini-x64-6.4.2.iso]: 1353314304, available: 1841299456
<SNIP>
```

IOS XR 32-Bit to IOS XR 64-Bit Migration Steps

Step 1.	Upgrade current cXR to image 613/622/+
Step 2.	Ensure all supported RSP/RP/FC/LC loaded in the setup
Step 3.	Upgrade FPDs and reload is optional Setup a user with “root-lr” privileges ONLY in the XR plane
Step 4.	Clean up harddisk:/, harddiskb:/, backup cfg to harddiskb:/ via **/pkg/bin/resize_eusb in the XR plane**
Step 5.	Copy crypto file, admin, XR config file to harddiskb:/
Step 6.	Copy 613/622/+ migration tar file to harddisk:/ (Available in CCO or GISO) * GISO = Mini.iso + RPMs + SMUs + Config

IOS XR 32-Bit to IOS XR 64-Bit Migration Steps (Cont.)

Step 7.	Execute **migrate_to_eXR** script with reload enabled(-r)
Step 8.	System automatically will boot up eXR
Step 9.	If Auto FPD enabled, FPDs will get upgraded
Step 10.	Force upgrade FPD to upgrade golden FPD
Step 11.	Reboot system for FPDs to take effect.

IOS XR 64-Bit Migration Script

- Migrate to Cisco IOS XR 64-Bit Software by using the script migrate_to_eXR, available in /pkg/bin/.

- The migration script performs these tasks:
  - Backs up XR plane and admin plane config to hardisk:/
  - Copies GRUB files to /harddisk:/efi/boot/
  - Sets the boot mode on active RSP/RP to boot from harddiskb:/
  - Sets the boot mode on standby RSP/RP to boot from active RSP/RP

Migrate using eUSB Boot

```
[RP/0/RSP0/CPU0:ios# copy <img-location>/asr9k-mini-x64-migrate_to_eXR.tar6.1.3 harddisk:/asr9k-mini-x64-migrate_to_eXR.tar-6.1.3

[RP/0/RSP0/CPU0:ios#run /pkg/bin/migrate_to_eXR -m eusb –r
```

Configuration: Numbering Scheme

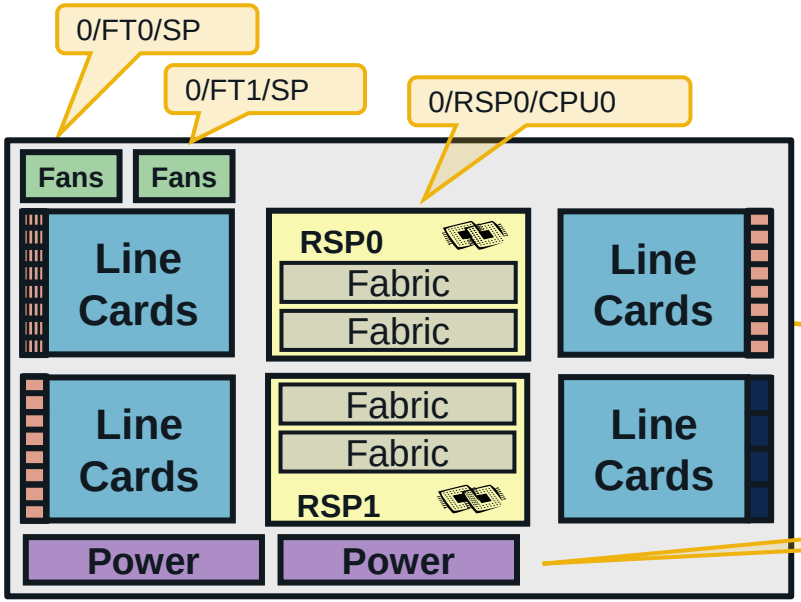
Inventory and Addressing

Exec and admin modes

- XR interfaces use 4-5 layer naming
  - Rack (0 for single chassis)
  - /Slot (0+ for LC, RSP0, RSP1)
  - /Bay (0, 0+ on MOD and SIP)
  - /Port (0+)
  - .Subinterface/EFP (optional)

```
RP/0/RSP1/CPU0:viking-1#show platform
Node              Type                  State         Config State
--------------------------------------------------------------------0/RSP0/CPU0
A9K-RSP-4G(Standby)  IOS XR RUN   PWR,NSHUT,MON
0/RSP1/CPU0       A9K-RSP-4G(Active)    IOS XR RUN    PWR,NSHUT,MON
0/0/CPU0          A9K-40GE-E            IOS XR RUN    PWR,NSHUT,MON
0/1/CPU0          A9K-8T/4-B            IOS XR RUN    PWR,NSHUT,MON
```

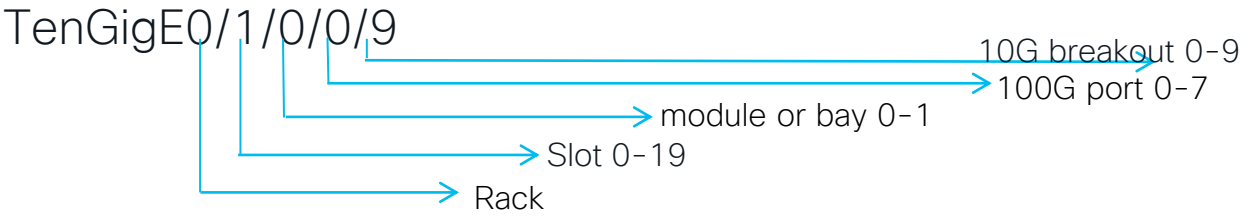
```
RP/0/RSP1/CPU0:viking-1#admin show platform
Wed Jul 22 09:23:32.482 EST
Node              Type                  State         Config State
--------------------------------------------------------------------
0/RSP0/CPU0       A9K-RSP-4G(Standby)   IOS XR RUN    PWR,NSHUT,MON
0/RSP1/CPU0       A9K-RSP-4G(Active)    IOS XR RUN    PWR,NSHUT,MON
0/FT0/SP          FAN TRAY              READY
0/FT1/SP          FAN TRAY              READY
0/0/CPU0          A9K-40GE-E            IOS XR RUN    PWR,NSHUT,MON
0/1/CPU0          A9K-8T/4-B            IOS XR RUN    PWR,NSHUT,MON
0/PM0/SP          A9K-3KW-AC            READY         PWR,NSHUT,MON
0/PM1/SP          A9K-3KW-AC            READY         PWR,NSHUT,MON
```



interface ten[Rack/Slot/Bay/Port.Sub]

0/PM1/SP

Interface/Component Numbering Scheme




```
RP/0/RP0/CPU0:asr9k#show platform
Node          Type                 State          Config State
--------------------------------------------------------------------------------
<snip>
0/1/CPU0      A9K-24x10GE-SE       IOS XR RUN     PWR,NSHUT,MON
```

TenGigE0/1/0/0
.
up to
TenGigE0/1/0/23

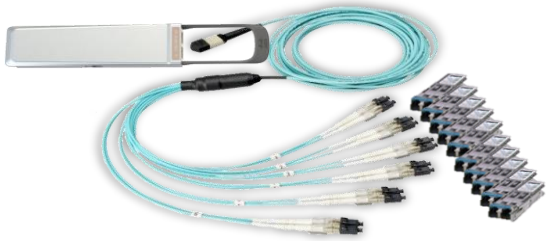
port numbering always begins at 0 for any LineCard, MPA, or SPA

Rack/Slot/Module/Port

```
RP/0/RSP0/CPU0:asr9001(admin)#show platform
Node                   Type                              State                Config State
-----------------------------------------------------------------------------------------------------
0/RSP0/CPU0      ASR9001-RP(Active)   IOS XR RUN    PWR,NSHUT,MON
0/FT0/SP             FAN TRAY                    READY
0/0/CPU0             ASR9001-LC                IOS XR RUN    PWR,NSHUT,MON
0/0/0                    A9K-MPA-4X10GE       OK                     PWR,NSHUT,MON
0/0/1                    A9K-MPA-4X10GE       OK                     PWR,NSHUT,MON
0/PM0/SP            A9K-750W-AC             READY                PWR,NSHUT,MON

RP/0/RSP0/CPU0:asr9001-nV-Edge(admin)#show platform
Node                   Type                              State                 Config State
-----------------------------------------------------------------------------------------------------
0/RSP0/CPU0     ASR9001-RP(Active)   IOS XR RUN       PWR,NSHUT,MON
0/FT0/SP             FAN TRAY                   READY
0/0/CPU0            ASR9001-LC                IOS XR RUN       PWR,NSHUT,MON
0/0/0                   A9K-MPA-20X1GE        OK                       PWR,NSHUT,MON
0/0/1                   A9K-MPA-4X10GE        OK                       PWR,NSHUT,MON
0/PM0/SP          A9K-750W-DC               READY                PWR,NSHUT,MON
1/RSP0/CPU0   ASR9001-RP(Active)     IOS XR RUN      PWR,NSHUT,MON
1/FT0/SP            FAN TRAY                    READY
1/0/CPU0          ASR9001-LC                  IOS XR RUN       PWR,NSHUT,MON
1/0/0                 A9K-MPA-20X1GE          OK                       PWR,NSHUT,MON
1/0/1                 A9K-MPA-2X10GE          OK                       PWR,NSHUT,MON
1/PM0/SP         A9K-750W-DC                 READY                PWR,NSHUT,MON
```

A cluster node

ASR9K CPAK - Mode-Change CLI



- Configuration happens at exec-config level
- Example Config:
- *hw-module location 0/0/cpu0* ***port 0*** *breakout* ***2xFortyGigE***
- *hw-module location 0/0/cpu0* ***port 1*** *breakout* ***10xTenGigE***
- *Decide which Port ? and at what Speed ?*
- *Each CPAK port can have different speed (10,40 or 100). 100G is default.*

- Reload of Line card is not necessary after rate change on ASR9K.

ASR9K CPAK – 5 Tuple Port Numbering

- 10G interface: *Interface Tengig0/5/0/1/0-9 (qty=10 TenGig ports)*
- *Example: Interface Tengig 0 / 5 / 0 / 1 / 0*

- *Rack, Slot, Instance (Bay/EP), Port (physical plughole), Breakout port*
- *40G Interface[1]: Interface FortyGige0/0/0/0/0-1 (qty=2 FortyGig ports)*
- There is no Gige interface in 530
- 5 Tuple only applies to 10G and 40G[1] speeds
- 100G speeds remain as 4 Tuple. (HunGig0/0/0/0)

** 40G is not supported in 5.3*

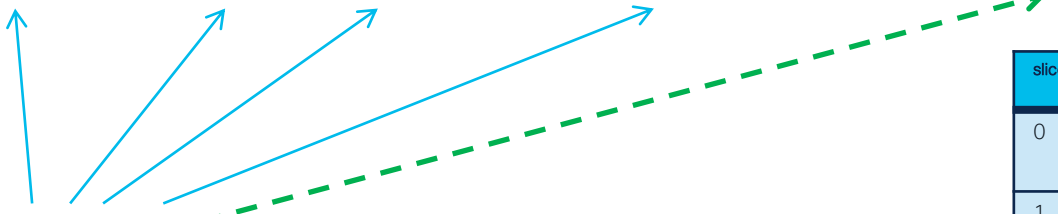
ASR9K CPAK - 5 Tuple Port Numbering - Example 1

Node, port numbering - with slice A9K-8x100GE-SE

```
RP/0/RP0/CPU0:ASR9K#show run | i module
Building configuration...
hw-module location 0/1/CPU0 port 7 breakout 10xTenGigE
```

- 10G interfaces uses 5th octet for breakout-10GE-port numbering

TenGigE<rack_num>/<slot_id>/<module/Bay>/<physical-port_num>/<breakout-port-num>



TenGigE0/1/0/7/0
TenGigE0/1/0/7/1

Rack / Slot / Module(Bay) / Port / Logical slice port

0	1	0	7	0
0	1	0	7	1

slice	100GE port	10GE port
0	0	0-9
	1	0-9
1	2	0-9
	3	0-9
2	4	0-9
	5	0-9
3	6	0-9
	7	0-9

ASR9K CPAK - 5 Tuple Port Numbering - Example 2

Interface Order

hw-module location 0/1/CPU0 port 7 breakout 10xTenGigE

- 10G interfaces uses 5th octet for breakout-10GE-port numbering

TenGigE<rack_num>/<slot_id>/<module/Bay>/<physical-port_num>/<breakout-port-num>

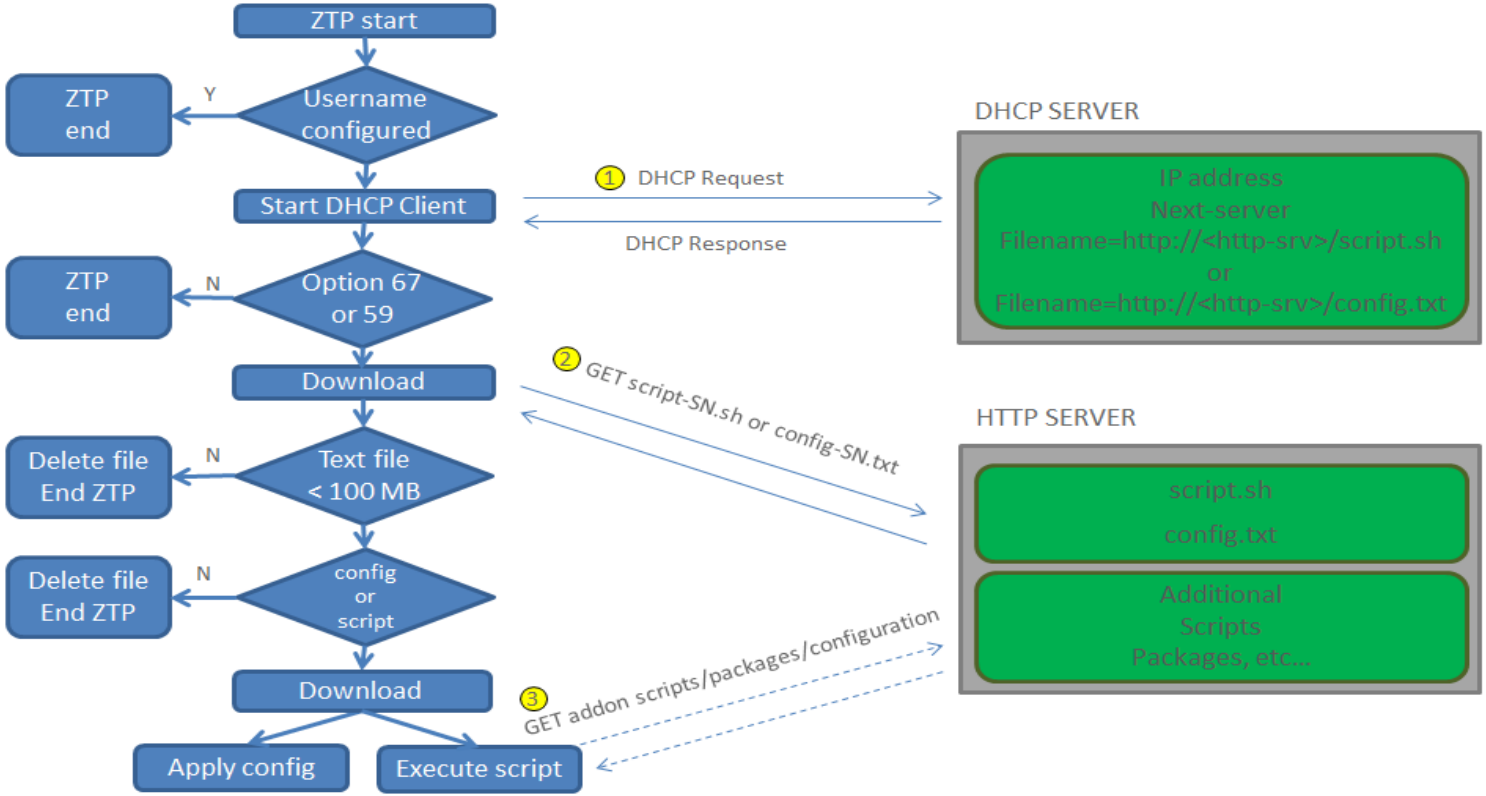
A9K-8x100GE-SE

```
HundredGigE0/1/0/0              10.1.50.1       Down                Down
HundredGigE0/1/0/1              unassigned      Down                Down
HundredGigE0/1/0/2              unassigned      Shutdown            Down
HundredGigE0/1/0/3              unassigned      Shutdown            Down
HundredGigE0/1/0/4              unassigned      Shutdown            Down
HundredGigE0/1/0/5              unassigned      Up                  Up
HundredGigE0/1/0/6              unassigned      Shutdown            Down
TenGigE0/1/0/7/0                10.1.10.1       Up                  Up
TenGigE0/1/0/7/1                10.1.11.1       Up                  Up
<snip>
TenGigE0/1/0/7/8                10.1.12.1       Up                  Up
TenGigE0/1/0/7/9                10.1.14.1       Up                  Up
```

1) Baseline this is an 8x100GE LC ports 0-7
2) We have broken out 100GE port7 into a 10x10GE breakout, so we no longer have a 100GE port 7
3) We now have 10x10GE ports from TenGigE0/1/0/7/0 - TenGigE0/1/0/7/9
4) We still have 7x100GE ports left, HundredGigE0/1/0/0 - HundredGigE0/1/0/6

Zero Touch Provisioning

ZTP Tools & Process: DHCP & HTTP



ZTP: Outcomes and Process

- Upgrade: Download and install packages
- Script: Download and run a shell script
- Configure: Download and commit a config file
- You need to:
  - Test
  - Verify: Error checking – build in process
  - Validate: Validate the node functionality [control and forwarding] – automate

HTTP Server: XR Config file

```
cisco@mamikhai-ubuntu:~$ more /var/www/html/ztp/pe125-config-initial.txt
!! IOS XR Configuration version = 6.2.1
!
hostname PE125
group ISIS-P2P
 router isis 'ISIS'
  interface 'GigabitEthernet0/0/0/0\..+'
   point-to-point
  !
 !
end-group
clock timezone EST America/New_York
domain name cisco.com
.
interface Loopback0
 ipv4 address 10.101.125.1 255.255.255.255
 ipv6 address 2001:db8:125::1/128
!
interface GigabitEthernet0/0/0/0
 mtu 4470
!
interface GigabitEthernet0/0/0/0.300 l2transport
 encapsulation dot1q 300
!
interface GigabitEthernet0/0/0/0.1225
.
```

File start:
!! IOS XR == XR config

HTTP Server: XR Script

```
cisco@mamikhai-ubuntu:~$ more /var/www/html/ztp/pe125-script.sh
#!/bin/bash

source ztp_helper.sh
config_file='/disk0:/ztp/tmp/config.txt'
config_log='/disk0:/ztp/customer/config-log.txt'

/bin/touch $config_log

if [ -f $config_file ]; then
  /bin/rm -f $config_file
else
  /bin/touch $config_file
fi

echo 'username cisco' >> $config_file
echo ' group root-lr' >> $config_file
echo ' group cisco-support' >> $config_file
echo ' secret cisco' >> $config_file
echo 'interface MgmtEth0/RP0/CPU0/0' >> $config_file
echo ' ipv4 address 192.168.30.125 255.255.255.0' >> $config_file
echo ' no shutdown' >> $config_file
echo 'netconf-yang agent' >> $config_file
echo ' ssh' >> $config_file
echo 'ssh server v2' >> $config_file
echo 'ssh server netconf vrf default' >> $config_file
```

File start:
#!/bin/bash == script

HTTP Server: XR Script – continued

```
xrapply_with_reason 'Initial ZTP config' $config_file


if [[ -z $(xrcmd "show crypto key mypubkey rsa") ]]; then
    echo ”2048" | xrcmd "crypto key generate rsa"
else
    echo -ne "yes\n 2048\n" | xrcmd "crypto key generate rsa"
fi


xrcmd 'show running-config' >> $config_log
xrcmd 'show configuration failed' >> $config_log
xrcmd 'show crypto key mypubkey rsa' >> $config_log
```

+ can prep ssh ☺ [for NETCONF]

```
cisco@mamikhai-ubuntu:~$ ll /var/www/html/ztp/
total 40
drwxr-xr-x 2 root root 4096 Oct 22 08:52 ./
drwxr-xr-x 3 root root 4096 Aug  7 11:26 ../
-rw-r--r-- 1 root root  513 Oct 18 19:46 pe125-config-initial.txt
-rw-r--r-- 1 root root 8503 Oct 18 08:14 pe125-config.txt
.
-rw-r--r-- 1 root root 1090 Oct 22 08:52 pe125-script.sh
```

Served by HTTP server

DHCP Server: Initial Parameters & Pointer

```
cisco@mamikhai-ubuntu:~$ more /etc/dhcp/dhcpd.conf
.
# option definitions common to all supported networks...
option domain-name "cisco.com";
option domain-name-servers 171.70.168.183, 64.102.6.247;

default-lease-time 600;
max-lease-time 7200;

subnet 192.168.30.0 netmask 255.255.255.0 {
}

.


host PE125 {
  hardware ethernet 00:50:56:85:da:18;
  fixed-address 192.168.30.125;
  option routers 192.168.30.1;
  filename "http://192.168.30.101/ztp/pe125-script.sh";
}
```

ZTP Node Ready for NETCONF

```
<?xml version="1.0" encoding="UTF-8"?><data
xmlns="urn:ietf:params:xml:ns:netconf:base:1.0"
xmlns:nc="urn:ietf:params:xml:ns:netconf:base:1.0">
.
   <interface-configuration>
    <active>act</active>
    <interface-name>MgmtEth0/RP0/CPU0/0</interface-
name>
    <ipv4-network xmlns="http://cisco.com/ns/yang/Cisco-
IOS-XR-ipv4-io-cfg">
     <addresses>
      <primary>
       <address>192.168.30.125</address>
       <netmask>255.255.255.0</netmask>
      </primary>
     </addresses>
    </ipv4-network>
   </interface-configuration>
.
  <ssh xmlns="http://cisco.com/ns/yang/Cisco-IOS-XR-
crypto-ssh-cfg">
   <server>
    <v2/>
    <netconf-vrf-table>
     <vrf>
      <vrf-name>default</vrf-name>
      <enable/>
     </vrf>
    </netconf-vrf-table>
   </server>
  </ssh>
.
```

```
<netconf-yang xmlns="http://cisco.com/ns/yang/Cisco-
IOS-XR-man-netconf-cfg">
   <agent>
    <ssh>
     <enable/>
    </ssh>
   </agent>
  </netconf-yang>
.
  <aaa xmlns="http://tail-f.com/ns/aaa/1.1">
   <authentication>
    <users>
     <user>
      <name>cisco</name>
      <uid>9000</uid>
      <gid>100</gid>
<password>$1$glU0$OEyQD/4ePFuNrZ2d0xtQo0</password>
<ssh_keydir>/var/confd/homes/cisco/.ssh</ssh_keydir>
      <homedir>/var/confd/homes/cisco</homedir>
     </user>
    </users>
    <groups>
     <group>
      <name>aaa-r</name>
      <gid>100</gid>
      <users>%%__system_user__%%</users>
     </group>
     <group>
      <name>admin-r</name>
      <gid>100</gid>
      <users>%%__system_user__%%</users>
.
```

Configuration: NetConf-YANG

Configure NetConf over SSH

Pre-requisites: Manageability Package and SSH Service

```
RP/0/RSP1/CPU0:rasr9000-2w-a#admin show
install active summary
Fri Dec  4 00:44:26.004 EST
Default Profile:
  SDRs:
    Owner
  Active Packages:
    disk0:asr9k-services-infra-5.3.1
    disk0:asr9k-bng-px-5.3.1
    disk0:asr9k-video-px-5.3.1
    disk0:asr9k-mgbl-px-5.3.1
    disk0:asr9k-mcast-px-5.3.1
    disk0:asr9k-k9sec-px-5.3.1
    disk0:asr9k-services-px-5.3.1
    disk0:asr9k-optic-px-5.3.1
    disk0:asr9k-mpls-px-5.3.1
    disk0:asr9k-doc-px-5.3.1
    disk0:asr9k-fpd-px-5.3.1
    disk0:asr9k-9000v-nV-px-5.3.1
    disk0:asr9k-li-px-5.3.1
    disk0:asr9k-mini-px-5.3.1
```

```
RP/0/RSP1/CPU0:rasr9000-2w-a#crypto key
generate rsa
Fri Dec  4 00:44:47.039 EST
The name for the keys will be: the_default
  Choose the size of the key modulus in the
range of 512 to 4096 for your General
Purpose Keypair. Choosing a key modulus
greater than 512 may take a few minutes.

How many bits in the modulus [1024]:
Generating RSA keys ...
Done w/ crypto generate keypair
[OK]

RP/0/RSP1/CPU0:rasr9000-2w-a#configure
Fri Dec  4 00:45:02.974 EST
RP/0/RSP1/CPU0:rasr9000-2w-a(config)#ssh
server v2
RP/0/RSP1/CPU0:rasr9000-2w-a(config)#commit
Fri Dec  4 00:45:10.994 EST
RP/0/RSP1/CPU0:rasr9000-2w-a(config)#
RP/0/RSP1/CPU0:rasr9000-2w-a#
```

Configure NetConf over SSH

NETCONF – YANG – SSH Port

```
RP/0/RSP1/CPU0:rasr9000-2w-a#show running-
config | begin netconf
Fri Dec  4 02:45:23.593 EST
Building configuration...
netconf agent tty
!
netconf-yang agent
 ssh
!
ssh server v2
ssh server netconf port 830
.
```

```
RP/0/0/CPU0:PE178#ssh 10.101.137.1 username cisco
source-interface loopback 0 command netconf format
.
Password:

.
<?xml version="1.0" encoding="UTF-8"?>
<hello
xmlns="urn:ietf:params:xml:ns:netconf:base:1.0”>
    <capabilities>
        <capability>
            urn:ietf:params:netconf:base:1.0
        </capability>
        <capability>

urn:ietf:params:netconf:capability:candidate:1.0
        </capability>
        <capability>

urn:ietf:params:netconf:capability:notification:1.0
        </capability>
    </capabilities>
    <session-id>
        285212672
    </session-id>
</hello>
]]>]]>
```

NETCONF Example: <get-config> Operation

```
<?xml version="1.0" encoding="UTF-8"?>
<rpc message-id="101"
xmlns="urn:ietf:params:xml:ns:netconf:base:1
.0">
   <get-config>
     <source>
       <running/>
     </source>
     <filter>
       <Configuration>
          <InterfaceConfigurationTable>
              <InterfaceConfiguration>
                  <Naming>
                      <Active>act</Active>
                      <InterfaceName
Match="Loopback.*"/>
                  </Naming>
              </InterfaceConfiguration>
          </InterfaceConfigurationTable>
       </Configuration>
     </filter>
   </get-config>
</rpc>
]]>]]>
```

```
<?xml version="1.0" encoding="UTF-8"?>
<rpc-reply message-id="101"
xmlns="urn:ietf:params:xml:ns:netconf:base:1.0">
    <data>
        <Configuration>
            <InterfaceConfigurationTable MajorVersion="5"
MinorVersion="4">
                <InterfaceConfiguration>
                    <Naming>
                        <Active>
                            act
                        </Active>
                        <InterfaceName>
                            Loopback0
                        </InterfaceName>
                    </Naming>
                    <InterfaceVirtual>
                        true
                    </InterfaceVirtual>
                    <IPV4Network MajorVersion="6"
MinorVersion="4">
                        <Addresses>
                            <Primary>
                                <Address>
                                    10.101.137.1
                                </Address>
                                <Netmask>
                                    255.255.255.255
                                </Netmask>
                            </Primary>
                        </Addresses>
                    </IPV4Network>
                    <IPV6Network MajorVersion="2"
MinorVersion="3">
                        .
```

NETCONF-YANG

Sessions

```
RP/0/0/CPU0:PE178#show netconf-yang clients
Fri Dec  4 12:51:32.840 EST
Netconf clients
client session ID|     NC version|    client connect time|        last OP time|        last OP type|    <lock>|
           20890|           1.1|         0d  0h  6m 12s|            12:45:19|                 get|       No|

RP/0/0/CPU0:PE178#show netconf-yang statistics
Fri Dec  4 12:51:37.999 EST
Summary statistics
         # requests|            total time|   min time per request|   max time per request|   avg time per request|
other             0|       0h  0m  0s   0ms|       0h  0m  0s   0ms|       0h  0m  0s   0ms|       0h  0m  0s   0ms|
close-session     0|       0h  0m  0s   0ms|       0h  0m  0s   0ms|       0h  0m  0s   0ms|       0h  0m  0s   0ms|
kill-session      0|       0h  0m  0s   0ms|       0h  0m  0s   0ms|       0h  0m  0s   0ms|       0h  0m  0s   0ms|
get-schema       57|       0h  0m  2s 489ms|       0h  0m  0s   0ms|       0h  0m  0s 319ms|       0h  0m  0s   8ms|
get               6|       0h  0m  0s 669ms|       0h  0m  0s   9ms|       0h  0m  0s 299ms|       0h  0m  0s 111ms|
get-config        0|       0h  0m  0s   0ms|       0h  0m  0s   0ms|       0h  0m  0s   0ms|       0h  0m  0s   0ms|
edit-config       0|       0h  0m  0s   0ms|       0h  0m  0s   0ms|       0h  0m  0s   0ms|       0h  0m  0s   0ms|
commit            0|       0h  0m  0s   0ms|       0h  0m  0s   0ms|       0h  0m  0s   0ms|       0h  0m  0s   0ms|
cancel-commit     0|       0h  0m  0s   0ms|       0h  0m  0s   0ms|       0h  0m  0s   0ms|       0h  0m  0s   0ms|
.

RP/0/0/CPU0:PE178#show tcp brief | include 830
Fri Dec  4 12:51:50.959 EST
0x1015b21c 0x60000000      0      0  :::830                 :::0                   LISTEN
0x101577b8 0x00000000      0      0  :::830                 :::0                   LISTEN
0x1015f728 0x60000000      0      0  10.101.178.1:830       192.168.30.102:35406   ESTAB
0x1015b064 0x60000000      0      0  0.0.0.0:830            0.0.0.0:0              LISTEN
0x10153fa0 0x00000000      0      0  0.0.0.0:830            0.0.0.0:0              LISTEN
```

XR XML Schema: Config & Operational

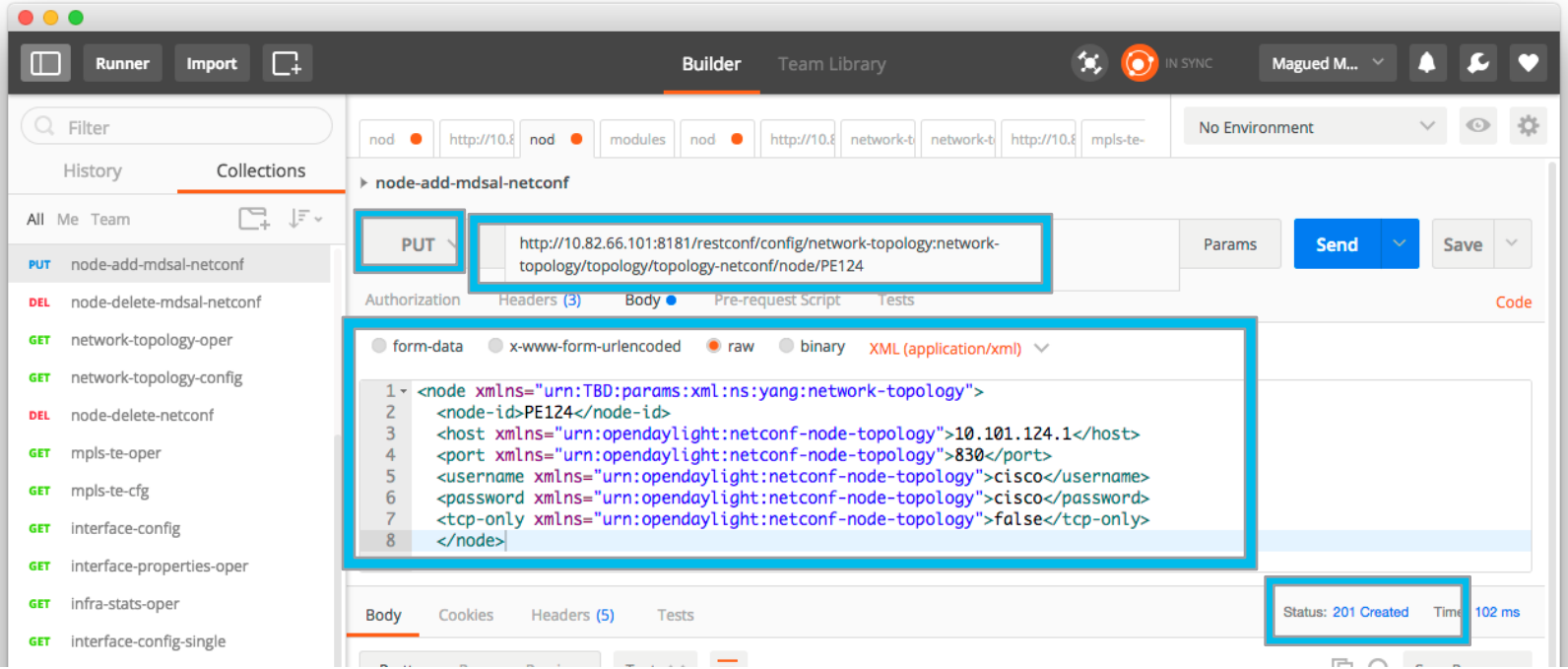
```
RP/0/0/CPU0:PE178#show xml schema
Thu Sep  3 11:06:50.353 EDT
Username:cisco
Password:
Enter 'help' or '?' for help
xml-schema[config]:> ls
[container]                     AAA
[container]                     AIB
[container]                     ARP
[container]                     BMP
[container]                     BGP
[container]                     BulkStatistics
[container]                     LACP
[container]                     CallHome
[container]                     CDP
[container]                     CEM
[container]                     NETCONF_YANG
[container]                     Clock
[container]                     Crypto
[container]
AddressPoolService
[container]                     DHCPv6
[container]                     HardwareModule
[container]                     Exception
[container]                     EIGRP
[container]                     PolicyManager
[container]                     ERP
```

```
xml-schema[config]:> oper
xml-schema[oper]:> ls
[container]                     AAA
[container]                     AIB
[container]                     ARP_GMP
[container]                     ARP
[container]                     BGP
[container]
BundlesAdjacency
[container]                     Bundles
[container]
BundleInformation
[container]                     LACPBundles
[container]
LACPBundleMembers
[container]                     LACPData
[container]                     CDP
[container]                     CEM
[container]                     CHDLC_MA
[container]                     Cinetd
[container]                     CLNS
[container]
CryptoCertificateAuthority
[container]                     CryptoEngine
[container]                     IKE
[container]
IPSecStaticSessionInfo
[container]
IPSecDynamicSessionInfo
.
```

```
.
[container]                     ErrorDisable
[container]                     ES_ACL
[container]                     EventManager
[container]                     LR
[container]                     HSRP
[container]                     HTTPC
[container]                     HTTP
[container]
nVSatelliteTable
[container]
GenericInterfaceListTable
[container]
GlobalInterfaceConfiguration
[container]
InterfaceConfigurationTable
[container]                     NSR
[container]                     Onep
[container]                     RCC
[container]                     SLA
[container]                     Statistics
[container]                     Syslog
[container]                     SyslogService
[container]                     LISP
[container]                     IPAddressPool
[container]                     BFD
[container]                     IPDomain
.
```

Open DayLight: NETCONF Mounting Nodes

Setting up network nodes to be controlled by the controller



Configuration:
Command Line [CLI]

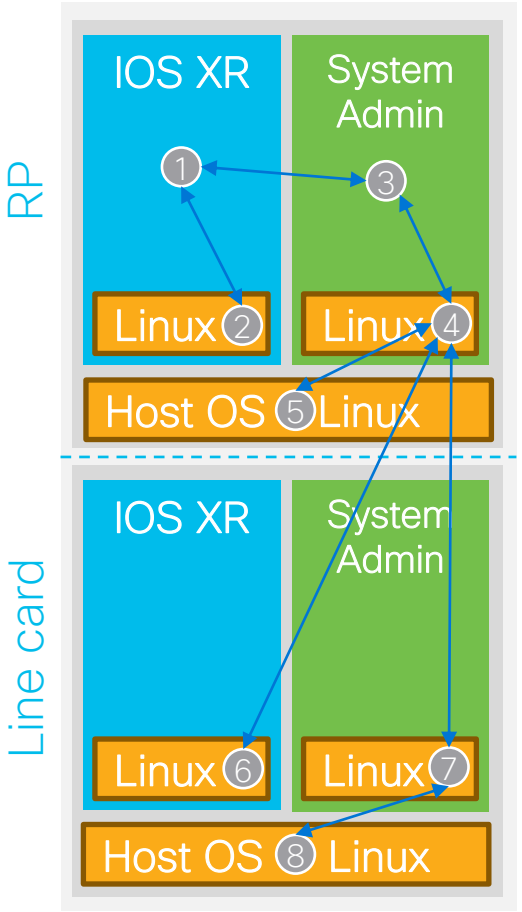
IOS XR 64 Bit CLI Access

RP XR VM, RP Admin VM, RP Host

```
RP/0/RSP0/CPU0:CORE-TOP#run                                          1

[xr-vm_node0_RSP0_CPU0:~]$
[xr-vm_node0_RSP0_CPU0:~]$exit                                       2

RP/0/RSP0/CPU0:CORE-TOP#
RP/0/RSP0/CPU0:CORE-TOP#admin                                        1

root connected from 192.0.16.4 using ssh on sysadmin-vm:0_RSP0
sysadmin-vm:0_RSP0# run                                              3

[sysadmin-vm:0_RSP0:~]$
[sysadmin-vm:0_RSP0:~]$chvrf 0 bash
[sysadmin-vm:0_RSP0:~]$ssh my_host                                   4

[host:~]$                                                            5
```

Note Exit from any prompt:
a) hit 'Ctrl-d'
b) type 'exit'

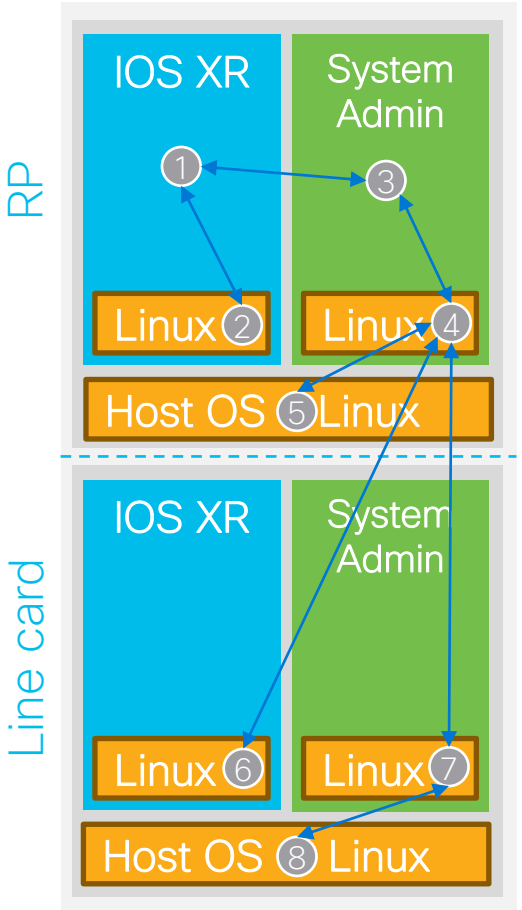
Note "chvrf 0 bash" enables ssh to hostnames:
- ssh my_host
- ssh lc<n>_xr
- ssh lc<n>_admin



IOS XR 64 Bit CLI Access

LC XR VM, LC Admin VM, LC Host

```
RP/0/RSP0/CPU0:CORE-TOP#                                          1
RP/0/RSP0/CPU0:CORE-TOP#admin
aleks connected from 192.0.16.4 using ssh on sysadmin-vm:0_RSP0   3
sysadmin-vm:0_RSP0# run
[sysadmin-vm:0_RSP0:~]$chvrf 0 bash                               4
[sysadmin-vm:0_RSP0:~]$ssh lc5_xr
Last login: Sun Nov 10 17:02:15 2019 from 192.0.16.1              6
[xr-vm_node0_5_CPU0
[xr-vm_node0_5_CPU0:~]$exit
Connection to lc5_xr closed.                                      4
[sysadmin-vm:0_RSP0:~]$ssh lc5_admin
Last login: Sun Nov 10 17:02:25 2019 from 192.0.16.1              7
[sysadmin-vm:0_5:~]$
[sysadmin-vm:0_5:~]$ ssh my_host
Last login: Sun Nov 10 17:02:32 2019 from 10.0.2.15               8
[host:~]$
```



IOS XR 64 Bit CLI Access Example

RAM available to IOS XR on 32-bit vs 64-bit

32-bit

```
RP/0/RSP1/CPU0:ariad#sh platform | i 8x100GE
Node               Type                       State             Config State
0/3/CPU0           A9K-8X100GE-SE             IOS XR RUN        PWR,NSHUT,MON
RP/0/RSP1/CPU0:ariad#sh memory summary location 0/3/CPU0
node:      node0_3_CPU0
------------------------------------------------------------------
Physical Memory: 24576M total
 Application Memory : 24253M (18554M available)
 Image: 82M (bootram: 82M)
 Reserved: 224M, IOMem: 0, flashfsys: 0
 Total shared window: 431M
```

64-bit

```
RP/0/RSP0/CPU0:CORE-TOP#sh platform | i 8X100GE
0/5/CPU0           A9K-8X100GE-TR                IOS XR RUN        NSHUT
RP/0/RSP0/CPU0:CORE-TOP#sh memory summ loc 0/5/cpu0
node:      node0_5_CPU0
------------------------------------------------------------------
 Physical Memory: 10691M total (4939M available)
 Application Memory : 10691M (4939M available)
 Image: 4M (bootram: 0M)
 Reserved: 0M, IOMem: 0M, flashfsys: 0M
 Total shared window: 311M
```

Both LCs are 8x100G Tomahawk LC

- TR/SE have the same size RAM available to LC CPU
- TR/SE differ in RAM available to NP

Different RAM size available to IOS XR

XR Configuration Key Concepts

- Two Stage Commit
- Config History Database
- Rollback
- Atomic vs. Best Effort
- Multiple Config Sessions

Commit

Two stage

Enter Proposed Changes

```
interface gig 0/3/0/0
  ipv4 address 9.9.9.9/24

router ospf 100
  area 0
    interface gig 0/3/0/0
  area 1
    interface ten 0/2/0/0.1
```

Target Configuration

Commit

Changes take effect

Active Configuration
Before Commit

```
hostname Leif
line default
 exec-timeout 1440 0
!
taskgroup ops
 task read boot
 task write boot
 task execute bgp
!
router static
```

Active Configuration
After Commit

```
hostname Leif
line default
 exec-timeout 1440 0
!
Interface GigabitEthernet0/3/0/0
  ipv4 address 9.9.9.9/24
!
taskgroup ops
 task read boot
 task write boot
 task execute bgp
!
router ospf 100
  area 0
    interface GigabitEthernet0/3/0/0
  area 1
    interface TenGigE0/2/0/0.1
!
router static
```

Commit: Verification

Syntax by line, then logic at commit

Syntax Check
after each line

```
interface gig 0/3/0/0
  ipv4 address 9.9.9.9/24

router ospf 100
  area 0
    interface gig 0/3/0/0
  area 1
    interface ten 0/2/0/0.1
```

Target Configuration

Semantic Check
during commit

Active Configuration
Before Commit

```
hostname Leif
line default
 exec-timeout 1440 0
!
taskgroup ops
 task read boot
 task write boot
 task execute bgp
!
router static
```

Active Configuration
After Commit

```
hostname Leif
line default
 exec-timeout 1440 0
!
Interface GigabitEthernet0/3/0/0
  ipv4 address 9.9.9.9/24
!
taskgroup ops
 task read boot
 task write boot
 task execute bgp
!
router ospf 100
  area 0
    interface GigabitEthernet0/3/0/0
  area 1
    interface TenGigE0/2/0/0.1
!
router static
```

Commit Behavior: Atomic

Commit all or nothing [default behavior]

Syntax Check
after each line
PASSES

```
Interface GigabitEthernet0/3/0/0
  ipv4 address 9.9.9.9/24
taskgroup bgp
  task read bgp
  task write bgp
```

Target Configuration

Semantic Check
during commit
FAILS
BGP cannot be
taskgroup name

Active Configuration
Before Commit

```
hostname odin
line default
 exec-timeout 1440 0
!
taskgroup ops
 task read boot
 task write boot
 task execute bgp
!
router static
 address-family ipv4 unicast
  0.0.0.0/0 7.1.9.1
  7.7.7.77/32 7.1.9.1
```

Active Configuration
After Commit

No Change

```
hostname odin
line default
 exec-timeout 1440 0
!
taskgroup ops
 task read boot
 task write boot
 task execute bgp
!
router static
 address-family ipv4 unicast
  0.0.0.0/0 7.1.9.1
  7.7.7.77/32 7.1.9.1
```

Commit Behavior: Best Effort

Commit valid sections

Syntax Check
after each line
PASSES

```
Interface GigabitEthernet0/3/0/0
  ipv4 address 9.9.9.9/24
taskgroup bgp
  task read bgp
  task write bgp
```

Target Configuration

Semantic Check
during commit
FAILS
BGP cannot be
taskgroup name

Active Configuration
Before Commit

```
hostname odin
line default
 exec-timeout 1440 0
!
taskgroup ops
 task read boot
 task write boot
 task execute bgp
!
router static
 address-family ipv4 unicast
  0.0.0.0/0 7.1.9.1
  7.7.7.77/32 7.1.9.1
```

Active Configuration
After Commit

Partial Commit

```
hostname Olav
line default
 exec-timeout 1440 0
!
Interface GigabitEthernet0/3/0/0
  ipv4 address 9.9.9.9/24
!
taskgroup ops
 task read boot
 task write boot
 task execute bgp
!
```

Concurrent Config Sessions

Concurrent and exclusive

Enter Proposed Changes

```
interface gig 0/3/0/0
  ipv4 address 9.9.9.9/24

router ospf 100
  area 0
    interface gig 0/3/0/0
  area 1
    interface gig 0/4/0/0
```

First to Commit

Normal Commit
only first user's changes

Enter Proposed Changes

```
interface gig 0/3/0/0
  ipv4 address 9.9.9.7/24

router ospf 100
  area 2
    interface gig 0/3/0/0
  area 4
    interface gig 0/4/0/0
```

Second to Commit

Use **`config exclusive`** to block other users from committing

```
One or more commits have occurred from
other configuration sessions since this
session started or since the last commit
was made from this session.
You can use the 'show configuration commit
changes' command to browse the changes.
Do you wish to proceed with this commit
anyway? [no]:
```

Atomic Configuration Replace – NEW Behavior

Replace/remove/add interface config lines witout removing interface

1 Original Configuration

```
RP/0/RSP0/CPU0:PE1#sh run int gigabitEthernet 0/0/0/19
Mon Feb 16 13:00:32.153 UTC
interface GigabitEthernet0/0/0/19
 description ***AAABBBCCC***
 cdp
 ipv4 address 13.3.5.5 255.255.255.0
 negotiation auto
 shutdown
 load-interval 30
!
```

2 Target Configuration

```
RP/0/RSP0/CPU0:PE1(config)#no interface GigabitEthernet0/0/0/19
RP/0/RSP0/CPU0:PE1(config)#
RP/0/RSP0/CPU0:PE1(config)#interface GigabitEthernet0/0/0/19
RP/0/RSP0/CPU0:PE1(config-if)# ipv6 address 2603:10b0:100:10::21/126
RP/0/RSP0/CPU0:pE1(config-if)# commit
```

Example:
Consider an interface with a target config where all config lines are new

3 Committed Configuration

```
RP/0/RSP0/CPU0:PE1#show configuration commit changes last 1
Mon Feb 16 13:15:36.655 UTC
Building configuration...
!! IOS XR Configuration 5.1.2
interface GigabitEthernet0/0/0/19
 no description ***AAABBBCCC***
 no cdp
 no ipv4 address 13.3.5.5 255.255.255.0
 ipv6 address 2603:10b0:100:10::21/126
 no negotiation auto
 no shutdown
 no load-interval 30
!
end
```

NEW Behavior:
When issuing the “no” interface config, the system does not destroy the subtree but instead performs a SET of new config and REMOVE of unwanted config lines

Monitoring Configuration

Configuration, commits, changes, sessions

```
RP/0/RSP0/CPU0:rasr9000-2w-b#show running-config
Mon Jan  5 00:35:19.951 EST
Building configuration...
!! IOS XR Configuration 5.1.3
!! Last configuration change at Tue Dec  2 22:19:25 2014 by cisco
!
service unsupported-transceiver
hostname rasr9000-2w-b
clock timezone EST -5
.
RP/0/RSP0/CPU0:rasr9000-2w-b#show configuration commit list
Mon Jan  5 00:35:34.747 EST
SNo. Label/ID    User      Line                Client      Time Stamp
~~~~ ~~~~~~~~~   ~~~~      ~~~~                ~~~~~~      ~~~~~~~~~~
1    1000000033  cisco     con0_RSP0_CPU0      CLI         Tue Dec  2 22:19:25 2014
2    1000000032  cisco     con0_RSP0_CPU0      CLI         Tue Dec  2 22:15:54 2014
3    1000000031  cisco     con0_RSP0_CPU0      CLI         Tue Dec  2 22:14:18 2014
.
RP/0/RSP0/CPU0:rasr9000-2w-b#show configuration commit changes last 5
Mon Jan  5 00:36:17.813 EST
Building configuration...
!! IOS XR Configuration 5.1.3
interface GigabitEthernet0/1/0/4
 ipv4 address 10.6.7.6 255.255.255.0
.
RP/0/RSP1/CPU0:rasr9000-2w-b# show config sessions
Current Configuration Session  Line       User     Date                      Lock
00000051-004c4104-00000000     con0_RSP1_ ww       Tue Jul 21 16:58:22 2009
.
RP/0/RSP1/CPU0:rasr9000-2w-b# rollback configuration last 1
```

Configuration Consistency

If the system detects inconsistency: A downgrade example

```
RP/0/RSP0/CPU0:Jan  2 14:11:35.876 : cfgmgr-rp[161]: %MGBL-CONFIGCLI-3-BATCH_CONFIG_FAIL : 3
config(s) failed during startup. To view failed config(s) use the command - "show configuration
failed startup"

RP/0/RSP0/CPU0:Jan  2 14:11:35.884 : cfgmgr-rp[161]: %MGBL-CONFIG-3-INCONSISTENCY_ALARM : A
configuration inconsistency alarm has been raised. Configuration commits will be blocked until
'clear configuration inconsistency' command has been run to synchronize persistent configuration
with running configuration.

RP/0/RSP0/CPU0:rasr9000-2w-b#configure
Thu Jan  9 20:51:57.470 EST
This SDR's running configuration is inconsistent with persistent configuration.
No configuration commits for this SDR will be allowed until a 'clear configuration
inconsistency' command is performed.

RP/0/RSP0/CPU0:rasr9000-2w-b#show configuration persistent diff
Thu Jan  9 20:54:02.570 EST
Building configuration...
!! IOS XR Configuration 4.2.3
End
```

Configuration Consistency

If the system detects inconsistency: A downgrade example

```
RP/0/RSP0/CPU0:rasr9000-2w-b#show configuration failed startup
Thu Jan  9 21:00:33.389 EST
!!14:11:35 UTC Thu Jan 02 2014
!! SYNTAX/AUTHORIZATION ERRORS: This configuration failed due to
!! one or more of the following reasons:
!!  - the entered commands do not exist,
!!  - the entered commands have errors in their syntax,
!!  - the software packages containing the commands are not active,
!!  - the current user is not a member of a task-group that has
!!    permissions to use the commands.

router bgp 65001
l2vpn
 xconnect group PW-10
  p2p PW-10
   neighbor ipv4 192.168.10.2 pw-id 10
    pw-class PW-GRE

RP/0/RSP0/CPU0:rasr9000-2w-b#clear configuration inconsistency
Thu Jan  9 20:54:46.435 EST

Creating any missing directories in Configuration File system...OK
Initializing Configuration Version Manager...OK
Syncing commit database with running configuration...OK
RP/0/RSP0/CPU0:Jan  9 20:54:48.946 : cfgmgr-rp[161]: cfgmgr_retry_mgr_sysdb_notification_fn Item is deleted
RP/0/RSP0/CPU0:Jan  9 20:54:48.948 : cfgmgr_cfs_check[65784]: %MGBL-CONFIG-3-INCONSISTENCY_ALARM : A configuration
inconsistency alarm has been cleared.  Configuration commits within this SDR are no longer blocked.
```

Agenda

- ✓ **System Architecture**: System anatomy & health
- ✓ **Operating System & Configuration: IOS-XR & configuration models**
- ➢ **Control, Management, Security**: Processing of control & exceptions
- ➢ **Transit Packet/Frame Journey**: Life of L3/L2 unicast/multicast
- ➢ **MPLS Operation**: Processing, forwarding & L3/L2 service operation
- ➢ **Troubleshooting**: Diagnostics, counters, drops, and packet capture

3 Control, Management, & Security

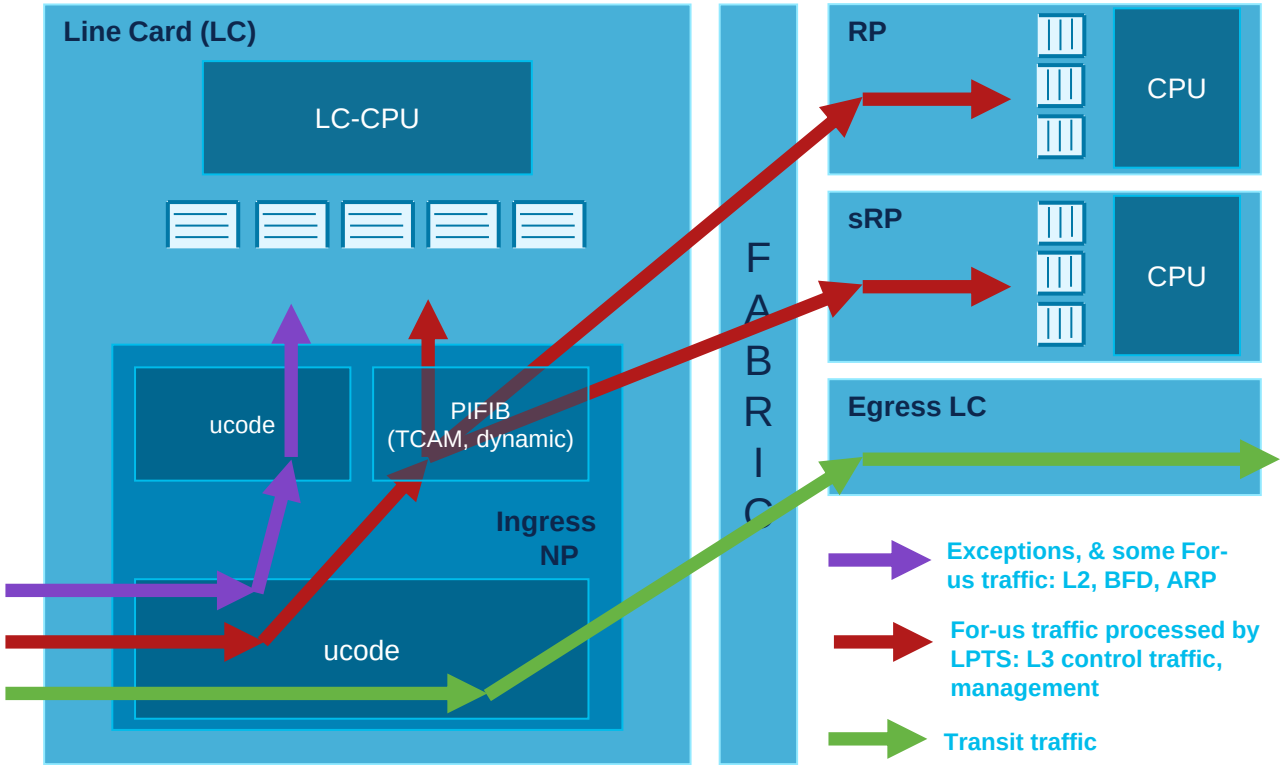
cisco Live!

For Us, Exceptions, & Resource Protection

Traffic: Transit, For us, and Exceptions

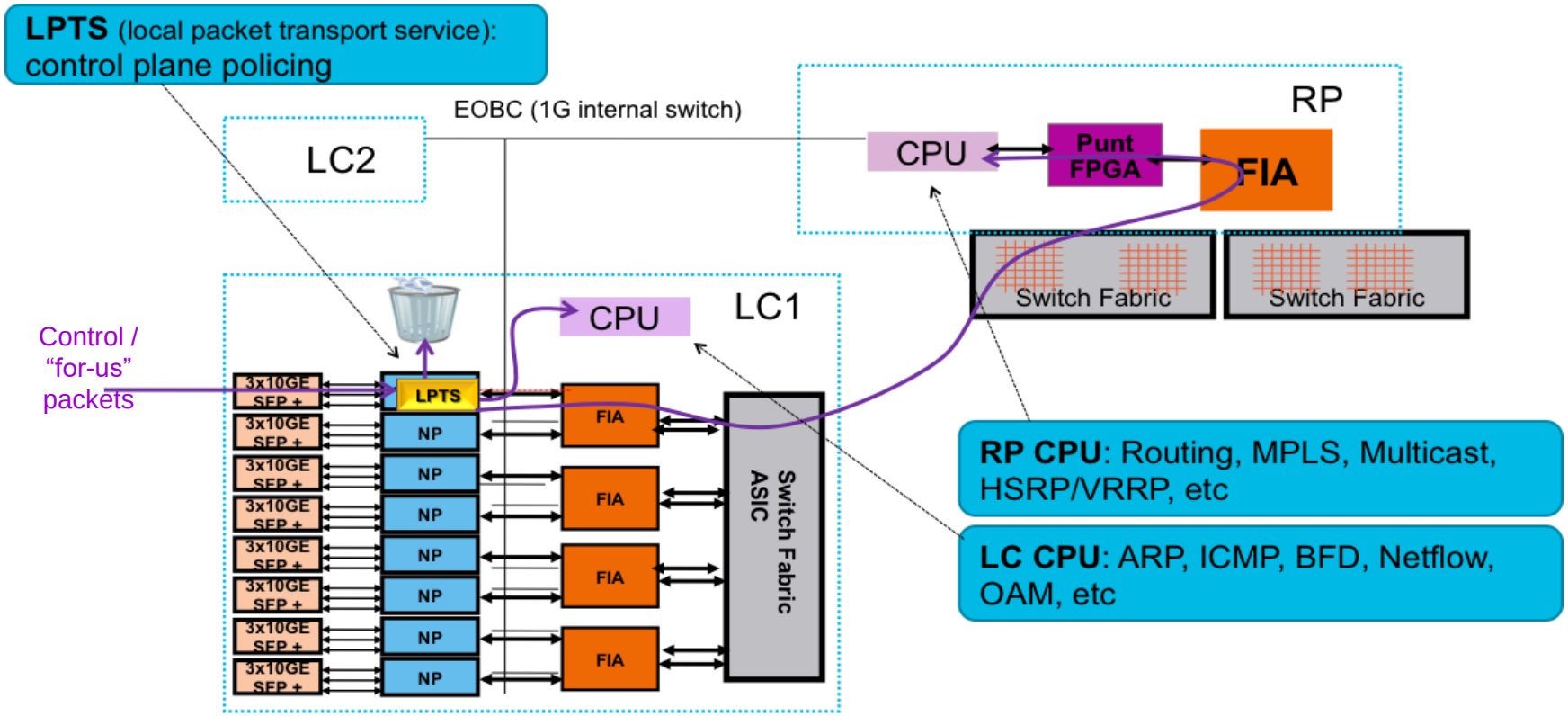
Differentiate on ingress NP

- **Transit**
  - Look up, re-write, forward
- **For us**
  - Destined to RP, or link local scope
  - Punt to RP or ingress LC CPU
- **Exception**
  - MTU failure, TTL failure, etc. Should have been transit
  - Punt to LC CPU



For Us Frame Path

From ingress NP to RP CPU or LC CPU



For Us Frame Path

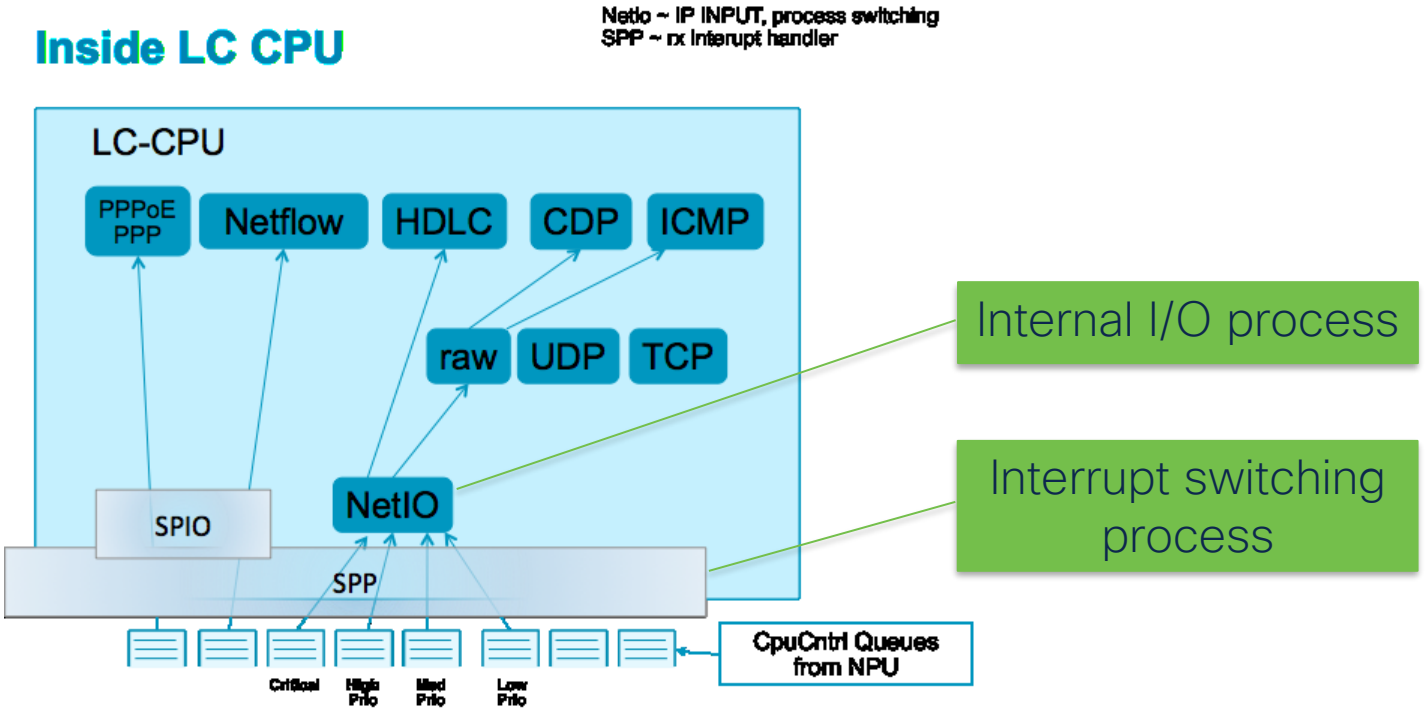
The internal FIB [IFIB]

```
RP/0/RSP0/CPU0:rasr9000-2w-b#show lpts ifib brief
Wed Dec 28 11:58:43.726 EST
 Slice     VRF-ID    L4      Interface        Dlvr          Local-Address,Port Remote-Address,Port
 --------- --------- ------- ---------------  ------------- ----------------------------------
 RAWIP4    TRAFFIC   112     Te0/0/0/2.200    0/RSP0/CPU0 224.0.0.18 any
 RAWIP4    default   RSVP    Gi0/1/0/3.400    [0x0003]      any any
 RAWIP4    default   RSVP    BE1              [0x0003]      any any
 RAWIP4    default   RSVP    Te0/0/0/4.100    [0x0003]      any any
 RAWIP4    default   IGMP    any              [0x0003]      any any
 BGP4      default   TCP     any              0/RSP0/CPU0 10.101.188.1,179 10.100.102.1,48462
 BGP4      default   TCP     any              0/RSP0/CPU0 any,179 10.100.102.1
 BGP6      default   TCP     any              0/RSP0/CPU0 2001:db8:1:a06::,179 2001:db8:1:c06::
 BGP6      default   TCP     any              0/RSP0/CPU0 2001:db8:1:a06::,60698 2001:db8:1:c06::,179
 UDP4      default   UDP     any              [0x0003]      10.101.188.1,646 10.100.108.1
 UDP4      default   UDP     any              [0x0003]      10.101.188.1,646 10.101.111.1
 TCP4      default   TCP     Mg0/RSP1/CPU0/0  0/RSP0/CPU0 any,23 any
 TCP4      default   TCP     any              [0x0003]      10.101.188.1,59192 10.101.111.1,646
 TCP4      default   TCP     Gi0/1/0/1        0/RSP0/CPU0 any,38751 any
 TCP4      default   TCP     Mg0/RSP1/CPU0/0  0/RSP0/CPU0 any,38751 any
 TCP4      default   TCP     any              [0x0003]      10.101.188.1,63675 10.100.108.1,646
 ISIS      default   -       BE1              [0x0003]      - -
 ISIS      default   -       Te0/0/0/4.100    [0x0003]      - -
```

Physical slot mask: 0003 is first 2 slots in 9006: RSP0 & RSP1

For Us Frame Path

From ingress NP to RP CPU or LC CPU



For Us Frame Path

Processes to watch on RP CPU

- netio on RP CPU
- Example for BGP from unknown
  - LPTS relaxed for simulation

```
RP/0/RSP0/CPU0:rasr9k-1y#show processes cpu location 0/RSP0/CPU0 |
exclude " 0%       0%       0%"
Wed Nov 28 01:36:52.203 UTC

CPU utilization for one minute: 26%; five minutes: 25%; fifteen minutes: 22%

PID    1Min    5Min    15Min Process
94243    3%      3%       3% spp
254074  23%     22%      19% netio

RP/0/RSP0/CPU0:rasr9k-1y#show lpts pifib hardware police location 0/0/CPU0 | e$
Wed Nov 28 01:23:10.907 UTC
------------------------------------------------------------
               Node 0/0/CPU0:
------------------------------------------------------------
 Burst = 100ms for all flow types
------------------------------------------------------------
FlowType              Policer Type    Cur. Rate  Def. Rate  Accepted            Dropped
--------------------- ------- ------- ---------- ---------- ------------------- -----------------
-
.
BGP-default           108     Local   150000     1500       89395477            3845915191
.
TCP-default           164     Local   150000     2000       49872016            8066163019
.
------------------------
statistics:
Packets accepted by deleted entries: 19477
Packets dropped by deleted entries: 0
Run out of statistics counter errors: 0
```

For Us Frame Path

Processes to watch on LC CPU

- netio and spp on LC CPU
  - netio for internal in/out (like ip input in IOS)
  - spp for software switched (similar to interrupt switching in IOS)
- Example for for-us fragments
  - LPTS relaxed for simulation

```
RP/0/RSP0/CPU0:rasr9k-1y#show processes cpu location 0/0/CPU0 |
exclude " 0%       0%       0%"
Wed Nov 28 01:28:52.281 UTC

CPU utilization for one minute: 46%; five minutes: 48%; fifteen
minutes: 39%

PID    1Min     5Min     15Min Process
45085   22%      23%       22% spp
180316  23%      23%       23% netio

RP/0/RSP0/CPU0:rasr9k-1y#show lpts pifib hardware police location 0/0/CPU0
Wed Nov 28 01:23:10.907 UTC
--------------------------------------------------------------
               Node 0/0/CPU0:
--------------------------------------------------------------
 Burst = 100ms for all flow types
--------------------------------------------------------------
FlowType              Policer Type    Cur. Rate  Def. Rate  Accepted             Dropped
--------------------- ------- ------- ---------- ---------- -------------------- ------------------
-
Fragment              101     Local   100000     2500       142076716            5033837819
.
UDP-default           163     Local   1000000    3500       38336274             2376859
.
------------------------
statistics:
Packets accepted by deleted entries: 19477
Packets dropped by deleted entries: 0
Run out of statistics counter errors: 0
```

For Us Frame Path

Traffic rate from/to each CPU

```
RP/0/RSP0/CPU0:rasr9k-1y#show netio rates
location 0/RSP0/CPU0
Thu Jan  3 06:56:28.745 UTC

Netio packet rate for node 0/RSP0/CPU0
----------------------------------
Current rate       (updated 0 seconds ago):
  Input:           82811 pkts/s
  Output:            100 pkts/s
  Driver Output:     100 pkts/s

 1 minute rate     (updated 7 seconds ago):
  Input:           82668 pkts/s
  Output:             98 pkts/s
  Driver Output:      98 pkts/s

 5 minute rate     (updated 7 seconds ago):
  Input:           57073 pkts/s
  Outpu
  Drive
```

RSP: routing protocols, management, etc.

```
RP/0/RSP0/CPU0:rasr9k-1y#show netio rates
location 0/0/CPU0
Thu Jan  3 06:56:20.329 UTC

Netio packet rate for node 0/0/CPU0
----------------------------------
Current rate       (updated 0 seconds ago):
  Input:           14759 pkts/s
  Output:              0 pkts/s
  Driver Output:   14760 pkts/s

 1 minute rate     (updated 0 seconds ago):
  Input:           14770 pkts/s
  Output:              0 pkts/s
  Driver Output:   14771 pkts/s

 5 minute rate     (updated 0 seconds ago):
  Input:           10178 pkts/s
  Output:              0 pkts/s
  Drive
```

LC: fragments, BFD, ARP, L2, etc.

Control Plane Protection

LPTS flow type policers

```
RP/0/RSP0/CPU0:rasr9k-1y#show lpts pifib hardware police location 0/0/CPU0
Wed Dec 28 06:07:36.931 EST
--------------------------------------------------------
                Node 0/0/CPU0:
--------------------------------------------------------
 Burst = 100ms for all flow types
--------------------------------------------------------
```

FlowType	Policer	Type	Cur. Rate	Def. Rate	Accepted	Dropped	TOS Value
unconfigured-default	100	Static	2500	2500	0	0	01234567
Fragment	101	Local	0	2500	0	0	01234567
OSPF-mc-known	102	Static	2000	2000	0	0	01234567
OSPF-mc-default	103	Local	0	1500	53	26	01234567
OSPF-uc-known	104	Static	2000	2000	0	0	01234567
OSPF-uc-default	105	Local	0	1000	0	0	01234567
ISIS-known	143	Static	2000	2000	20890	0	01234567
ISIS-default	144	Local	0	1500	0	0	01234567
.							
BGP-known	106	Static	2500	2500	4070	0	01234567
BGP-cfg-peer	107	Static	2000	2000	17	0	01234567
BGP-default	108	Local	0	1500	335787	15570288947	01234567
PIM-mcast-default	109	Local	0	2000	0	0	01234567
PIM-mcast-known	176	Static	2000	2000	0	0	01234567
PIM-ucast	110	Static	1500	1500	0	0	01234567
IGMP	111	Static	3000	3000	0	0	01234567

Control Plane Protection

LPTS flow type policers

FlowType	Policer	Type	Cur. Rate	Def. Rate	Accepted	Dropped	TOS Value
ICMP-local	112	Static	1500	1500	20044	0	01234567
ICMP-app	152	Local	100	1500	0	0	01234567
ICMP-control	140	Static	1000	1000	0	0	01234567
ICMP-default	153	Local	100	1500	0	0	01234567
ICMP-app-default	152	Local	100	1500	0	0	01234567
LDP-TCP-known	113	Static	2500	2500	0	0	01234567
LDP-TCP-cfg-peer	114	Static	2000	2000	0	0	01234567
LDP-TCP-default	115	Local	0	1500	0	0	01234567
LDP-UDP	116	Static	2000	2000	0	0	01234567
All-routers	117	Local	0	1000	0	0	01234567
.							
RSVP-UDP	118	Static	2000	2000	0	0	01234567
RSVP-default	154	Local	0	500	0	0	01234567
RSVP-known	177	Static	7000	7000	0	0	01234567
IKE	119	Static	100	100	0	0	01234567
IPSEC-known	120	Static	400	400	0	0	01234567
IPSEC-default	121	Local	0	100	0	0	01234567
MSDP-known	122	Static	300	300	0	0	01234567
.							
MIPv6	88	Static	5000	5000	0	0	01234567
DHCPv4	92	Static	4000	4000	0	0	01234567
DHCPv6	93	Static	4000	4000	0	0	01234567
.							

Control Plane Protection

LPTS flow types: BGP example

Established session packets

Configured peer packets

BGP packets from unknown

```
RP/0/RSP0/CPU0:rasr9k-1y#show lpts pifib hardware police location 0/0/CPU0
Wed Nov 28 03:01:48.271 UTC
--------------------------------------------------------
                Node 0/0/CPU0:
--------------------------------------------------------
 Burst = 100ms for all flow types
--------------------------------------------------------
FlowType          Policer Type  Cur. Rate Def. Rate  Accepted        Dropped
----------------- ------- ----- --------- ---------  --------------- -----------
.
BGP-known         106    Local  50000      2500       2590            0

BGP-cfg-peer      107    Static 2000       2000       13              0

BGP-default       108    Local  400000     1500       138918630       3848639925
.

-------------------------
statistics:
Packets accepted by deleted entries: 19477
Packets dropped by deleted entries: 0
Run out of statistics counter errors: 0
```

Control Plane Protection

Customize LPTS flow rates

```
RP/0/RSP0/CPU0:rasr9k-1y(config)#lpts pifib hardware police location 0/0/CPU0
RP/0/RSP0/CPU0:rasr9k-(config-pifib-policer-per-node)# flow isis default rate 0
RP/0/RSP0/CPU0:rasr9k-(config-pifib-policer-per-node)# flow bgp configured rate 500
RP/0/RSP0/CPU0:rasr9k-(config-pifib-policer-per-node)# flow bgp default rate 0
RP/0/RSP0/CPU0:rasr9k-(config-pifib-policer-per-node)# flow pim multicast default rate 0
RP/0/RSP0/CPU0:rasr9k-(config-pifib-policer-per-node)# flow icmp application rate 100
RP/0/RSP0/CPU0:rasr9k-(config-pifib-policer-per-node)# flow icmp default rate 100
RP/0/RSP0/CPU0:rasr9k-(config-pifib-policer-per-node)# flow ldp tcp default rate 0
RP/0/RSP0/CPU0:rasr9k-(config-pifib-policer-per-node)# flow all-routers rate 0
.
RP/0/RSP0/CPU0:rasr9k-1y#show running-config lpts pifib hardware police location 0/0/CPU0
Sun Dec  2 06:29:11.493 UTC
lpts pifib hardware police location 0/0/CPU0
 flow bgp default rate 0
 flow pim multicast default rate 0
 flow icmp application rate 100
 flow icmp default rate 100
 flow ldp tcp default rate 0
 flow all-routers rate 0
 flow lmp tcp default rate 0
 flow lmp udp rate 0
 flow rsvp default rate 0
 flow ipsec default rate 0
 flow msdp default rate 0
 flow ssh known rate 0
 flow ssh default rate 0
```

Control Plane Protection

LPTS flow policers

```
RP/0/RSP0/CPU0:rasr9k-1y#show lpts pifib hardware police location 0/0/CPU0
Sun Dec  2 06:32:04.344 UTC
--------------------------------------------------------
                Node 0/0/CPU0:
--------------------------------------------------------
 Burst = 100ms for all flow types
--------------------------------------------------------
FlowType                Policer Type    Cur. Rate  Def. Rate  Accepted            Dropped
----------------------- ------- ------- ---------- ---------- ------------------- --------------------
unconfigured-default    100     Static  2500       2500       0                   0
Fragment                101     Local   0          2500       0                   0
OSPF-mc-known           102     Static  2000       2000       0                   0
OSPF-mc-default         103     Local   0          1500       54                  27
OSPF-uc-known           104     Static  2000       2000       0                   0
OSPF-uc-default         105     Local   0          1000       0                   0
ISIS-known              143     Static  2000       2000       21078               0
ISIS-default            144     Local   0          1500       0                   0
TCP-known               156     Static  2500       2500       0                   0
TCP-listen              157     Static  2500       2500       0                   0
TCP-cfg-peer            158     Static  2000       2000       0                   0
TCP-default             164     Local   0          2000       95977990            1995220219679
Mcast-known             159     Static  2500       2500       0                   0
RADIUS                  174     Local   0          2000       0                   0
TACACS                  175     Static  2000       2000       0                   0
NTP-default             126     Local   0          200        0                   0
NTP-known               180     Local   0          200        0                   0
```

Control Plane Protection

LPTS PIFIB

- LPTS is the group of processes to transport for-us packets
  - Destination is either RP CPU's or ingress LC CPU
  - 5 queues of different priorities in Typhoon NP
- LPTS policers
  - Configured LC flow rate applied to LC, if not then configured global flow rate applied, if not then a default rate applied
  - Enforced by each NP
  - Flow entries created and installed based on: configuration and neighbor flow state [e.g. BGP TCP]

For Us Packet Forwarding Entries

LPTS flow entries

```
RP/0/RSP0/CPU0:ASR9006-2w-a.PE2#show lpts pifib hardware entry statistics location 0/0/CPU0
.
Offset L3   VRD id       L4     Intf            Dest       Pkts/Drops      laddr,Port raddr,Port
------ ---- ------------ ------ --------------- ---------- --------------- --------------------
8      IPV4 *            any    any             Local      0/0             any,any any,any
9      CLNS *            -      BE1             LM[3]      0/0             - -
10     CLNS *            -      Te0/0/0/1       LM[3]      59571/0         - -
11     CLNS *            -      Te0/0/0/4.100   LM[3]      0/0             - -
12     CLNS *            -      any             LU(30)     8/0             - -
13     IPV4 *            ICMP   any             Local      0/0             any,any any,ECHO
14     IPV4 default      RSVP   Te0/0/0/1       Local      15120/0         any,any any,any
15     IPV4 default      TCP    any             LM[3]      16991/0         any,65264 10.10.1.1,179
16     IPV4 default      TCP    any             LU(30)     19377/0         any,42370 10.10.1.1,646
.

------------------------
statistics:

Type            Num. Entries         Pkts
------          ------------         ----
IPv4            58                   151029/0
IPv6            39                   0/0
Packets accepted by deleted entries: 5
Packets dropped by deleted entries: 0
Run out of statistics counter errors: 0
```

For Us Packet Forwarding Entries

LPTS flow entries

```
RP/0/RSP0/CPU0:rasr9k-1y#show lpts pifib
hardware entry location 0/0/CPU0
Sun Dec  2 00:46:50.573 UTC

               Node: 0/0/CPU0:
----------------------------------------
M - Fabric Multicast;
L - Listener Tag; T - Min TTL;
F - Flow Type;
DestNode - Destination Node;
DestAddr - Destination Fabric queue;
SID - Stream ID;
Po - Policer; Ct - Stats Counter;
Lp - Lookup priority; Sp - Storage
Priority;
Ar - Average rate limit; Bu - Burst;
HAr - Hardware Average rate limit; HBu
- Hardware Burst;
Cir - Committed Information rate in
HAL
Rsp - Relative sorting position;
Rtp - Relative TCAM position;
na - Not Applicable or Not Available
.
```

Show flow policers in LC TCAM.

BGP-known
Session already established. Flow parameters in hardware policer.

```
.
-----------------------------------------------
VRF ID              : 0x60000000
Destination IP      : 2001:db8:1:a06::
Source IP           : 2001:db8:1:c06::
Is Fragment         : 0
Interface           : any
M/L/T/F             :
1/IPv6_STACK/0/BGP-known
DestNode            : 48
DestAddr            : 48
SID                 : 7
L4 Protocol         : TCP
TCP flag byte       : any
Source port         : Port:60698
Destination Port    : 179
Ct                  : 0x5f0690
Accepted/Dropped    : 3189/0
Lp/Sp               : 1/255
# of TCAM entries   : 1
HPo/HAr/HBu/Cir     :
15794309/2500pps/1250ms/2500pps
State               : Entry in TCAM
Rsp/Rtp             : 5/15
-----------------------------------------------
.
```

For Us Packet Forwarding Entries

LPTS flow entries

```
------------------------------------------------
VRF ID              : 0x60000000
Destination IP      : any
Source IP           : 192.1.1.2
Is Fragment         : 0
Interface           : any
M/L/T/F             :
0/IPv4_STACK/255/BGP-known
DestNode            : 48
DestAddr            : 48
SID                 : 7
L4 Protocol         : TCP
TCP flag byte       : any
Source port         : Port:179
Destination Port    : 41243
Ct                  : 0x5f0670
Accepted/Dropped    : 0/0
Lp/Sp               : 1/255
# of TCAM entries   : 1
HPo/HAr/HBu/Cir     :
15794309/2500pps/1250ms/2500pps
State               : Entry in TCAM
Rsp/Rtp             : 6/16
------------------------------------------------
```

BGP-known
Active session with a configured peer.

BGP-cfg-peer
Open to receiving peer attempts to establish.

```
------------------------------------------------
VRF ID              : 0x60000000
Destination IP      : 2001:db8:1:a06::
Source IP           : 2001:db8:1:c06::
Is Fragment         : 0
Interface           : any
M/L/T/F             :
0/IPv6_LISTENER/0/BGP-cfg-peer
DestNode            : 48
DestAddr            : 48
SID                 : 8
L4 Protocol         : TCP
TCP flag byte       : any
Source port         : Port:any
Destination Port    : 179
Ct                  : 0x5f0340
Accepted/Dropped    : 0/0
Lp/Sp               : 1/255
# of TCAM entries   : 1
HPo/HAr/HBu/Cir     :
15794310/2000pps/1000ms/2000pps
State               : Entry in TCAM
Rsp/Rtp             : 7/17
------------------------------------------------
```

For Us Packet Forwarding Entries

LPTS flow entries

```
------------------------------------------------
VRF ID              : any
Destination IP      : any
Source IP           : any
Is Fragment         : 0
Interface           : any
M/L/T/F             : 0/BGP4_FM/0/BGP-
default
DestNode            : 48
DestAddr            : 48
SID                 : 9
L4 Protocol         : TCP
TCP flag byte       : any
Source port         : Port:179
Destination Port    : any
Ct                  : 0x5f01b0
Accepted/Dropped    : 300890/13952472426
Lp/Sp               : 1/0
# of TCAM entries   : 1
HPo/HAr/HBu/Cir     :
15794311/0pps/750ms/1pps
State               : Entry in TCAM
Rsp/Rtp             : 13/23
------------------------------------------------
```

BGP-default
Any TCP from port 179 not matching previous entries.

BGP-default
Any TCP to port 179 not matching previous entries.

```
------------------------------------------------
VRF ID              : any
Destination IP      : any
Source IP           : any
Is Fragment         : 0
Interface           : any
M/L/T/F             : 0/BGP4_FM/0/BGP-
default
DestNode            : 48
DestAddr            : 48
SID                 : 9
L4 Protocol         : TCP
TCP flag byte       : any
Source port         : Port:any
Destination Port    : 179
Ct                  : 0x5f01a0
Accepted/Dropped    : 0/0
Lp/Sp               : 1/0
# of TCAM entries   : 1
HPo/HAr/HBu/Cir     :
15794311/0pps/750ms/1pps
State               : Entry in TCAM
Rsp/Rtp             : 15/25
------------------------------------------------
```

For Us Packet Forwarding Entries

LPTS flow entries

```
------------------------------------------------
VRF ID              : any
Destination IP      : any
Source IP           : any
Is Fragment         : 0
Interface           : any
M/L/T/F             : 0/TCP4_FM/0/TCP-
default
DestNode            : 48
DestAddr            : 48
SID                 : 9
L4 Protocol         : TCP
TCP flag byte       : any
Source port         : Port:any
Destination Port    : any
Ct                  : 0x5f0170
Accepted/Dropped    :
95947801/1817465391676
Lp/Sp               : 1/0
# of TCAM entries   : 1
HPo/HAr/HBu/Cir     :
15794367/0pps/1000ms/1pps
State               : Entry in TCAM
Rsp/Rtp             : 24/34
------------------------------------------------
```

TCP-default
Any IPv4 TCP not matched by previous entries.

Raw-default
Any IPv4 not matched by previous entries.

```
------------------------------------------------
VRF ID              : any
Destination IP      : any
Source IP           : any
Is Fragment         : 0
Interface           : any
M/L/T/F             : 0/RAWIP4_FM/0/Raw-
default
DestNode            : 48
DestAddr            : 48
SID                 : 9
L4 Protocol         : any
Source port         : any
Destination Port    : any
Ct                  : 0x5f01f0
Accepted/Dropped    : 10272/18857
Lp/Sp               : 1/0
# of TCAM entries   : 1
HPo/HAr/HBu/Cir     :
15794370/0pps/1250ms/1pps
State               : Entry in TCAM
Rsp/Rtp             : 28/38
------------------------------------------------
```

Control Plane Protection

LPTS: PIFIB ACL-Based Policers

```
RP/0/RSP0/CPU0:rasr9000-2w-b#show running-config
ipv4 access-list
Thu Apr  3 18:21:35.034 EDT
ipv4 access-list PE
 10 remark PE LOOPBACKS
 20 permit ipv4 10.101.0.1 0.0.255.0 any
 40 deny ipv4 any any
!
ipv4 access-list CORE
 10 permit ipv4 10.100.0.0/16 any
 20 deny ipv4 any any
!
ipv4 access-list OFFENDERS
 10 permit ipv4 host 172.19.19.1 any
 20 permit ipv4 host 172.19.19.15 any
 30 permit ipv4 172.19.19.224/29 any
 40 deny ipv4 any any
!
.
```

```
RP/0/RSP0/CPU0:rasr9000-2w-b#show running-config
lpts pifib hardware police
Thu Apr  3 18:25:22.831 EDT
lpts pifib hardware police
 acl PE rate 11000
 flow bgp known rate 6000
 flow bgp configured rate 1000
 flow bgp default rate 0
 acl CORE rate 33000
 acl OFFENDERS rate 0
!
```

"per-ACL" PPS rate

Control Plane Protection

LPTS: PIFIB ACL-Based Policers

```
RP/0/RSP0/CPU0:rasr9000-2w-b#show lpts pifib hardware entry brief location 0/1/CPU0
Thu Apr  3 18:28:57.713 EDT

             Node: 0/0/CPU0:
----------------------------------------
L3 - L3 Protocol;L4 - Layer4 Protocol; Intf - Interface;
Dest - Destination Node; V - Virtual;
na - Not Applicable or Not Available;
LU - Local chassis fabric unicast;
LM - Local chassis fabric multicast;
RU - Multi chassis fabric unicast;
RM - Multi chassis fabric multicast;
def - default

Offset L3   VRF id       L4     Intf            Dest       laddr,Port raddr,Port          acl name
------ ---- ------------ ------ --------------- ---------  ----------                     ----------------------------
.
18     IPV4 default      RSVP   Gi0/1/0/3.400   Local      any,any any,any
19     IPV4 default      TCP    any             LU(30)     any,179 10.100.104.1,28603     CORE
20     IPV4 default      TCP    any             LU(30)     any,40607 10.100.102.1,179     CORE
21     IPV4 default      TCP    any             LM[3]      any,38362 10.100.108.1,646     CORE
22     IPV4 default      UDP    any             LM[3]      any,646 192.168.10.2,any
23     IPV4 default      UDP    any             LM[3]      any,646 10.100.108.1,any       CORE
24     IPV4 default      TCP    any             LU(30)     any,179 10.100.102.1,any       CORE
25     IPV4 default      TCP    any             LU(30)     any,179 10.100.104.1,any       CORE
26     IPV4 default      TCP    any             LU(30)     any,23 any,any
.
```

Control Plane Protection

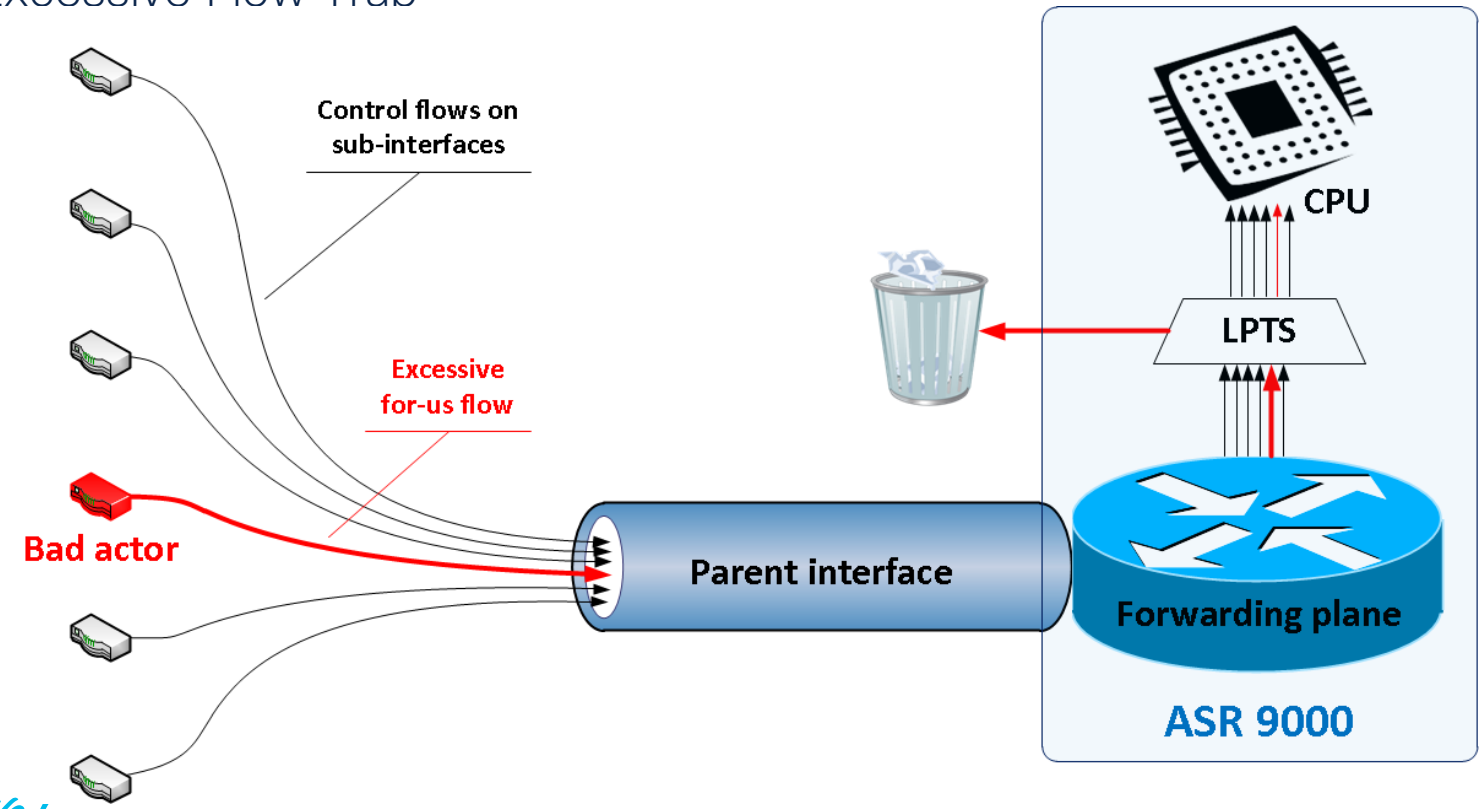
LPTS: PIFIB ACL-Based Policers

```
RP/0/RSP0/CPU0:rasr9000-2w-b#show lpts pifib hardware entry
type ipv4 start-index 12 num-entries 7 location 0/1/CPU0
Thu Apr  3 18:40:54.467 EDT
.
VRF ID              : 0x60000000
Destination IP      : any
Source IP           : 10.100.104.1
Is Fragment         : 0
Interface           : any
M/L/T/F             : 0/IPv4_STACK/0/BGP-known
DestNode            : 48
DestAddr            : 48
SID                 : 7
L4 Protocol         : TCP
TCP flag byte       : any
Source port         : Port:28603
Destination Port    : 179
Ct                  : 0x612050
Accepted/Dropped    : 5058/0
Lp/Sp               : 1/255
# of TCAM entries  : 1
HPo/HAr/HBu/Cir/acl:
14876914/33000pps/33000ms/33000pps/CORE
.
```

```
VRF ID              : 0x60000000
Destination IP      : any
Source IP           : 10.100.108.1
Is Fragment         : 0
Interface           : any
M/L/T/F             : 1/IPv4_LISTENER/0/LDP-UDP
DestNode            : FGID 3
DestAddr            : 3
SID                 : 7
L4 Protocol         : UDP
Source port         : Port:any
Destination Port    : 646
Ct                  : 0x612060
Accepted/Dropped    : 16214/0
Lp/Sp               : 1/255
# of TCAM entries  : 1
HPo/HAr/HBu/Cir/acl:
14876914/33000pps/33000ms/33000pps/CORE
State               : Entry in TCAM
Rsp/Rtp             : 16/30
-------------------------------------------------
--
```

Control Plane Protection

LPTS Excessive Flow Trap



Control Plane Protection

LPTS Excessive Flow Trap: Configuration

```
RP/0/RSP0/CPU0:rasr9000-2w-b#show running-config lpts punt excessive-flow-trap
Tue Mar 11 11:47:47.820 EDT
lpts punt excessive-flow-trap
 penalty-rate arp 50
 penalty-rate icmp 50
 penalty-rate igmp 100
 penalty-rate ip 100
 penalty-timeout arp 5
 penalty-timeout icmp 5
 penalty-timeout igmp 2
 penalty-timeout ip 4
 non-subscriber-interfaces mac
!
```

- Policing per Src MAC on main interfaces
- Policing per Src MAC on BNG sub-interfaces
- Policing per sub-int on non-BNG
  - Change with "`non-subscriber-interfaces mac`"

- Policing for-us from offending source instead of dropping for flow type from all peers
- Penalizing "bad actor" on major protocols: IP, IGMP, ICMP, ARP, DHCP, PPP, PPPoE, L2TP. Potentially impacting several protocols from offending peer
- Not enabled by default
- Check if default penalty rates and timeouts are acceptable in your case

Control Plane Protection

LPTS Excessive Flow Trap: Default & Configured Penalties

```
RP/0/RSP0/CPU0:rasr9000-2w-b#show lpts punt excessive-flow-trap information
.
              Police          Penalty
              Rate (pps)      Timeout (mins)
 Protocol    Default Config   Default Config   Punt Reasons
 --------    --------------   --------------   ----------------
 ARP            10      50        15      5     ARP
                                                Reverse ARP
                                                Dynamic ARP Inspection (DAI)

 ICMP           10      50        15      5     ICMP
.

 IGMP           10     100        15      2     IGMP
                                                IGMP Snoop
                                                MLD Snoop

 IPv4/v6        10     100        15      4     IP Subscriber (IPSUB)
                                                IPv4 options
                                                IPv4 FIB
                                                IPv4 TTL exceeded
                                                IPv4 fragmentation needed
                                                IPv4/v6 adjacency
                                                IPV4/v6 unknown IFIB
                                                UDP-known
 .
```

Control Plane Protection

LPTS Excessive Flow Example: A VRRP Flooding

```
LC/0/0/CPU0:Mar 11 12:52:09.059 : flowtrap[187]: %OS-FLOWTRAP-4-BAD_ACTOR_INTF_DETECTED : Excessive VRRP
flow detected on interface TenGigE0/0/0/5.511. The interface will be penalty-policed at 10 pps for 15
minutes.

RP/0/RSP0/CPU0:rasr9000-2w-b#show lpts pifib hardware police location 0/0/CPU0
.
FlowType               Policer Type    Cur. Rate  Def. Rate  Accepted             Dropped              TOS Value
---------------------- ------- ------- ---------- ---------- -------------------- -------------------- ----------
.
VRRP                   148     Static  1000       1000       804133               40681182             01234567
.

RP/0/RSP0/CPU0:rasr9000-2w-b#show lpts punt excessive-flow-trap all location 0/0/CPU0
Tue Mar 11 13:04:35.545 EDT
           Interface: TenGigE0/0/0/5.511
          Intf Handle: 0x04001740                          Location: 0/0/CPU0
             Protocol: ****                             Punt Reason: VRRP
        Penalty Rate: 10 pps                         Penalty Timeout: 15 mins
      Time Remaining: 8 mins 22 secs

LC/0/0/CPU0:Mar 11 13:24:33.899 : flowtrap[187]: %OS-FLOWTRAP-4-BAD_ACTOR_INTF_CLEARED : Interface
TenGigE0/0/0/5.511 cleared from penalty-policing by timeout.
```

Control Plane Protection

LPTS Excessive Flow Example: A VRRP Flooding

```
RP/0/RSP0/CPU0:rasr9000-2w-b#show vrrp
Tue Mar 11 13:07:23.623 EDT
IPv4 Virtual Routers:
                     A indicates IP address owner
                     | P indicates configured to preempt
                     | |
Interface   vrID Prio A P State    Master addr     VRouter addr
Te0/0/0/5.500  100  100   P Master   local           172.24.1.1
Te0/0/0/5.501  101   90   P Backup   172.24.1.2      172.24.1.1
Te0/0/0/5.502  102  100   P Master   local           172.24.1.1
Te0/0/0/5.503  103   90   P Backup   172.24.1.2      172.24.1.1
Te0/0/0/5.504  104  100   P Master   local           172.24.1.1
Te0/0/0/5.505  105   90   P Backup   172.24.1.2      172.24.1.1
Te0/0/0/5.506  106  100   P Master   local           172.24.1.1
Te0/0/0/5.507  107   90   P Backup   172.24.1.2      172.24.1.1
Te0/0/0/5.508  108  100   P Master   local           172.24.1.1
Te0/0/0/5.509  109   90   P Backup   172.24.1.2      172.24.1.1
Te0/0/0/5.510  110  100   P Master   local           172.24.1.1
Te0/0/0/5.511  111   90   P Master   local           172.24.1.1
Te0/0/0/5.512  112  100   P Master   local           172.24.1.1
Te0/0/0/5.513  113   90   P Backup   172.24.1.2      172.24.1.1
Te0/0/0/5.514  114  100   P Master   local           172.24.1.1
Te0/0/0/5.515  115   90   P Backup   172.24.1.2      172.24.1.1
Te0/0/0/5.516  116  100   P Master   local           172.24.1.1
Te0/0/0/5.517  117   90   P Backup   172.24.1.2      172.24.1.1
```

Control Plane Protection

LPTS exception punt

- Handles transit exceptions, some protocols, and snooping
  - Exceptions are transit that needs special processing [examples: MTU failure, TTL exhaustion]
  - Some protocols handled by LC CPU [BFD, ARP, CDP]
  - IGMP snooping
- Punted to LC CPU
  - Exception is IGMP snooping, punted to RSP CPU's
- Policers
  - Configured LC rate applied to LC, if not then a default rate applied
  - No global rate configuration option. But a pre-configuration per LC option
  - Enforced by each NP's microcode
  - More policers are added in newer releases

Control Plane Protection

Exception punt policers

```
RP/0/RSP0/CPU0:rasr9k-1y#show lpts pifib hardware static-police location 0/0/CPU0
Sun Dec  2 06:42:23.474 UTC
-------------------------------------------------------------
                Node 0/0/CPU0:
-------------------------------------------------------------
 Burst = 100ms for all flow types
-------------------------------------------------------------
Punt Reason              SID             Flow Rate  Burst Rate Accepted    Dropped    Destination
------------------------ --------------- ---------  ---------- ----------- ---------- -----------
PUNT_INVALID             NETIO_LOW       100        20         0           0          Local
.
PUNT_ARP                 ARP             1000       200        1888820     0          Local
.
PUNT_IGMP_SNOOP          NETIO_MED       4000       2000       0           0          0x0030 (0/RSP0/CPU0)
PUNT_MLD_SNOOP           NETIO_MED       4000       2000       0           0          0x0030 (0/RSP0/CPU0)
.
PUNT_IPv4_OPTIONS        NETIO_LOW       5000       1000       0           0          Local
.
PUNT_FOR_ICMP            NETIO_LOW       250        200        0           0          Local
PUNT_TTL_EXCEEDED        NETIO_LOW       2000       400        0           0          Local
PUNT_FRAG_NEEDED         NETIO_LOW       1000       400        0           0          Local
PUNT_PPPOE_FRAG_NEEDED   NETIO_LOW       1000       400        0           0          Local
PUNT_IPV4_BFD            BFD             12800      3500       0           0          Local
.
PUNT_IPV6_LINK_LOCAL     NETIO_HI        2000       2000       0           0          Local
PUNT_IPV6_SRC_LINK_LOCAL NETIO_HI        2000       2000       0           0          Local
PUNT_IPV6_HOP_BY_HOP     NETIO_LOW       5000       1000       2533        0          Local
PUNT_IPV6_TTL_ERROR      NETIO_LOW       2000       400        0           0          Local
.
```

Control Plane Protection

Customize punt policer rates

```
RP/0/RSP0/CPU0:rasr9k-1y(config)#lpts punt police location 0/0/CPU0
RP/0/RSP0/CPU0:rasr9k-1(config-punt-policer-per-node)# protocol cdp rate 50
RP/0/RSP0/CPU0:rasr9k-1(config-punt-policer-per-node)# protocol arp rate 5000
RP/0/RSP0/CPU0:rasr9k-1(config-punt-policer-per-node)# protocol ipv4 options rate 100
RP/0/RSP0/CPU0:rasr9k-1(config-punt-policer-per-node)# exception icmp rate 200
RP/0/RSP0/CPU0:rasr9k-1(config-punt-policer-per-node)# exception ipv4 ttl-error rate 500
.
RP/0/RSP0/CPU0:rasr9k-1y#show running-config lpts punt police location 0/0/CPU0
Sun Dec  2 07:05:30.358 UTC
lpts punt police location 0/0/CPU0
 exception invalid rate 400
 protocol cdp rate 50
 protocol arp rate 5000
 protocol ipv4 options rate 100
 exception icmp rate 200
 exception ipv4 ttl-error rate 500
 exception ipv4 fragment rate 10000
 exception adjacency rate 300
 exception acl-deny rate 50
 exception ipv6 ttl-error rate 500
 exception ipv6 fragment rate 10000
 exception mpls fragment rate 10000
 exception mpls ttl-error rate 500
!
```

Control Plane Protection

Customize punt policer rates – pre-configure

```
RP/0/RSP0/CPU0:rasr9k-1y(config)#lpts punt police location preconfigure 0/4/CPU0
RP/0/RSP0/CPU0:rasr9k-1(config-punt-policer-per-node)# protocol cdp rate 50
RP/0/RSP0/CPU0:rasr9k-1(config-punt-policer-per-node)# protocol arp rate 5000
RP/0/RSP0/CPU0:rasr9k-1(config-punt-policer-per-node)# protocol ipv4 options rate 100
RP/0/RSP0/CPU0:rasr9k-1(config-punt-policer-per-node)# exception icmp rate 200
RP/0/RSP0/CPU0:rasr9k-1(config-punt-policer-per-node)# exception ipv4 ttl-error rate 500
.
RP/0/RSP0/CPU0:rasr9k-1y#show running-config lpts punt police location 0/4/CPU0
Sun Dec  2 07:05:30.358 UTC
lpts punt police location 0/4/CPU0
 exception invalid rate 400
 protocol cdp rate 50
 protocol arp rate 5000
 protocol ipv4 options rate 100
 exception icmp rate 200
 exception ipv4 ttl-error rate 500
 exception ipv4 fragment rate 10000
 exception adjacency rate 300
 exception acl-deny rate 50
 exception ipv6 ttl-error rate 500
 exception ipv6 fragment rate 10000
 exception mpls fragment rate 10000
!
```

Control Plane Protection

Monitoring Hints

- A TCL script to periodically check and log excessive drops: https://supportforums.cisco.com/sites/default/files/legacy/1/5/2/116251-IOS-XR_LPTS_Alerting.tar.gz
  - lpts-threshold-alerting.tcl[65755]: LPTS threshold (80%) exceeded for flow type Raw-default on 0/2/0, 102.513333333% of 250 pps in last 60 seconds
- To clear punt/exception Accepted/Dropped counters:
  - #clear controller np counters all location …

Management

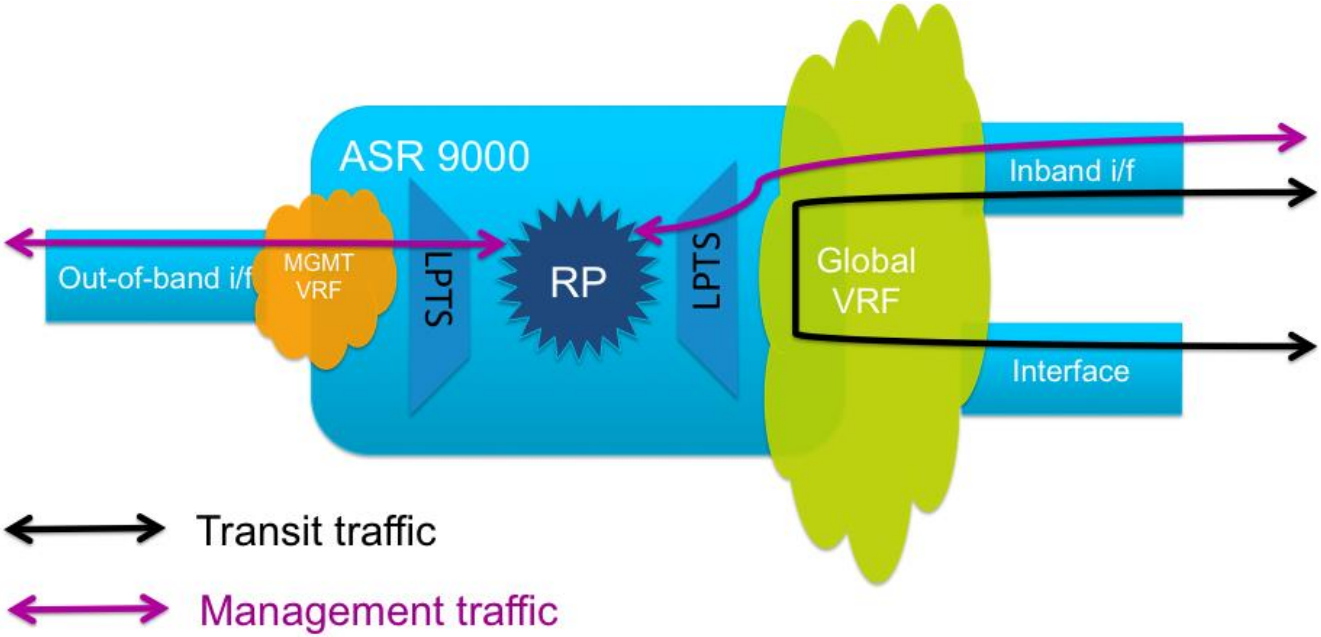
Management

Management Plane Protocols

- FTP
- NETCONF SSH
- NetFlow
- NTP
- RADIUS
- SCP
- SFTP
- SNMP
- SSH
- Syslog
- TACACS+
- Telnet
- Telemetry protocols
- TFTP

Management Traffic

Management interfaces: Out-of-band, in-band, and “global”



- No communication permitted between inband and out-of-band
- Management VRF is not necessary, but looks cleaner

Management Traffic

Out-of-Band: Virtual address, interfaces, and protocols

```
RP/0/RSP0/CPU0:rasr9k-1y#show running-config
.
vrf MGMT
 address-family ipv4 unicast
 !
 address-family ipv6 unicast
 !
!
rp mgmtethernet forwarding ! If LC<>Mgmt forwarding is needed
[not recommended]
!
ipv4 virtual address vrf MGMT 172.16.99.2/16
ipv4 virtual address use-as-src-addr
ipv6 virtual address vrf MGMT 2001:db8:1:ace:99::2/64
ipv6 virtual address use-as-src-addr
!
interface Loopback61 ! If needed
 vrf MGMT
 ipv4 address 172.16.100.100 255.255.255.255
 ipv6 address 2001:db8:1:ace:61::/128
!
interface MgmtEth0/RSP0/CPU0/0
 vrf MGMT
 ipv4 address 172.16.99.52 255.255.0.0
 ipv6 address 2001:db8:1:ace:99::52/64
!
interface MgmtEth0/RSP1/CPU0/0
 vrf MGMT
 ipv4 address 172.16.99.62 255.255.0.0
 ipv6 address 2001:db8:1:ace:99::62/64
!
```

```
!
key chain OSPF-MGMT
 key 1
  accept-lifetime 00:00:00 january 01 2022 23:59:59
december 31 2024
  key-string password 153B382537
  send-lifetime 00:00:00 january 01 2022 23:59:59
december 31 2024
  cryptographic-algorithm HMAC-MD5
 !
 accept-tolerance 90000
!
router ospf OSPF
 vrf MGMT
  router-id 172.24.100.100
  area 0
   authentication message-digest keychain OSPF-MGMT
   interface Loopback1
   !
   interface MgmtEth0/RSP0/CPU0/0
   !
   interface MgmtEth0/RSP1/CPU0/0
   !
  !
 !
```

Management Traffic

Out-of-Band: Virtual address, interfaces, and protocols

```
RP/0/RSP0/CPU0:rasr9k-1y#show running-config
control-plane management-plane out-of-band
Wed Dec  5 00:45:07.132 UTC
control-plane
 out-of-band
   vrf MGMT
   interface MgmtEth0/RSP0/CPU0/0
    allow SSH peer
     address ipv4 172.16.0.0/16
     address ipv6 2001:db8:1:ace::/64
    !
    allow NETCONF peer
     address ipv4 172.16.0.0/16
     address ipv6 2001:db8:1:ace::/64
    !
    .
   interface MgmtEth0/RSP1/CPU0/0
    allow SSH peer
     address ipv4 172.16.0.0/16
     address ipv6 2001:db8:1:ace::/64
    !
    allow NETCONF peer
     address ipv4 172.16.0.0/16
     address ipv6 2001:db8:1:ace::/64
    .
```

```
RP/0/RSP0/CPU0:rasr9k-1y#show mgmt-plane
Wed Dec  5 00:46:26.162 UTC

Management Plane Protection

.

outband interfaces
----------------------
interface - MgmtEth0/RSP0/CPU0/0
   ssh configured -
           peer v4 allowed - 172.16.0.0/16
           peer v6 allowed - 2001:db8:1:ace::/64
   netconf configured -
           peer v4 allowed - 172.16.0.0/16
           peer v6 allowed - 2001:db8:1:ace::/64
.
```

Management Traffic

In-band: If OOB is not available

```
RP/0/RSP0/CPU0:rasr9k-1y#show running-config
control-plane management-plane inband
Tue Dec 11 23:05:11.597 UTC
control-plane
 management-plane
  inband
   interface TenGigE0/0/0/2
    allow SSH peer
     address ipv4 192.168.1.0/24
     address ipv6 2001:db8:1:f192::/64
    !
   !
  !
 !
!
```

```
RP/0/RSP0/CPU0:rasr9k-1y#show mgmt-plane
Wed Dec  5 00:46:26.162 UTC

Management Plane Protection

inband interfaces
----------------------
interface - TenGigE0_0_0_2/
  ssh configured -
          peer v4 allowed - 192.168.1.0/24
          peer v6 allowed - 2001:db8:1:f192::/64

outband interfaces
----------------------
interface - MgmtEth0/RSP0/CPU0/0
  ssh configured -
          peer v4 allowed - 172.16.0.0/16
          peer v6 allowed - 2001:db8:1:ace::/64
  netconf configured -
          peer v4 allowed - 172.16.0.0/16
          peer v6 allowed - 2001:db8:1:ace::/64
.
```

Management Protection

Authentication, authorization, and accounting

```
tacacs source-interface Loopback1 vrf MGMT
tacacs-server host 172.16.1.98 port 49
 key 7 13061E010803
!
taskgroup operation
 task read bgp
 task read isis
 task write ospf
 inherit taskgroup operator
!
taskgroup provisioning
 inherit taskgroup netadmin
 description PROVISIONING GROUP
!
usergroup PROVISIONING
 taskgroup netadmin
 taskgroup provisioning
!
aaa authentication login default local

aaa accounting exec default start-stop group
tacacs+ none
aaa authorization exec default group tacacs+
local none
aaa authorization commands default group tacacs+
none
aaa authentication login default group tacacs+
local
```

```
RP/0/RSP0/CPU0:rasr9k-1y(config)#do show aaa
taskgroup
Wed Dec  5 01:40:50.022 UTC
Task group 'operation'
  Inherits from task group 'operator'

  Task IDs included directly by this group:
  Task:                  bgp  : READ
  Task:                 isis  : READ
  Task:                 ospf  : READ    WRITE

  Task group 'operation' has the following combined
set
  of task IDs (including all inherited groups):
  Task:       basic-services  : READ    WRITE
EXECUTE    DEBUG
  Task:                  bgp  : READ
  Task:                  cdp  : READ
  Task:                 diag  : READ
  Task:           ext-access  : READ
EXECUTE
  Task:                 isis  : READ
  Task:              logging  : READ
  Task:                 ospf  : READ    WRITE

Task group 'provisioning'
  Inherits from task group 'netadmin'
  .
```

More Control Plane Security

What's at stake?

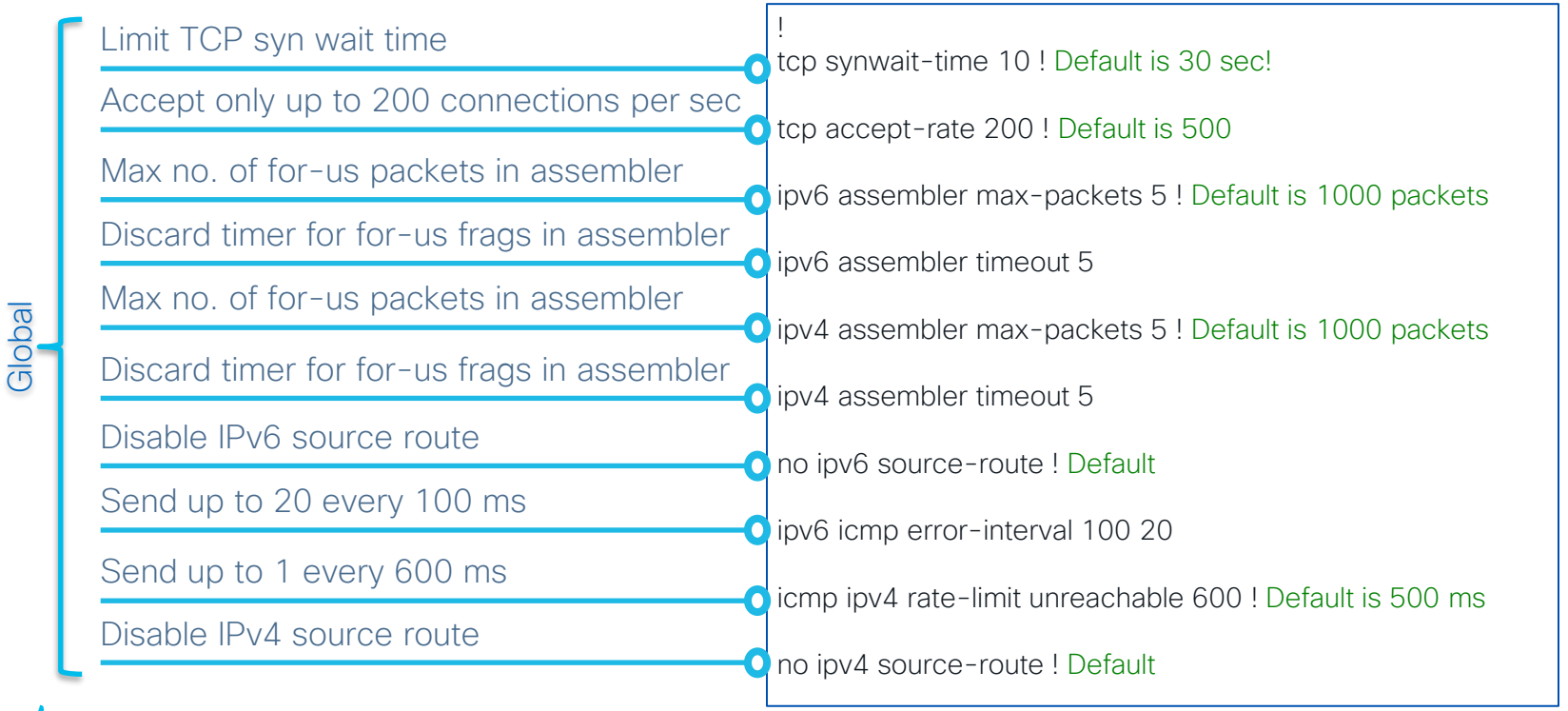
Security is CIA

- Confidentiality:
  - SP info
  - Customer info
  - Traffic
- Integrity:
  - SP traffic
  - Customer traffic
  - Control information
- Availability:
  - Operating environment
  - Operating parameters

- Threats are: environmental/natural/physical – human – technical

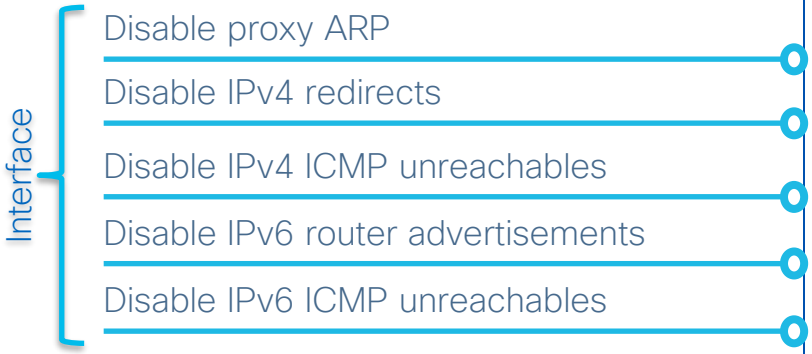
Control Plane Protection

Some Essential & Easy Protections



Control Plane Protection

Some Essential & Easy Protections




```
!
interface GigabitEthernet0/0/1/0
 no proxy-arp ! Default

 no ipv4 redirects ! Default

 ipv4 unreachables disable ! Breaks PMTUD!

 ipv6 nd suppress-ra ! RA only needed on host nets

 ipv6 unreachables disable
```

Control Plane Protection

Routing & MPLS Protocol Security

```
key chain ISIS
 key 1
  accept-lifetime 00:00:00 january 01 2012
23:59:59 december 31 2014
  key-string password 153B382537
  send-lifetime 00:00:00 january 01 2012
23:59:59 december 31 2014
  cryptographic-algorithm HMAC-MD5
 !
router isis ISIS
 is-type level-2-only
 net 49.6539.1291.6800.1240.00
 nsf ietf
 lsp-password keychain ISIS
 address-family ipv4 unicast
  metric-style wide
 !
interface TenGigE0/0/0/2
  circuit-type level-2-only
  hello-padding disable
  hello-password keychain ISIS
  address-family ipv4 unicast
   metric 20
```

```
router ospf OSPF
 vrf MGMT
  router-id 172.24.100.100
  area 0
   authentication message-digest keychain
OSPF-MGMT
   interface Loopback1
   !
   interface MgmtEth0/RSP0/CPU0/0
   !
   interface MgmtEth0/RSP1/CPU0/0
   !
  !
 !
!



rsvp
 interface TenGigE0/0/0/2
 !
 authentication
  key-source key-chain RSVP-KEY
```

```
router bgp 6539
 nsr
 bgp router-id 192.168.1.240
 .
 !
 neighbor 192.1.1.2
  remote-as 64000
  password encrypted 1511021F0725
  ttl-security
  address-family ipv4 unicast
   route-policy CUSTOMER-A in
   maximum-prefix 200 70
   remove-private-AS
  !
!
```

Control Plane Protection

VPLS Control Security: MAC LIMIT

- MAC learning limit per bridge domain & per AC
  - BD limit = OR > aggregate for AC's
  - Default is 4K, but no action
- "No-flood" to prevent learning & flooding
  - Traffic from known MAC sources continues to be forwarded/flooded. Unicast From unknown (new) is dropped
  - Protection to both Control and Data planes
- Notification both: Syslog and SNMP

```
RP/0/RSP0/CPU0:rasr9k-1y#show running-config l2vpn
Fri Dec  7 23:23:15.872 UTC
l2vpn
 bridge group BRIDGES
  bridge-domain DOMAIN-A
   mac
    limit
     maximum 2000
     action no-flood
     notification both
    !
    secure
     action none
     logging
    !
   !
   .
   !
   interface GigabitEthernet0/0/1/10.100
    mac
     limit
      maximum 1000
      action no-flood
      notification both
```

Control Plane Protection

VPLS Control Security: MAC LIMIT

Notification: Limit exceeded and action being taken

```
RP/0/RSP0/CPU0:Dec  9 05:34:45.987 : l2vpn_mgr[1126]: %L2-L2VPN-6-MAC_LIMIT_AC_SET :
Number of MAC addresses in AC 'Gi0/0/1/19.101' has reached the configured MAC limit maximum,
MAC learning and unicast flooding disabled
```

AC showing configured limit/action and status

```
RP/0/RSP0/CPU0:rasr9k-1y#show l2vpn bridge-domain interface GigabitEthernet 0/0/1/19.101
detail
Sun Dec  9 06:10:10.934 UTC
Legend: pp = Partially Programmed.
Bridge group: BRIDGES, bridge-domain: DOMAIN-A, id: 0, state: up, ShgId: 0, MSTi: 0

 ACs: 2 (2 up), VFIs: 1, PWs: 0 (0 up), PBBs: 0 (0 up)
  List of ACs:
    AC: GigabitEthernet0/0/1/19.101, state is up

      MAC learning: disabled (MAC-limit action)
      Flooding:
        Broadcast & Multicast: enabled
          Unknown unicast: disabled (MAC-limit action)
      MAC aging time: 300 s, Type: inactivity
      MAC limit: 1000, Action: limit, no-flood, Notification: syslog, trap
      MAC limit reached: yes
      MAC port down flush: enabled
```

Back within limit

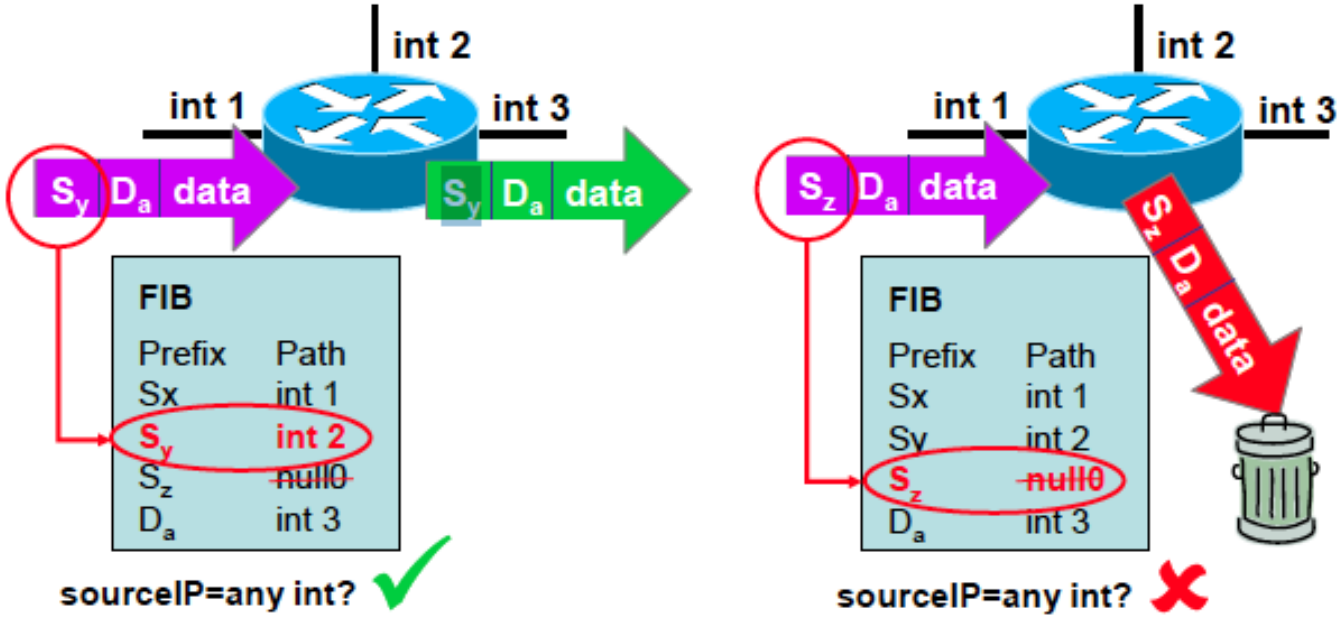
```
RP/0/RSP0/CPU0:Dec  9 08:00:17.577 : l2vpn_mgr[1126]: %L2-L2VPN-6-MAC_LIMIT_AC_CLEAR :
Number of MAC addresses in AC 'Gi0/0/1/19.101' has gone below 75 percent of the configured MAC
limit maximum, MAC learning and unicast flooding re-enabled
```

Securing the Data Plane

cisco Live!

Data Plane Protection

Reverse Path Forwarding Verification: uRPF Loose



- Defense against spoofed and bogus source packets.

Data Plane Protection

Reverse Path Forwarding Verification: uRPF Loose

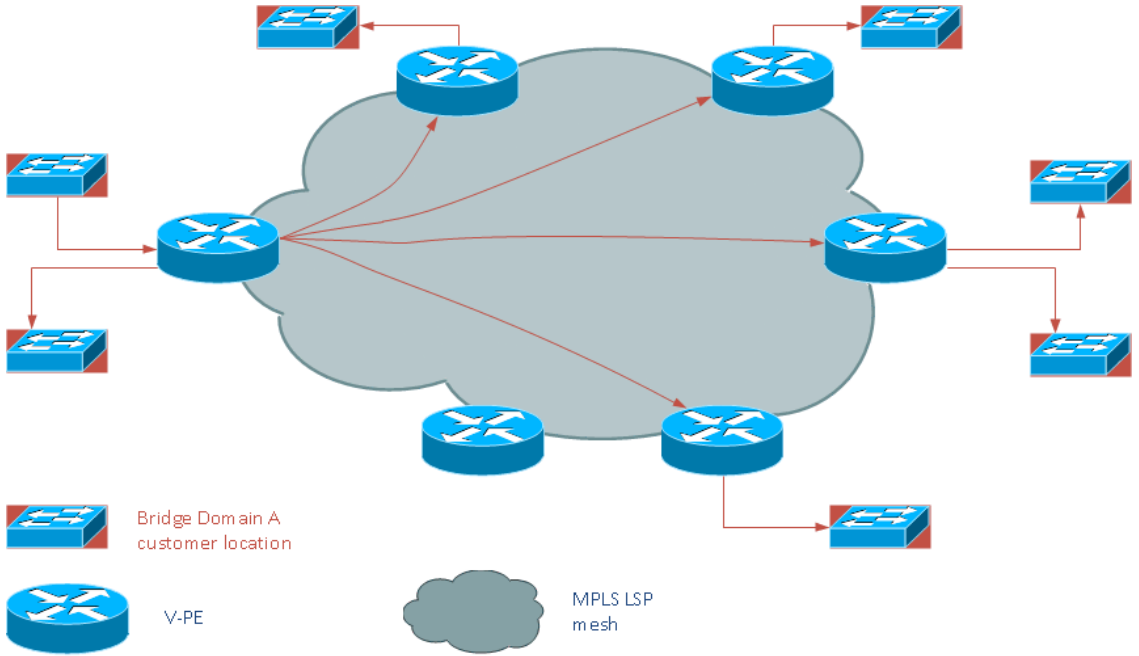
Routing to null0 (e.g. Bogons) results in source drops

Loose uRPF [via any] to account for asymmetric traffic and multi-homed customers

```
RP/0/RSP0/CPU0:rasr9k-1y#show running-config router static
Thu Dec  6 22:40:58.625 UTC
router static
 address-family ipv4 unicast
  10.0.0.0/8 Null0
  172.16.0.0/12 Null0
  192.168.0.0/16 Null0
!
 address-family ipv6 unicast
  2001:db8::/32 Null0
  fc00::/7 Null0
.
RP/0/RSP0/CPU0:rasr9k-1y#show running-config interface TenGigE 0/0/0/2
Thu Dec  6 22:30:17.910 UTC
interface TenGigE0/0/0/2
 ipv4 address 172.29.1.1 255.255.255.252
 ipv4 verify unicast source reachable-via any
 ipv6 verify unicast source reachable-via any
!
RP/0/RSP0/CPU0:rasr9k-1y#show ipv6 interface TenGigE 0/0/0/2
.
  Table Id is 0xe0800000
IP unicast RPF check is enabled
RPF mode loose
.
```

Data Plane Protection

VPLS Storm Control: Flooding is Costly



- Frame copies of unknown destination, multicast, and broadcast are flooded.

Data Plane Protection

VPLS Storm Control

- Sets the PPS limit per AC for forwarding/flooding ingress L2:
  - Broadcast [dest MAC FFFF.FFFF.FFFF]
  - Multicast [dest MAC DNDD.DDDD.DDDD where D=any, and N is odd]
  - Unknown unicast [dest unicast MAC is not in MAC table]
- Does not impact unicast with dest MAC in cache/table

```
RP/0/RSP0/CPU0:rasr9k-1y#show running-config l2vpn
Mon Dec 10 08:10:39.712 UTC
l2vpn
 bridge group BRIDGES
  bridge-domain DOMAIN-A
   interface GigabitEthernet0/0/1/10.100
   !
    storm-control unknown-unicast pps 2000
    storm-control multicast pps 6000
    storm-control broadcast pps 5000
   !
   interface GigabitEthernet0/0/1/19.101
   !
    storm-control unknown-unicast pps 2000
    storm-control multicast pps 6000
    storm-control broadcast pps 5000
   !
   vfi VFI-A
    vpn-id 65000
    autodiscovery bgp
     rd auto
     route-target 65000:1
```

Data Plane Protection

VPLS Storm Control

- For Data Centers: broadcast and unknown unicast rate is influenced by ARP and MAC aging on hosts and network devices
- To clear counters: “clear l2vpn forwarding counters”

```
RP/0/RSP0/CPU0:rasr9k-1y#show l2vpn bridge-domain interface GigabitEthernet
0/0/1/19.101 detail
Mon Dec 10 07:15:42.544 UTC
Legend: pp = Partially Programmed.
Bridge group: BRIDGES, bridge-domain: DOMAIN-A, id: 0, state: up, ShgId: 0, MSTi: 0

  List of ACs:
    AC: GigabitEthernet0/0/1/19.101, state is up

      MAC aging time: 300 s, Type: inactivity

      Storm Control:
       Broadcast: enabled(5000)
       Multicast: enabled(6000)
       Unknown unicast: enabled(2000)
      Static MAC addresses:
      Statistics:
        packets: received 4303565, sent 0
        bytes: received 258213900, sent 0
      Storm control drop counters:
        packets: broadcast 0, multicast 0, unknown unicast 108388233
        bytes: broadcast 0, multicast 0, unknown unicast 6503293980
      Dynamic ARP inspection drop counters:
        packets: 0, bytes: 0
```

Data Plane Protection

VPLS ARP Inspection

- Perform ARP/RARP checks:
  - All: Sender_MAC == Source_MAC
  - Replies: Target_MAC == Dest_MAC
  - ARP request source IPv4 is unicast
  - ARP reply dest IPv4 is unicast
- Configure at the BD level [Applies to all AC's] or the AC level

```
RP/0/RSP0/CPU0:rasr9k-1y#show running-config l2vpn
Tue Dec 11 03:50:31.612 UTC
l2vpn
 bridge group BRIDGES
  bridge-domain DOMAIN-A
  .
   dynamic-arp-inspection
    logging
    address-validation
     src-mac
     dst-mac
     ipv4
  !
  !
   interface GigabitEthernet0/0/1/10.100
   .
    dynamic-arp-inspection
     logging
     address-validation
      src-mac
      dst-mac
      ipv4
   !
   !
    storm-control unknown-unicast pps 2000
```

Data Plane Protection

VPLS ARP Inspection

- Violations are logged and dropped

```
RP/0/RSP0/CPU0:rasr9k-1y#show l2vpn forwarding interface GigabitEthernet 0/0/1/10.100 detail
location 0/0/CPU0
Tue Dec 11 06:29:08.357 UTC
Local interface: GigabitEthernet0/0/1/10.100, Xconnect id: 0x40001, Status: up
  .
    Flooding:
      Broadcast & Multicast: enabled
      Unknown unicast: enabled
    MAC aging time: 300 s, Type: inactivity
    MAC limit: 1000, Action: limit, no flood, Notification: syslog, trap
    MAC limit reached: no
    MAC Secure: enabled, Logging: disabled, Action: restrict
    DHCPv4 snooping: profile not known on this node, disabled
    Dynamic ARP Inspection: enabled, Logging: enabled
    Dynamic ARP Inspection Address Validation:
      IPv4 verification: enabled
      Source MAC verification: enabled
      Destination MAC verification: enabled
    IP Source Guard: disabled, Logging: disabled
    IGMP snooping profile: profile not known on this node
```

```
LC/0/0/CPU0:Jun 16 13:28:28.697 : l2fib[188]: %L2-L2FIB-5-SECURITY_DAI_VIOLATION_AC :
Dynamic ARP inspection in AC GigabitEthernet0_0_0_7.1000 detected violated packet - source MAC:
0000.0000.0065, destination MAC: 0000.0040.0000, sender MAC: 0000.0000.0064, target MAC:
0000.0000.0000, sender IP: 5.6.6.6, target IP: 130.10.3.2
.
```

Data Plane Protection

VPLS MAC Security

- If a cached MAC appears as frame source on another AC:
  - Log
  - Do not learn MAC
  - Drop frame
- A bridge domain level options is to shut down “offending” AC

```
RP/0/RSP0/CPU0:rasr9k-1y#show running-config l2vpn
Fri Dec 14 02:52:41.373 UTC
l2vpn
 bridge group BRIDGES
  bridge-domain DOMAIN-A
   mac
    limit
     .
    !
    secure
     action none
     logging
    !
   .
   !
   interface GigabitEthernet0/0/1/0.200
    mac
     limit
      maximum 1000
      action no-flood
      notification both
     !
     secure
      action restrict
      logging
     !
```

Data Plane Protection

VPLS MAC Security

To see the MAC table

```
RP/0/RSP0/CPU0:rasr9k-1y#show l2vpn forwarding bridge-domain mac-address location 0/0/CPU0
Fri Dec 14 02:48:57.535 UTC
 To Resynchronize MAC table from the Network Processors, use the command...
    l2vpn resynchronize forwarding mac-address-table location <r/s/i>

Mac Address    Type    Learned from/Filtered on    LC learned Resync Age        Mapped to
--------------------------------------------------------------------------------
0000.c001.0102 dynamic Gi0/0/1/19.101             0/0/CPU0   0d 0h 0m 15s      N/A
0000.c001.0103 dynamic Gi0/0/1/19.101             0/0/CPU0   0d 0h 0m 0s       N/A
.
0000.c001.015f dynamic Gi0/0/1/19.101             0/0/CPU0   0d 0h 0m 3s       N/A
0000.c001.0160 dynamic Gi0/0/1/19.101             0/0/CPU0   0d 0h 0m 3s       N/A
0000.c001.0161 dynamic Gi0/0/1/19.101             0/0/CPU0   0d 0h 0m 13s      N/A
.
0000.c001.0164 dynamic Gi0/0/1/0.200              0/0/CPU0   0d 0h 0m 20s      N/A
0000.c001.0166 dynamic Gi0/0/1/0.200              0/0/CPU0   0d 0h 0m 10s      N/A
0000.c001.0167 dynamic Gi0/0/1/0.200              0/0/CPU0   0d 0h 0m 17s      N/A
0000.c001.0168 dynamic Gi0/0/1/0.200              0/0/CPU0   0d 0h 0m 15s      N/A
0000.c001.0169 dynamic Gi0/0/1/0.200              0/0/CPU0   0d 0h 0m 0s       N/A
```

Violation detected, action taken

```
LC/0/0/CPU0:Dec 14 03:00:12.138 : l2fib[249]: %L2-L2FIB-5-
SECURITY_MAC_SECURE_VIOLATION_AC : MAC secure in AC GigabitEthernet0_0_1_0.200
detected violated packet - source MAC: 0000.c001.0160, destination MAC:
6c9c.ed2b.57dc; action: restrict
```

Data Plane Protection

Other Protections to Consider

- Routing:
  - Perimeter ACL [Also protects control and management planes]
  - Policing and admission control
  - No default route
  - Routing to Null0
  - BGP source/destination RTBH, and sinkholing
  - CSC label security
- Ethernet services:
  - DHCP snooping
  - IP source guard

Agenda

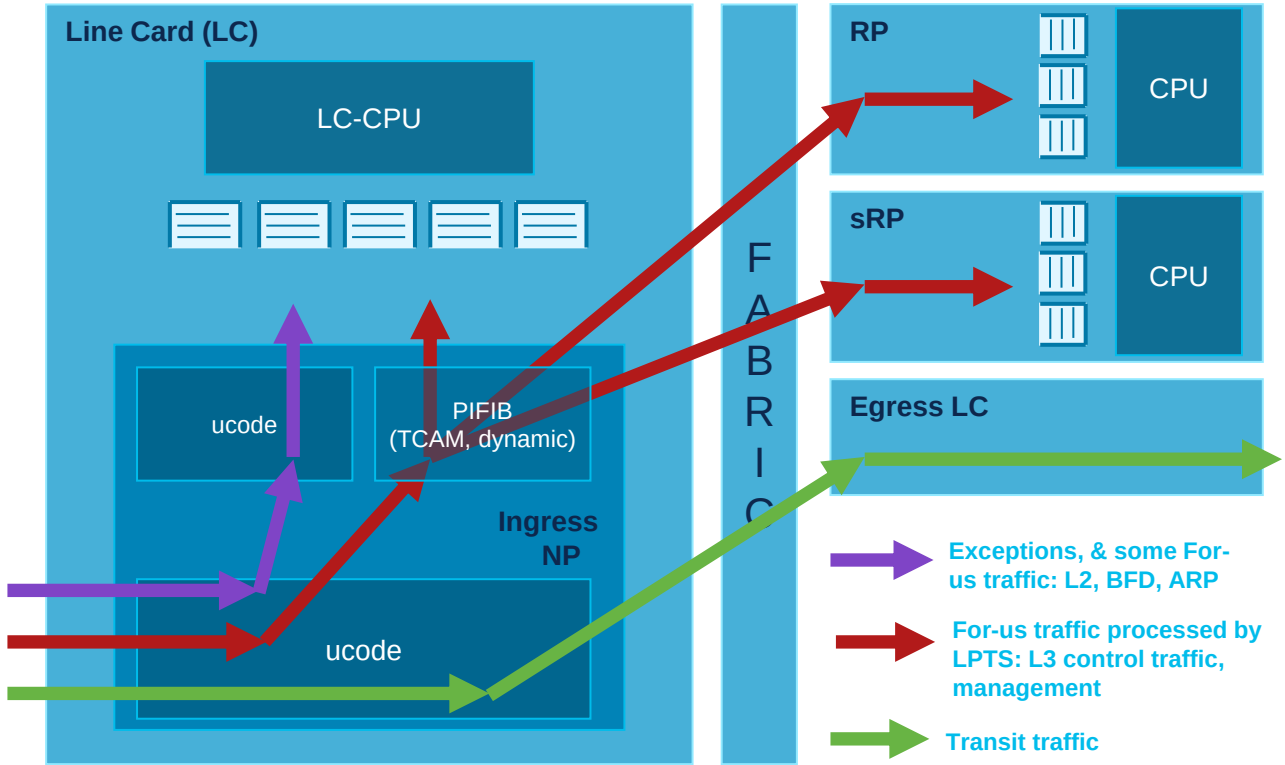
- ✓ **System Architecture**: System anatomy & health
- ✓ **Operating System & Configuration: IOS-XR & configuration models**
- ✓ **Control, Management, Security: Processing of control & exceptions**
- ➢ **Transit Packet/Frame Journey**: Life of L3/L2 unicast/multicast
- ➢ **MPLS Operation**: Processing, forwarding & L3/L2 service operation
- ➢ **Troubleshooting**: Diagnostics, counters, drops, and packet capture

4 Transit Packet/Frame Journey

Traffic: Transit, For us, and Exceptions

Differentiate on ingress NP

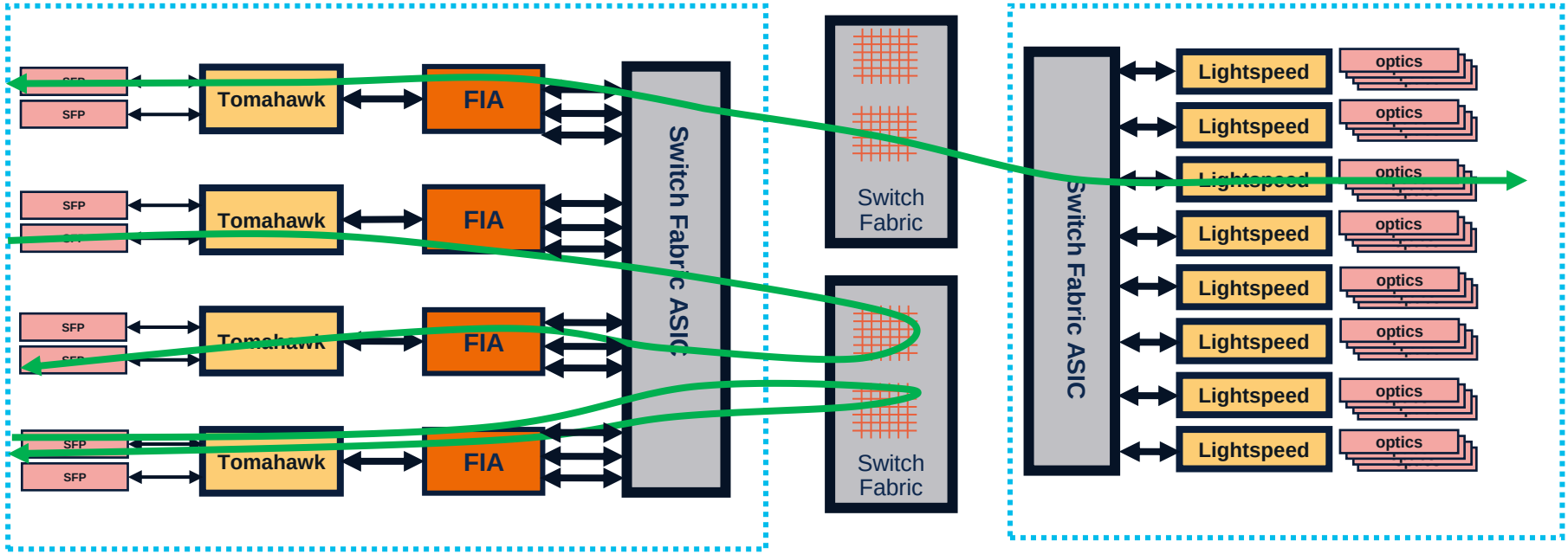
- **Transit**
  - Look up, re-write, forward
- **For us**
  - Destined to RP, or link local scope
  - Punt to RP or ingress LC CPU
- **Exception**
  - MTU failure, TTL failure, etc. Should have been transit
  - Punt to LC CPU



L3 Unicast Packet Forwarding

Unicast Transit Frame Path

Physical > NP > FIA > Fabric > [FIA > NP] > Physical



A9K-8X100GE

A99-32X100GE

Unicast Transit Frame Path

Forwarding

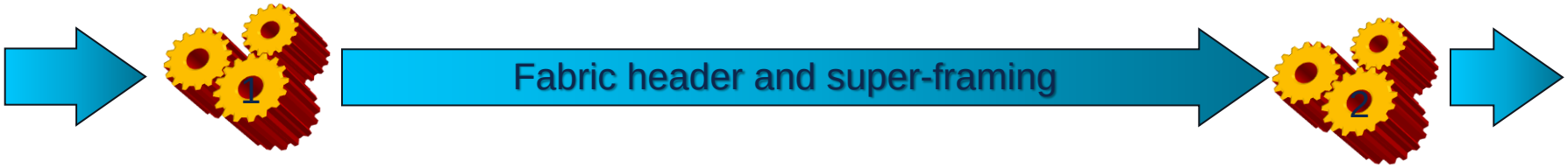
- All frames take same path stages
  - Ingress physical => ingress NP => FIA => fabric => FIA => egress NP => egress physical
  - Super-framing in fabric, and per super-frame load sharing
- Two stage forwarding
  - Ingress NP: to which egress port, ingress encap (if tunneling) and ingress features
  - Egress NP: Adjacency, encap, and egress features
- VOQ, and back-pressure signaling
  - Each FIA has a VQI per each egress port of 10 Gbps or higher
  - 3 priorities (VoQ) per VQI (imposed by ingress QoS)
  - Back pressure is signaled backwards from egress NP to ingress FIA for buffering

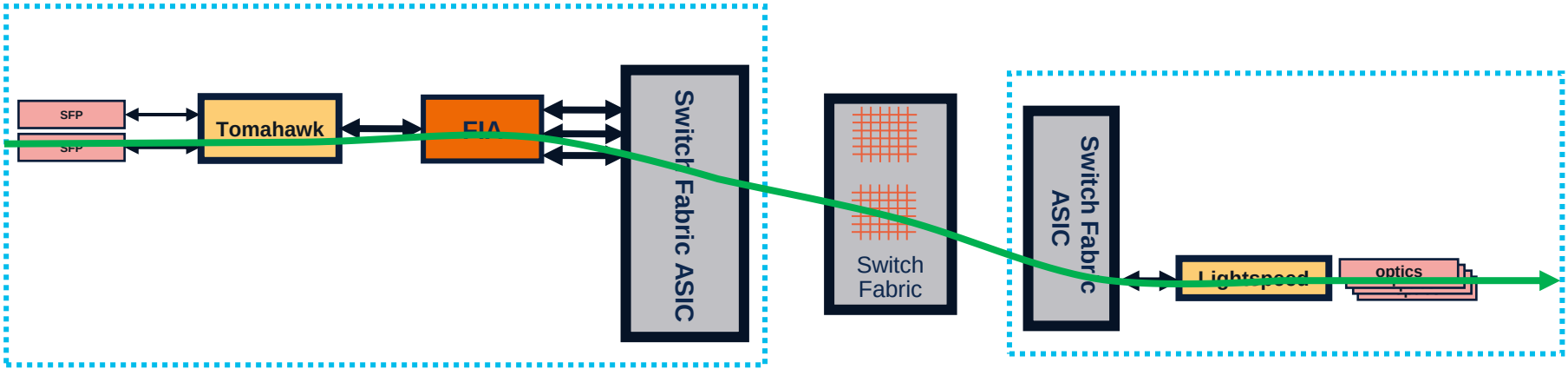
Unicast Two Stage Forwarding

By ingress NP and egress NP

First stage: lookup on **ingress NPU** →
Egress NPU (or **SFP**: switch fabric port)

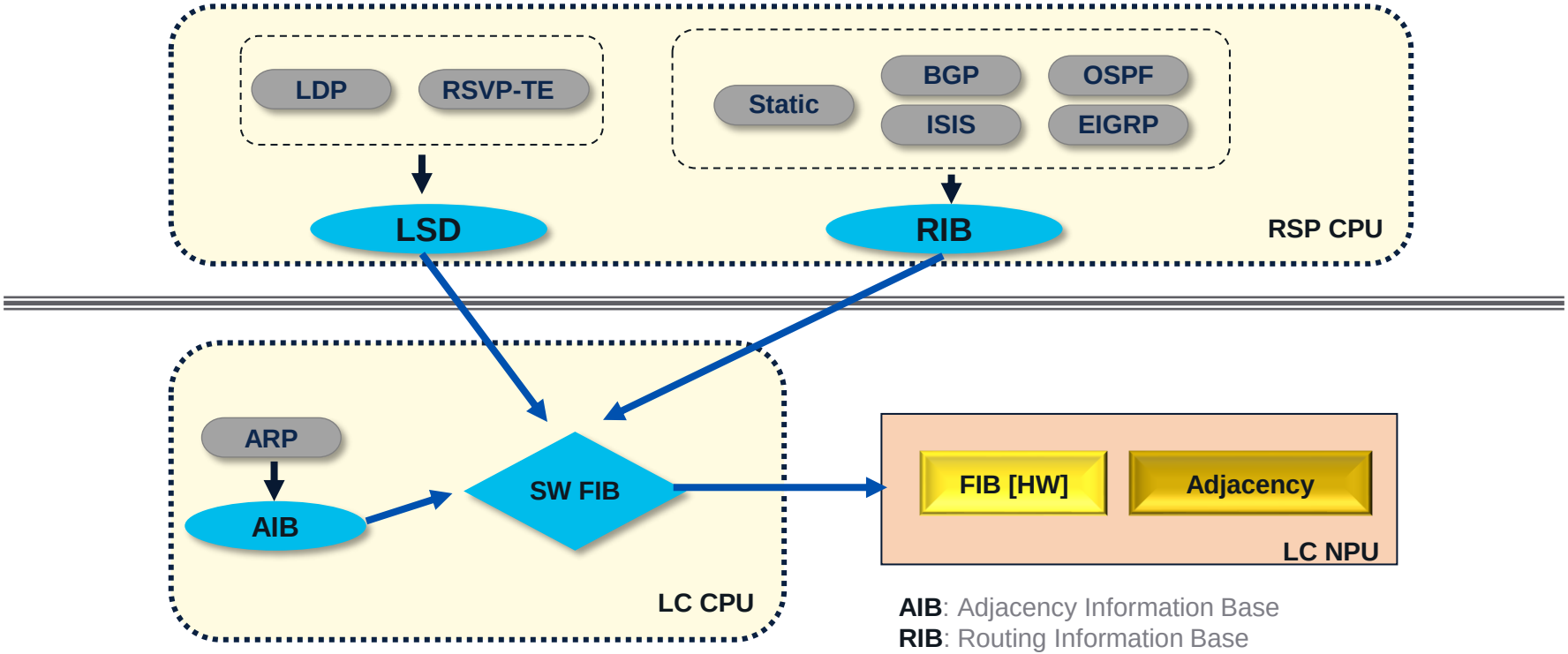
Second stage: lookup on **egress NPU** →
Egress port and rewrite information





The NP FIB

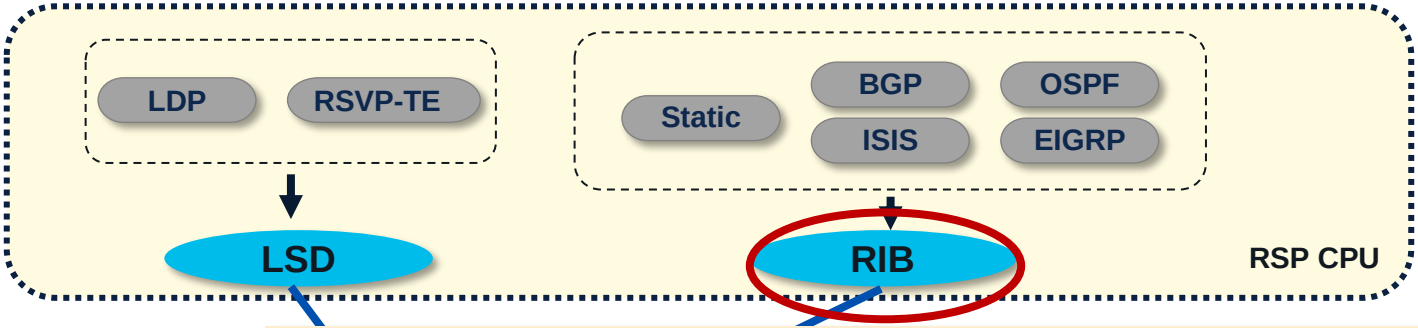
From RP control plane to data plane NP



AIB: Adjacency Information Base
RIB: Routing Information Base
FIB: Forwarding Information Base
LSD: Label Switch Database

The NP FIB

RIB info: example

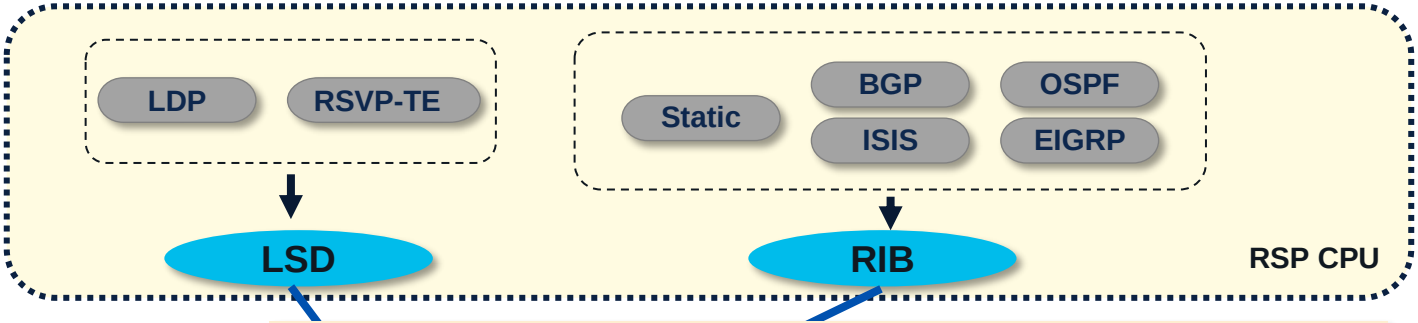


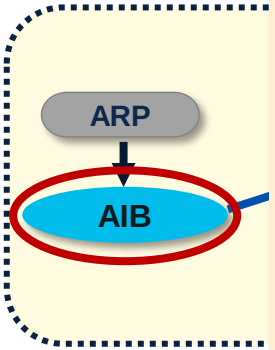

```
RP/0/RSP0/CPU0:asr#show route ipv6 2001:db8:1:c06::
Thu Jan  5 10:05:24.946 EST

Routing entry for 2001:db8:1:c06::/128
  Known via "isis 27064", distance 115, metric 20000, type level-2
  Installed Dec 28 14:11:48.864 for 1w0d
  Routing Descriptor Blocks
    fe80::12f3:11ff:fe2e:553c, from 2001:db8:1:c06::,
          via GigabitEthernet0/0/0/19
      Route metric is 20000
  No advertising protos.
```

RIB: Routing Information Base
FIB: Forwarding Information Base
LSD: Label Switch Database

The NP FIB

Line card adjacency

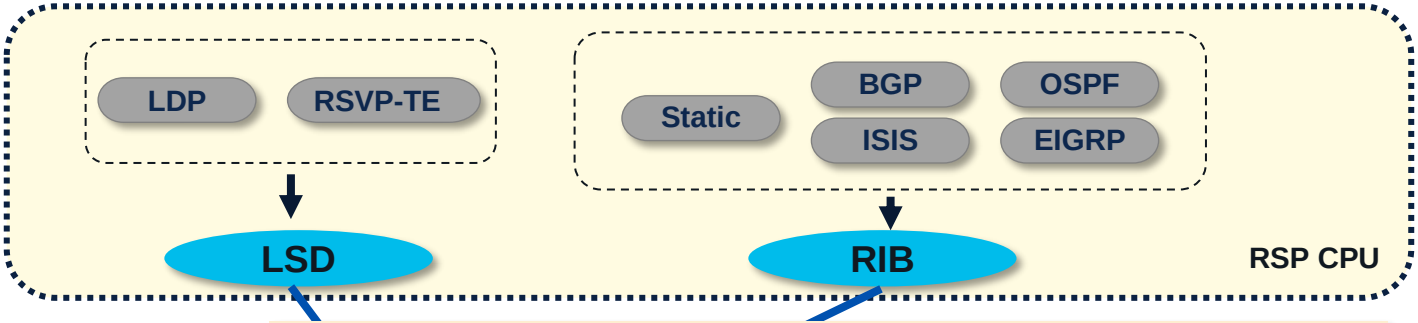


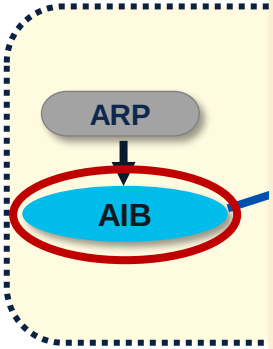



```
RP/0/RSP0/CPU0:asr#show adjacency summary location 0/0/CPU0
.
Adjacency table (version 93) has 41 adjacencies:
     17 complete adjacencies
      0 incomplete adjacencies
     24 interface adjacencies
      0 deleted adjacencies in quarantine list
      4 adjacencies of type IPv4
        .
      2 adjacencies of type IPv6
          2 complete adjacencies of type IPv6
        .
          1 multicast adjacency of type IPv6
     11 adjacencies of type MPLS
        .
```

The NP FIB

Line card adjacency



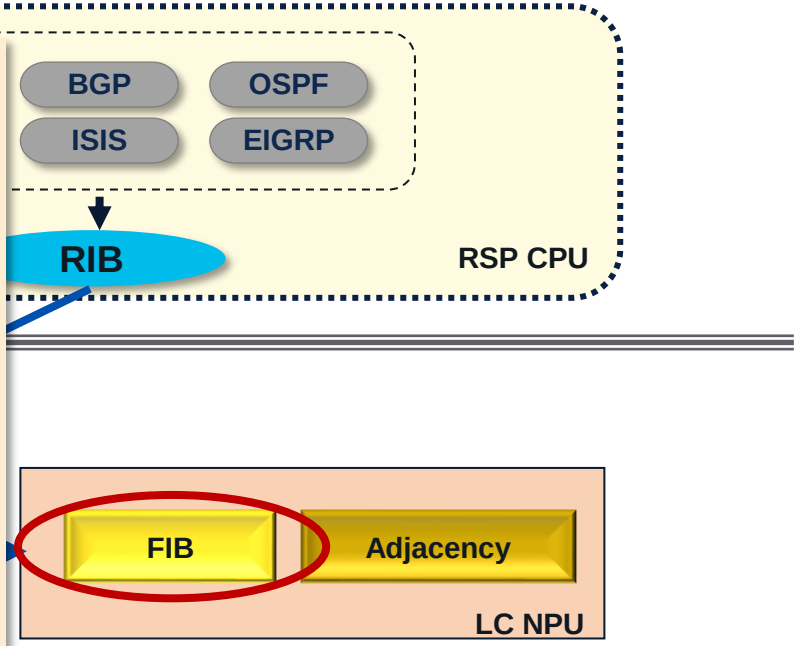



```
RP/0/RSP0/CPU0:asr#show adjacency GigabitEthernet 0/0/0/19
detail location 0/0/CPU0
.
Gi0/0/0/19                10.20.93.2                    44        2(    0) ipv4
                          10f3112e553c10f311366a190800
                          mtu: 4470, flags 1 0
                          0 packets, 0 bytes
Gi0/0/0/19                                              93        2(    0) mpls
                          fe80::12f3:11ff:fe2e:553c
                          10f3112e553c10f311366a198847
                          mtu: 4470, flags 1 0
                          0 packets, 0 bytes
Gi0/0/0/19                                              92        2(    0) ipv6
                          fe80::12f3:11ff:fe2e:553c
                          10f3112e553c10f311366a1986dd
                          mtu: 4470, flags 1 0
                          3690 packets, 265680 bytes
```

The NP FIB

FIB entry in NP: example

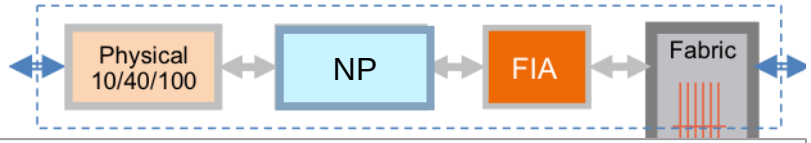
```
RP/0/RSP0/CPU0:asr#show cef ipv6 2001:db8:1:c06:: hardware ingress
loc 0/0/CPU0
.
   via fe80::12f3:11ff:fe2e:553c/128, GigabitEthernet0/0/0/19, 5
dependencies, weight 0, class 0 [flags 0x0]
    .
    next hop fe80::12f3:11ff:fe2e:553c/128
    .
TX H/W Result for NP:0 (index: 0x684c (BE)):
.
    uidb_index      : 0x1600 (LE)
    l3_mtu          : 4470
    adj_stats_index : 0x100661
    dest_mac        : 0x10f3.112e.553c
.
RX H/W Result on NP:0 [Adj ptr:0x18 (BE)]:
.
                rx_drop: 0

    FLAGS
                gre_adj: 0            pwhe_adj: 0
gre_ipv6_transp: 0
    sfp/vqi         : 0x5c
    egress np port  : 0x1d00
    if_handle       : 0x4000580
```



AIB: Adjacency Information Base
RIB: Routing Information Base
FIB: Forwarding Information Base
LSD: Label Switch Database

L3 Unicast Packet Journey

Mapping the port to NP and FIA

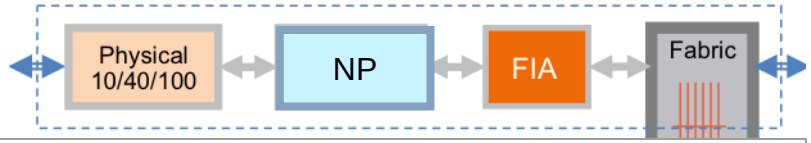



```
! Example: Path from GigabitEthernet0/0/1/0 192.3.1.2 TO TenGigE0/4/0/20.6 192.6.1.2

RP/0/RSP0/CPU0:rasr9k-1y#show controllers NP ports all location 0/0/CPU0
Fri Feb 22 15:57:32.307 UTC
                  Node: 0/0/CPU0:
----------------------------------------------------------------
NP Bridge Fia                          Ports
-- ------ --- --------------------------------------------------
0  --     0   TenGigE0/0/0/0, TenGigE0/0/0/1, TenGigE0/0/0/2, TenGigE0/0/0/3
1  --     1   GigabitEthernet0/0/1/0 - GigabitEthernet0/0/1/19

RP/0/RSP0/CPU0:rasr9k-1y#show controllers NP ports all location 0/4/CPU0
Fri Feb 22 15:55:22.370 UTC
                  Node: 0/4/CPU0:
----------------------------------------------------------------
NP Bridge Fia                          Ports
-- ------ --- --------------------------------------------------
0  --     0   TenGigE0/4/0/0, TenGigE0/4/0/1, TenGigE0/4/0/2
1  --     0   TenGigE0/4/0/3, TenGigE0/4/0/4, TenGigE0/4/0/5
2  --     1   TenGigE0/4/0/6, TenGigE0/4/0/7, TenGigE0/4/0/8
3  --     1   TenGigE0/4/0/9, TenGigE0/4/0/10, TenGigE0/4/0/11
4  --     2   TenGigE0/4/0/12, TenGigE0/4/0/13, TenGigE0/4/0/14
5  --     2   TenGigE0/4/0/15, TenGigE0/4/0/16, TenGigE0/4/0/17
6  --     3   TenGigE0/4/0/18, TenGigE0/4/0/19, TenGigE0/4/0/20
7  --     3   TenGigE0/4/0/21, TenGigE0/4/0/22, TenGigE0/4/0/23
```

Map the port to NP and FIA

L3 Unicast Packet Journey

The egress interface identifier




```
RP/0/RSP0/CPU0:rasr9k-1y#show controllers pm interface tenGigE 0/4/0/20.6
Fri Feb 22 16:45:22.404 UTC
Ifname(1): TenGigE0_4_0_20.6, ifh: 0xc001340 :
iftype                0x19
egress_uidb_index     0x1d
ingress_uidb_index    0x1d
port_num              0x14
subslot_num           0x0
phy_port_num          0x14
channel_id            0x6
channel_map           0x0
lag_id                0x0
virtual_port_id       0x0
switch_fabric_port    0x136
in_tm_qid_fid0        0x0
in_tm_qid_fid1        0x0
in_qos_drop_base      0x0
out_tm_qid_fid0       0x0          0x0          0x0          0x0          0x0          0x0          0x0
0x0
out_tm_qid_fid1       0x0          0x0          0x0          0x0          0x0          0x0          0x0
0x0
out_qos_drop_base     0x0
bandwidth             10000000 kbps
.
```

Get internal identifiers

Interface handle: unique to logical sub-interface

Fabric port, shared between all port sub-interfaces. The fabric destination.

L3 Unicast Packet Journey

Ingress NP FIB

```
RP/0/RSP0/CPU0:rasr9k-1y#show cef ipv4 192.6.1.2
hardware ingress location 0/0/CPU0
Fri Feb 22 17:40:35.887 UTC
192.6.1.0/24, version 364, attached, connected,
internal 0xc0000c1 (ptr 0x8856b534) [1], 0x0
(0x873dde50), 0x0 (0x0)
 Updated Feb 22 16:09:42.862
 remote adjacency to TenGigE0/4/0/20.6
 Prefix Len 24, traffic index 0, precedence
routine (0), priority 0
   via TenGigE0/4/0/20.6, 2 dependencies, weight
0, class 0 [flags 0x8]
    path-idx 0 [0x8a60a7bc 0x0]
    remote adjacency
LEAF - HAL pd context :
 sub-type : IPV4, ecd_marked:0,
has_collapsed_ldi:0, collapse_bwalk_required:0,
ecdv2_marked:0
Leaf H/W Result:

    Physical Result: 0x11dd0600 (LE)

    Raw Data0: 0x91ad1000 00000001 360c0013
40000000
    Raw Data1: 0x00000000 00000136 00180000
00000000
    .
```

```
RX H/W Result on NP:1 [Adj ptr:0x3a (BE)]:

    Raw Data0: 0x91000000 00000136 0c001340 00000000
    adj_resolve_control_byte0
                        match: 1
valid: 1
                    iptunl_adj: 0
                   remote_rack: 0

    adj_resolve_control_byte1
                     adj_down: 0
mgscp_en: 0
      rx_lag_hash_en: 0
          rx_lag_adj: 0

    adj_resolve_control_byte2
                    rx_lag_adj: 0
rx_adj_null0: 0
                   rp_destined: 0
rx_punt: 0
                       rx_drop: 0
    sfp/vqi          : 0x136
    if_handle        : 0xc001340
.
```

L3 Unicast Packet Journey

Egress NP FIB

```
RP/0/RSP0/CPU0:rasr9k-1y#show cef ipv4 192.6.1.2
hardware egress location 0/4/CPU0
Fri Feb 22 17:55:28.494 UTC
192.6.1.2/32, version 0, internal 0x4080001 (ptr
0x8efc2704) [1], 0x0 (0x8e0f2210), 0x0 (0x0)
 Updated Feb 22 16:13:35.351
 local adjacency 192.6.1.2
 Prefix Len 32, traffic index 0, Adjacency-
prefix, precedence routine (0), priority 0
   via 192.6.1.2, TenGigE0/4/0/20.6, 3
dependencies, weight 0, class 0 [flags 0x0]
    path-idx 0 [0x91a2cef8 0x0]
    next hop 192.6.1.2
    local adjacency
LEAF - HAL pd context :
 sub-type : IPV4, ecd_marked:0,
has_collapsed_ldi:0, collapse_bwalk_required:0,
ecdv2_marked:0
Leaf H/W Result:

    Physical Result: 0x11e80300 (LE)

    Raw Data0: 0x91ad1000 8a030001 360c0013
40400000
   .
```

```
TX H/W Result for NP:6 (index: 0x38a (BE)):
    Raw Data0: 0x91080000 1d000000 dc050000 400b5f00
    Raw Data1: 0x0000c006 01020000 00000000 00000000
    adj_resolve_control_byte0
                      reserved: 0
egr_uidb_internal: 1
                         match: 1
valid: 1
                    iptunl_adj: 0
    adj_resolve_control_byte1
                  tx_adj_null0: 0
tx_punt: 0                    tx_drop: 0
               default_action: 1
spare: 0
    adj_resolve_control_byte2
                         spare: 0
spare_cb: 0
    flags
        gre_adj          : 0
    uidb_index       : 0x1d00 (LE)
    reserve_pad_word: 0
    l3_mtu           : 1500
    reserve_pad_1    : 0
    adj_stats_index : 0x400b5f00
    dest_mac         : 0x0000.c006.0102
    ether reserved   : 0000000000000000
.
```

L3 Unicast Packet Journey

GRE encap case

```
RP/0/RSP0/CPU0:rasr9k-1y#show cef vrf DEF ipv4
172.25.25.2 hardware ingress location 0/4/CPU0
Sat Feb 23 14:35:00.017 UTC
172.25.25.0/24, version 1, attached, connected,
internal 0xc0000c1 (ptr 0x8e154de4) [1], 0x0
(0x8e0ec7c0), 0x0 (0x0)
 Updated Feb 21 16:28:04.573
 local adjacency point2point
 Prefix Len 24, traffic index 0, precedence
routine (0), priority 0
   via tunnel-ip25, 3 dependencies, weight 0,
class 0 [flags 0x8]
    path-idx 0 [0x90fdd3b4 0x0]
    local adjacency
LEAF - HAL pd context :
 sub-type : IPV4, ecd_marked:0,
has_collapsed_ldi:0, collapse_bwalk_required:0,
ecdv2_marked:0
Leaf H/W Result:

    Physical Result: 0x11be0200 (LE)

    Raw Data0: 0x11a50000 c9020000 00000000 00000000
    Raw Data1: 0x00000000 00000000 00180000 0000a2ff
    leaf_resolve_control_byte0
.
```

```
TX H/W Result for NP:6 (index: 0x28a (BE)):
.
    adj_resolve_control_byte0
                    reserved: 0
egr_uidb_internal: 1
                       match: 1
valid: 1
                  iptunl_adj: 1
.
    flags
      gre_adj      : 1
  uidb_index       : 0x1b00 (LE)
  reserve_pad_word: 0
  l3_mtu           : 1476
  reserve_pad_1    : 0
  adj_stats_index : 0x18005f00
  GRE Adj
    ip_src         : 172.20.20.1 ip_dst     : 172.20.20.2
    tos            : 0           ttl          : 0xff
    df             : 1           tos_reflect  : 1
    rsvd flag bits: 0            encap_checksum: 0x40a3
    vrf_id         : 0           reserved     : 0
.
```

L3 Unicast Packet Journey

GRE encap case: GRE adjacency

```
RP/0/RSP0/CPU0:rasr9k-1y#show cef vrf DEF
adjacency tunnel-ip 25 hardware ingress location
0/4/CPU0
Sat Feb 23 14:44:52.239 UTC
Display protocol is ipv4
Interface    Address
Type    Refcount

ti25         Prefix: 0.0.0.0/32
local   3
             Adjacency: PT:0x8aa0c0c8 0.0.0.0/32
             Interface: ti25
             GRE header:
0000004500400000a2fb2fff011414ac021414ac00080000
             GRE tunnel adjacency
             GRE tunnel info: 0x91b3b050 (0x1 3),
tos-propagate is set
             Interface Type: 0x25, Base Flags:
0x2001 (0x90fdd3b4)
             Nhinfo PT: 0x90fdd3b4, Idb PT:
0x8d8f8898, If Handle: 0x8000120
             Dependent adj type: remote
(0x90fdd460)
             Dependent adj intf: ti25
             Ancestor If Handle: 0x0
```

```
TX H/W Result for NP:6 (index: 0x28a (BE)):
.
    adj_resolve_control_byte0
                    reserved: 0       egr_uidb_internal:
1
                       match: 1                   valid:
1
                  iptunl_adj: 1
.
    flags
        gre_adj          : 1

    uidb_index       : 0x1b00 (LE)
    reserve_pad_word: 0
    l3_mtu           : 1476
    reserve_pad_1    : 0
    adj_stats_index : 0x18005f00

    GRE Adj
        ip_src   : 172.20.20.1      ip_dst   : 172.20.20.2
        tos      : 0                ttl      : 0xff
        df       : 1                tos_reflect : 1
        rsvd flag bits   : 0        encap_checksum   :
0x40a3
        vrf_id   : 0                reserved: 0
```

L3 Unicast Packet Journey

IP to MPLS-TE case

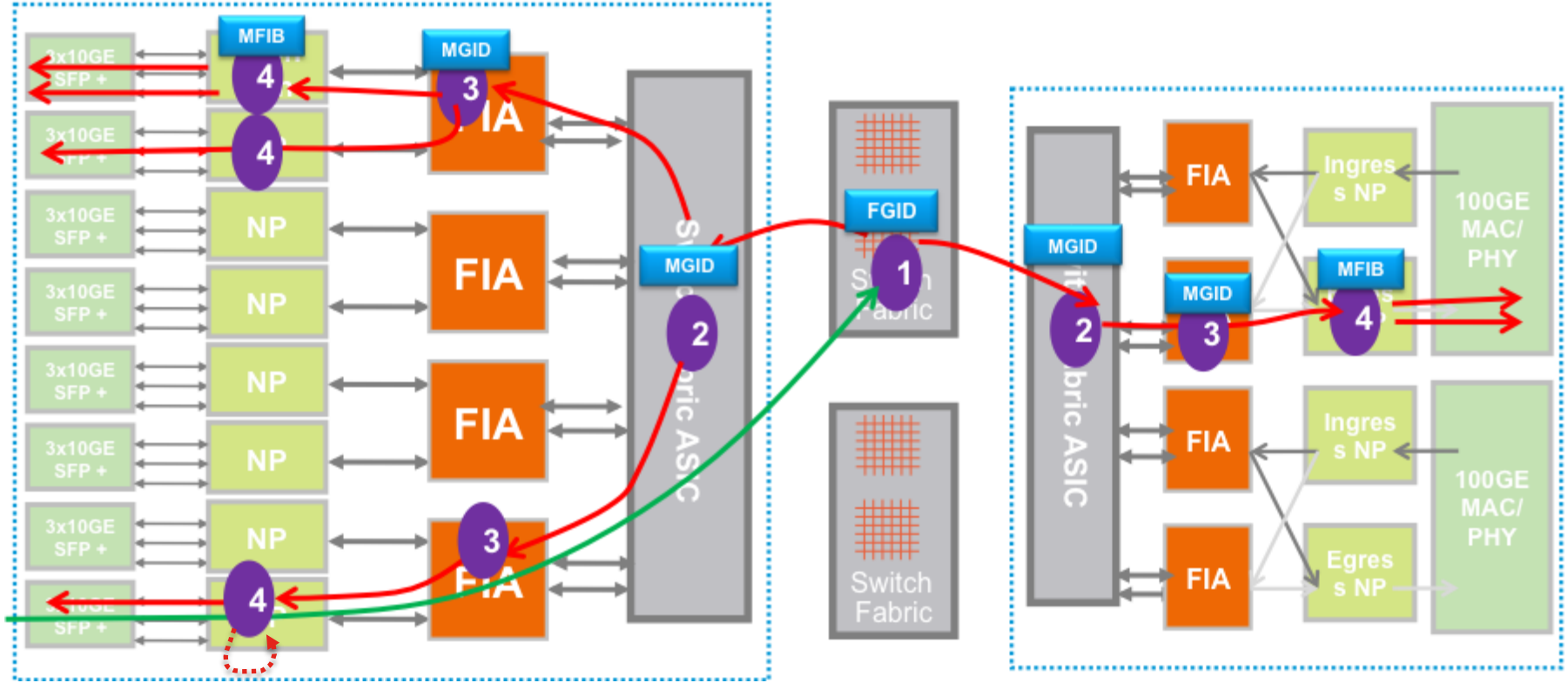
```
RP/0/RSP0/CPU0:rasr9k-1y#show cef ipv4 172.29.2.1
hardware ingress location 0/4/CPU0
Sat Feb 23 15:22:57.224 UTC
172.29.2.0/24, version 259, internal 0x4004001
(ptr 0x8efba154) [1], 0x0 (0x8e0ece00), 0x440
(0x90dca470)
 Updated Feb 22 11:03:15.593
 Prefix Len 24, traffic index 0, precedence
routine (0), priority 3
   via 192.168.20.242, tunnel-te200, 5
dependencies, weight 0, class 0 [flags 0x0]

 TE-NH H/W Result for 1st NP:0 (index: 0x3 (BE)):
.
          tunnel_over_tunnel: 0
spare: 0
.
    TE_local_label:
                        label: 16012
                          exp: 0
eos: 1
    TE_tunnel_label:
                        label: 0
                          exp: 0
eos: 1
         te_nh_stats_ptr: 0x880a5f
.
```

```
 RX H/W Result for 1st NP:0 (index: 0x38 (BE)):

    Raw Data0: 0x91000000 0000005c 00000640
00000000
    adj_resolve_control_byte0
                          match: 1
valid: 1
                    iptunl_adj: 0
                   remote_rack: 0

    adj_resolve_control_byte1
                     adj_down: 0
mgscp_en: 0
      rx_lag_hash_en: 0
          rx_lag_adj: 0

    adj_resolve_control_byte2
                    rx_lag_adj: 0
rx_adj_null0: 0
                   rp_destined: 0
rx_punt: 0
                       rx_drop: 0
    sfp/vqi          : 0x5c
    if_handle        : 0x640

.
```

L3 Multicast Packet Replication

Multicast Transit Frame Path

Replication stages: per LC, per FIA, per NP, per interface



Multicast Transit Frame Path

Lookups and replications

- Two MFIB lookups
  - Ingress NP tags frame with FGID and MGID, and ingress feature processing
    - FGID [Fabric Group ID]: to which LC's
    - MGID [Multicast Group ID]: Entry index to egress NP's
- Ingress replication
  - Only in case of routing mcast packets out of a bridge domain via BVI.
  - Original frame is forwarded at L2, at least one copy at L3 (depends on number of egress BVIs and xconnect distribution)
- Egress replication
  - 1st replication is at central switch fabric. Replicates to 1 copy per egress line card [FGID]
  - 2nd replication at each egress LC switch fabric. Replicates to 1 copy per egress FIA [MGID]
  - 3rd replication at each egress FIA. Replicates to 1 copy per egress NP [MGID]
  - 4th replication at each egress NP. Replicates to 1 copy for each egress interface [MFIB]
- Per flow load sharing
  - FIA hashes to LC fabric links
  - LC fabric hashes to RSP fabric links

Multicast Fabric Group ID

FGID = destination card

Slot		Slot Mask	
Logical	Physical	Binary	Hex
LC7	9	1000000000	0x0200
LC6	8	0100000000	0x0100
LC5	7	0010000000	0x0080
LC4	6	0001000000	0x0040
RSP0	5	0000100000	0x0020
RSP1	4	0000010000	0x0010
LC3	3	0000001000	0x0008
LC2	2	0000000100	0x0004
LC1	1	0000000010	0x0002
LC0	0	0000000001	0x0001

9906

9010

9910/12/22

Slot		Slot Mask	
Logical	Physical	Binary	Hex
LC3	5	0000100000	0x0020
LC2	4	0000010000	0x0010
LC1	3	0000001000	0x0008
LC0	2	0000000100	0x0004
RSP1	1	0000000010	0x0002
RSP0	0	0000000001	0x0001

Slot		Slot Mask	
Logical	Physical	Binary	Hex
LC19	21	10000 00000000 00000000	0x10 0000
LC1-18	3-20		
LC0	2	0000000100	0x0004
RP1	1	0000000010	0x0002
RP0	0	0000000001	0x0001

- Follows the sequence of slots in chassis

L3 Multicast Packet Journey

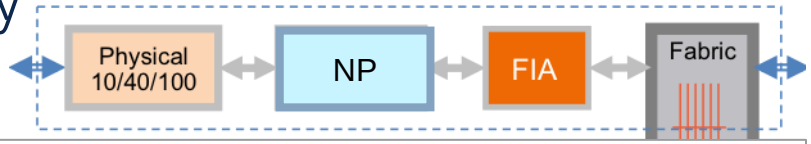
Reading the mRIB

```
RP/0/RSP0/CPU0:rasr9k-1y#show mrib route 232.1.1.1 172.30.1.1 detail
Tue Feb 26 17:15:05.039 UTC

IP Multicast Routing Information Base
Entry flags: L - Domain-Local Source, E - External Source to the Domain,
.

(172.30.1.1,232.1.1.1) Ver: 0x5180 RPF nbr: 172.29.1.2 Flags:,
 PD: Slotmask: 0x41
     MGID: 16903
  Up: 5d09h
  Incoming Interface List
    TenGigE0/0/0/2 Flags: A IC II LI, Up: 5d09h
  Outgoing Interface List
    TenGigE0/0/0/0 Flags: F IC NS II LI, Up: 4d08h
    TenGigE0/0/0/1 Flags: F IC NS II LI, Up: 4d08h
    TenGigE0/4/0/2.2 Flags: F NS LI, Up: 5d09h
    TenGigE0/4/0/20.6 Flags: F IC NS II LI, Up: 00:59:25
    GigabitEthernet0/0/1/19 Flags: F IC NS II LI, Up: 01:08:45
```

0x41 = 0100 0001:
LC0 & LC4 in 9010
[physical slots 0, 6]

L3 Multicast Packet Journey

Reading the NP mFIB

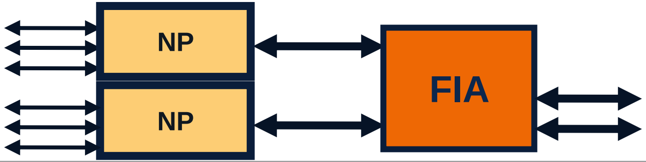



```
RP/0/RSP0/CPU0:rasr9k-1y#show mfib hardware route detail 232.1.1.1 172.30.1.1 location 0/4/CPU0
Tue Feb 26 18:09:54.515 UTC
LC Type: Typhoon A9K-24x10GE-TR
.
Source: 172.30.1.1       Group: 232.1.1.1       Mask: 64  RPF Int: Te0/0/0/2
MGID: 16903      MLI: 5          Fabric Slotmask: 0x41       FGID: 0x41
  Route Information
  ------------------------------------------------------------------------------
  NP B  S  DC PL PR PF DR RI          T  OC   MF TR TE TD CD MI   Base
  ------------------------------------------------------------------------------
  0  F  F  F  F  F  F  F  0x640       0  1    F  F  F  F  F  0x0  0x5100d4
  1  F  F  F  F  F  F  F  0x640       0  0    F  F  F  F  F  0x0  0x
  2  F  F  F  F  F  F  F  0x640       0  0    F  F  F  F  F  0x0  0x
  3  F  F  F  F  F  F  F  0x640       0  0    F  F  F  F  F  0x0  0x
  4  F  F  F  F  F  F  F  0x640       0  0    F  F  F  F  F  0x0  0x
  5  F  F  F  F  F  F  F  0x640       0  0    F  F  F  F  F  0x0  0x5100d4
  6  F  F  F  F  F  F  F  0x640       1  1    F  F  F  F  F  0x0  0x5100d4
  7  F  F  F  F  F  F  F  0x640       0  0    F  F  F  F  F  0x0  0x5100d4
  ------------------------------------------------------------------------------

  Software MGID Information
  ------------------------------------------------
  MGID: 16903 Mask: 0x41  Old MGID: 0 Old Mask: 0x1
  ------------------------------------------------
```

Outgoing interface count per NP

L3 Multicast Packet Journey

Reading egress LC Fabric and FIA replication




```
RP/0/RSP0/CPU0:rasr9k-1y#show controllers mgidprgm mgidindex 16903 location 0/0/CPU0
Tue Feb 26 17:35:10.026 UTC

Device          MGID-Bits       Client-Last-Modified
=======================================================

XBAR-0          11              MFIBV4
FIA-0           10              MFIBV4
FIA-1           10              MFIBV4

RP/0/RSP0/CPU0:rasr9k-1y#show controllers mgidprgm mgidindex 16903 location 0/4/CPU0
Tue Feb 26 17:35:15.417 UTC

Device          MGID-Bits       Client-Last-Modified
=======================================================

XBAR-0          1001            MFIBV4
FIA-0           1               MFIBV4
FIA-1           0               MFIBV4
FIA-2           0               MFIBV4
FIA-3           1               MFIBV4
```

MGID

Egress LC

Fabric to 1st & 2nd FIA

FIA to 2nd NP

Fabric to 1st & 4th FIA

None

FIA to 1st NP

L3 Multicast Packet Journey

Reading the hardware counters

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show mfib vrf TRAFFIC hardware route statistics 232.1.1.100
192.5.1.100 location 0/1/CPU0
Thu Jan  9 22:09:04.997 EST
LC Type: Typhoon A9K-MOD80-SE
------------------------------------------------------------------------
Legend:
    N:    NP ID                    R:    Received
    F:    Forwarded                P:    Punted to CPU
    ID:   Ingress Drop             ED:   Egress Drop

Source: 192.5.1.100  Group: 232.1.1.100  Mask:64
  ------------------------------------------------------------------------
  NP    R(packets:bytes)/F(packets:bytes)/P(packets)/ID(packets)/ED(packets)
  ------------------------------------------------------------------------
  0     18326252:4948100612 / 0:0 / 0 / 0 / 0
  1     0:0 / 18212304:4917314359 / 0 / 0 / 0
  ------------------------------------------------------------------------
  Interface Statistics:
  ------------------------------------------------------------------------
  C   Interface       F/P/D (packets:bytes)
  ------------------------------------------------------------------------
  1   Gi100/0/0/9     18212651:4917382603 / 0:0 / 0:0
  ------------------------------------------------------------------------
```

Received on NP0

Forwarded on NP1

Forwarded on interface [satellite]

L2 Frame Forwarding & Flooding

L2 Frame Journey

L2 frame forwarding/flooding

```
RP/0/RSP0/CPU0:rasr9k-1y#show l2vpn forwarding
bridge-domain BRIDGES:DOMAIN-A hardware ingress
detail location 0/4/CPU0
Sun Feb 24 13:53:34.530 UTC
Bridge-domain name: BRIDGES:DOMAIN-A, id: 0,
state: up
 MAC learning: enabled
 MAC port down flush: enabled
 Flooding:
   Broadcast & Multicast: enabled
   Unknown unicast: enabled
 MAC aging time: 300 s, Type: inactivity
 MAC limit: 2000, Action: limit, no flood,
Notification: syslog, trap
 MAC limit reached: no
 MAC Secure: enabled, Logging: enabled, Action: none
 DHCPv4 snooping: profile not known on this node
 Dynamic ARP Inspection: enabled, Logging: enabled
 Dynamic ARP Inspection Address Validation:
   IPv4 verification: enabled
   Source MAC verification: enabled
   Destination MAC verification: enabled
 IP Source Guard: disabled, Logging: disabled
 IGMP snooping: disabled, flooding: enabled
 Bridge MTU: 1500 bytes
 Number of bridge ports: 4
 Number of MAC addresses: 2002
.
```

```
 Bridge Domain: 0 NP 0
    Flags: Virtual Table, Multicast Flooding, Learn
Enable, No Learn, Learn Drop
    Num Members: 0, Learn Key: 0x00, Half Age: 5
    fgid shg0: 0x0001, fgid shg1: 0x0041, fgid
shg2: 0x0041
    PBB Core BD: 0, ISID: 0
 Bridge Domain: 0 NP 1
    Flags: Virtual Table, Multicast Flooding, Learn
Enable, No Learn, Learn Drop
    Num Members: 0, Learn Key: 0x00, Half Age: 5
    fgid shg0: 0x0001, fgid shg1: 0x0041, fgid
shg2: 0x0041
    PBB Core BD: 0, ISID: 0
.
 Bridge Domain: 0 NP 6
    Flags: Virtual Table, Multicast Flooding, Learn
Enable, No Learn, Learn Drop
    Num Members: 1, Learn Key: 0x00, Half Age: 5
    fgid shg0: 0x0001, fgid shg1: 0x0041, fgid
shg2: 0x0041
    PBB Core BD: 0, ISID: 0
    Bridge Port 0
      XID: 0x09b00001, Active
          virtual
      XID: 0x09b00001, Active
.
```

L2 Frame Journey

L2 frame forwarding/flooding

```
! CONTINUED
 TenGigE0/4/0/20.101, state: oper up
    Number of MAC: 0
    Statistics:
      packets: received 36731752, sent 14772099
      bytes: received 2203905120, sent 886325940
    Storm control drop counters:
      packets: broadcast 0, multicast 0, unknown
unicast 2961034169
      bytes: broadcast 0, multicast 0, unknown unicast
177662050140
    Dynamic arp inspection drop counters:
      packets: 0, bytes: 0
    IP source guard drop counters:
      packets: 0, bytes: 0
  Platform Bridge Port context:
  Ingress State: Bound
    Flags: DAI, DAI-ipv4, DAI-src-MAC, DAI-dst-MAC, DAI-log,
MAC-SEC, MAC-SEC-log,
           MAC-learn-disabled
  MAC Security Actions: Drop, No Notify
  Platform AC context:
  Ingress AC: VPLS, State: Bound
    Flags: Learn Limit - No Learn, Learn Limit - Drop, Storm
Control BCast,
           Storm Control MCast, Storm Control UCast, Port
Level MAC Limit
  XID: 0x09b00001, SHG: None
  Ingress uIDB: 0x001c, Egress uIDB: 0x001c, NP: 6,
Port Learn Key: 0
.
```

```
 NP6
    Ingress uIDB:
      Flags: DAI, DAI Notification, Dest MAC validation,
IP Addr Validation,
             L2PT, L2, Source MAC validation, Status, Ext
Required, VLAN Ops,
             VPLS
      Stats Ptr: 0x000000, uIDB index: 0x001c, Wire Exp
Tag: 1
      BVI Bridge Domain: 0, BVI Source XID: 0x00000000
      VLAN1: 0, VLAN1 etype: 0x0000, VLAN2: 0, VLAN2
etype: 0x0000
      L2 ACL Format: 0, L2 ACL ID: 0, IPV4 ACL ID: 0,
IPV6 ACL ID: 0
      QOS ID: 0, QOS Format ID: 0
      Local Switch dest XID: 0x09b00001
      UIDB IF Handle: 0x0c000042, Source Port: 0, Num
VLANs: 0
    Xconnect ID: 0x09b00001, NP: 6
      Type: AC
      Flags: Learn enable, Type 5, Learn limit no learn,
Learn limit drop,
             Broadcast storm control, Multicast storm
control, Unknown unicast storm control,
             VPLS
      uIDB Index: 0x001c
      Bridge Domain ID: 0, Stats Pointer: 0xf78122
      Storm Control enabled for: Broadcast, Multicast,
Unknown Unicast, Pointer: 0x00001801
.
```

L2 MAC

MAC learning and synchronization

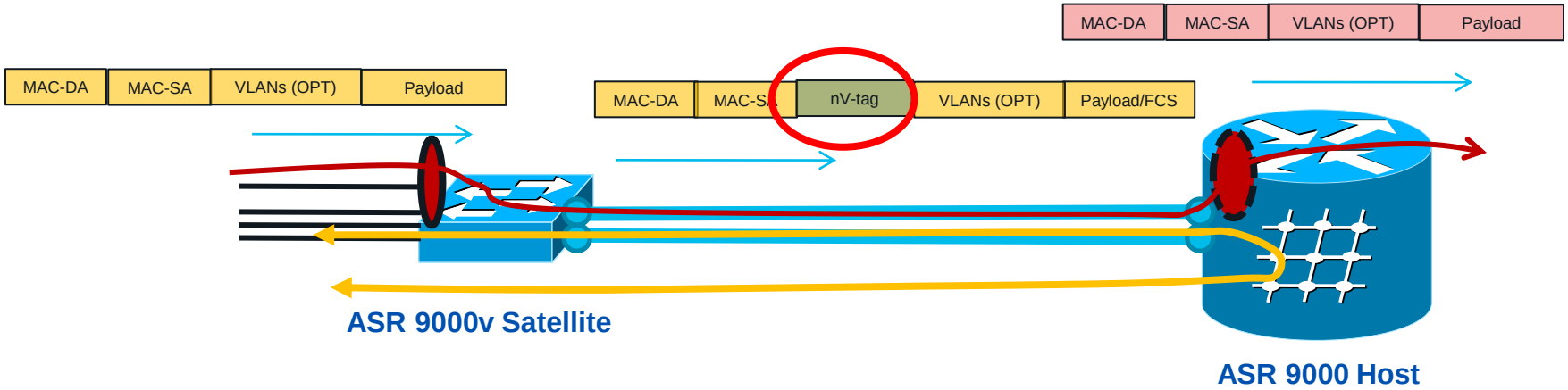
```
RP/0/RSP0/CPU0:rasr9k-1y#show l2vpn forwarding bridge-domain BRIDGES:DOMAIN-A mac-address
hardware ingress location 0/4/CPU0
Fri Feb 22 18:50:08.433 UTC
 To Resynchronize MAC table from the Network Processors, use the command...
    l2vpn resynchronize forwarding mac-address-table location <r/s/i>

Mac Address    Type    Learned from/Filtered on    LC learned Resync Age          Mapped to
--------------------------------------------------------------------------------
0000.c001.0167 dynamic Gi0/0/1/0.200              0/0/CPU0   0d 0h 0m 18s        N/A
0000.c001.016b dynamic Gi0/0/1/0.200              0/0/CPU0   0d 0h 0m 11s        N/A
0000.c001.016c dynamic Gi0/0/1/0.200              0/0/CPU0   0d 0h 0m 9s         N/A
0000.c001.016d dynamic Gi0/0/1/0.200              0/0/CPU0   0d 0h 0m 18s        N/A
0000.c001.016e dynamic Gi0/0/1/0.200              0/0/CPU0   0d 0h 0m 20s        N/A
0000.c001.016f dynamic Gi0/0/1/0.200              0/0/CPU0   0d 0h 0m 8s         N/A
0000.c001.0171 dynamic Gi0/0/1/0.200              0/0/CPU0   0d 0h 0m 17s        N/A
0000.c001.0102 dynamic Te0/4/0/20.101             0/4/CPU0   0d 0h 0m 16s        N/A
0000.c001.0104 dynamic Te0/4/0/20.101             0/4/CPU0   0d 0h 0m 20s        N/A
0000.c001.0105 dynamic Te0/4/0/20.101             0/4/CPU0   0d 0h 0m 8s         N/A
0000.c001.0106 dynamic Te0/4/0/20.101             0/4/CPU0   0d 0h 0m 9s         N/A
0000.c001.0107 dynamic Te0/4/0/20.101             0/4/CPU0   0d 0h 0m 18s        N/A
0000.c001.0108 dynamic Te0/4/0/20.101             0/4/CPU0   0d 0h 0m 15s        N/A
0000.c001.0109 dynamic Te0/4/0/20.101             0/4/CPU0   0d 0h 0m 3s         N/A
0000.c001.010a dynamic Te0/4/0/20.101             0/4/CPU0   0d 0h 0m 4s         N/A
.
```

ASR 9000 Satellite to/from Host

cisco Live!

ASR 9000v “Satellite”

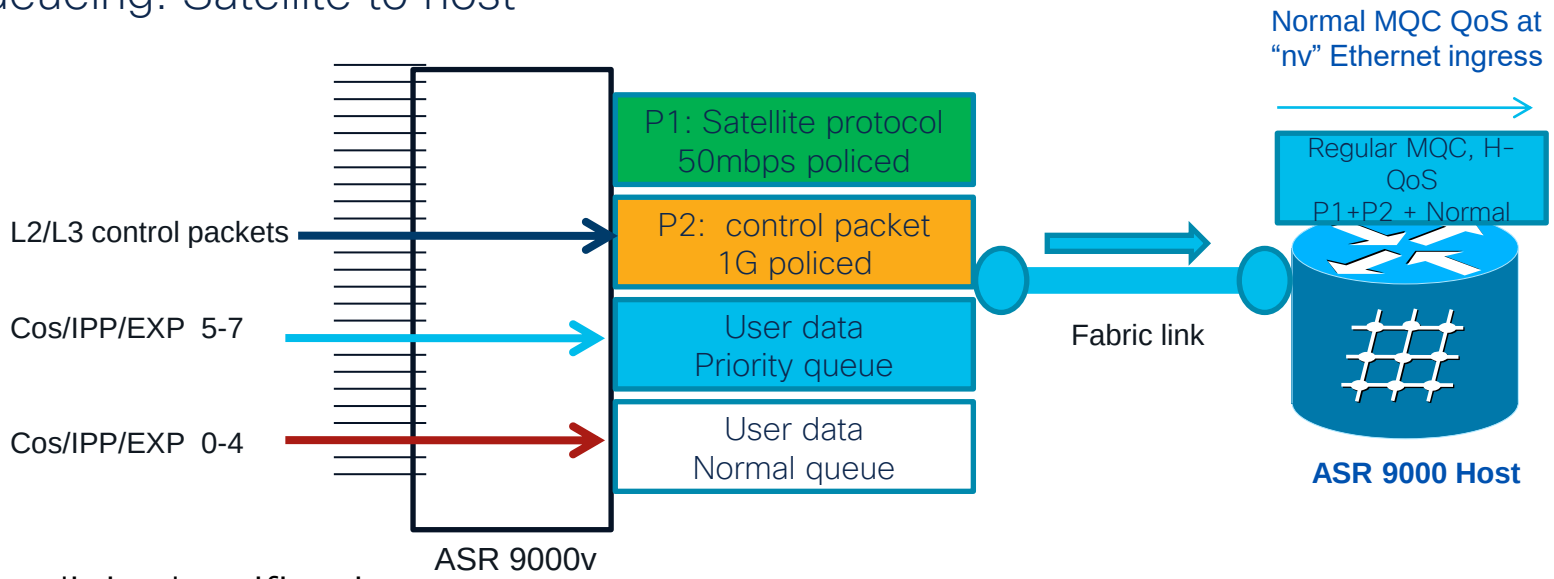
Traffic flow



- “nV” tag identifies the satellite port traffic
- No learning or switching on satellite
- L2/3/4 and ingress/egress QoS done on host

ASR 9000v “Satellite”

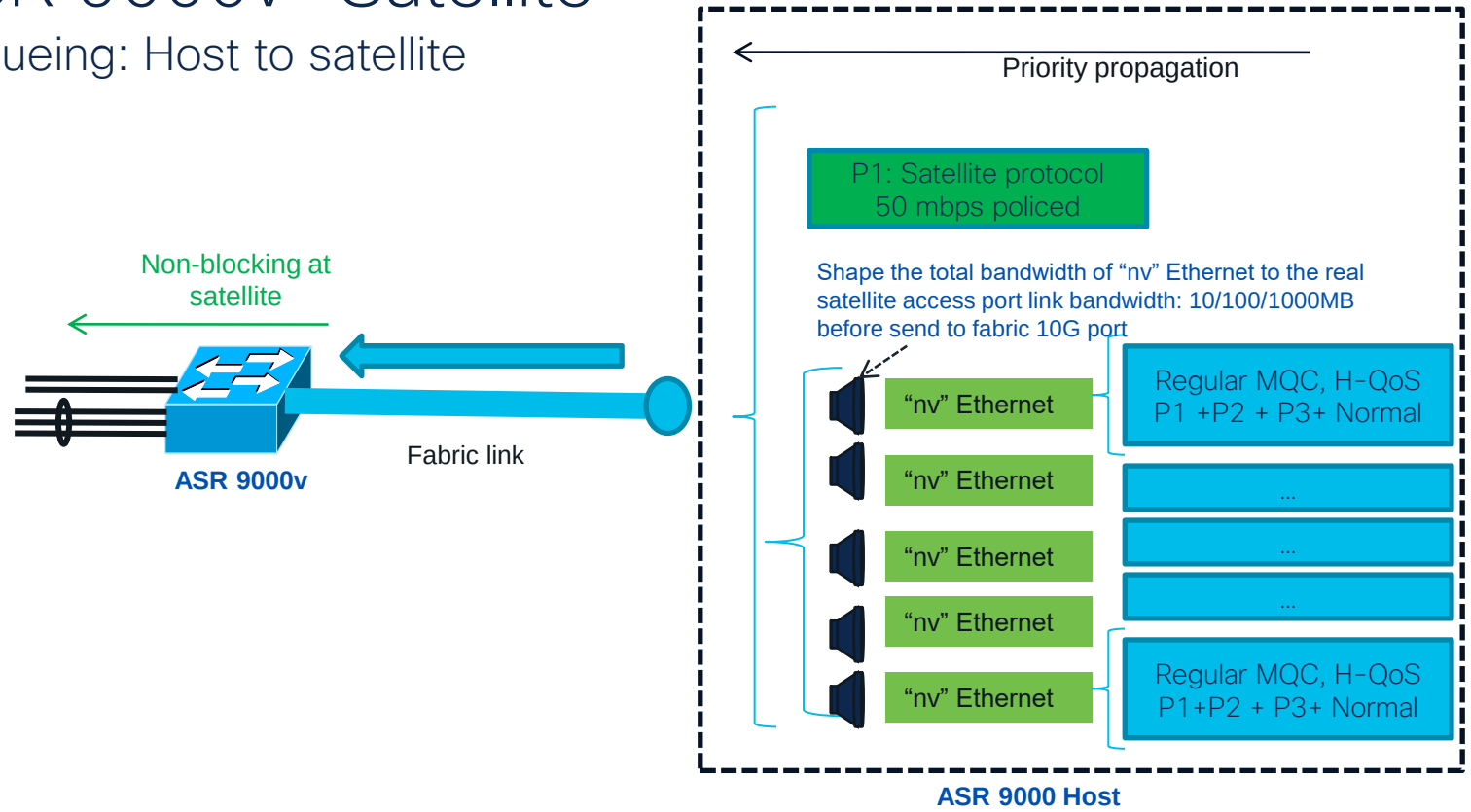
Queueing: Satellite to host



- Implicit classification
- P1 and P2 are strict priority
- User data priority:normal 100:1 bandwidth

ASR 9000v “Satellite”

Queueing: Host to satellite

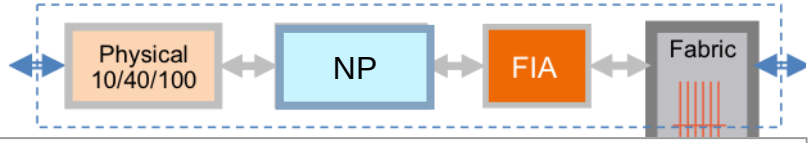


Under the Hood: Inside NP, FIA, and Fabric

cisco Live!

Frame Path: Stats & Drops

Reading the physical counters

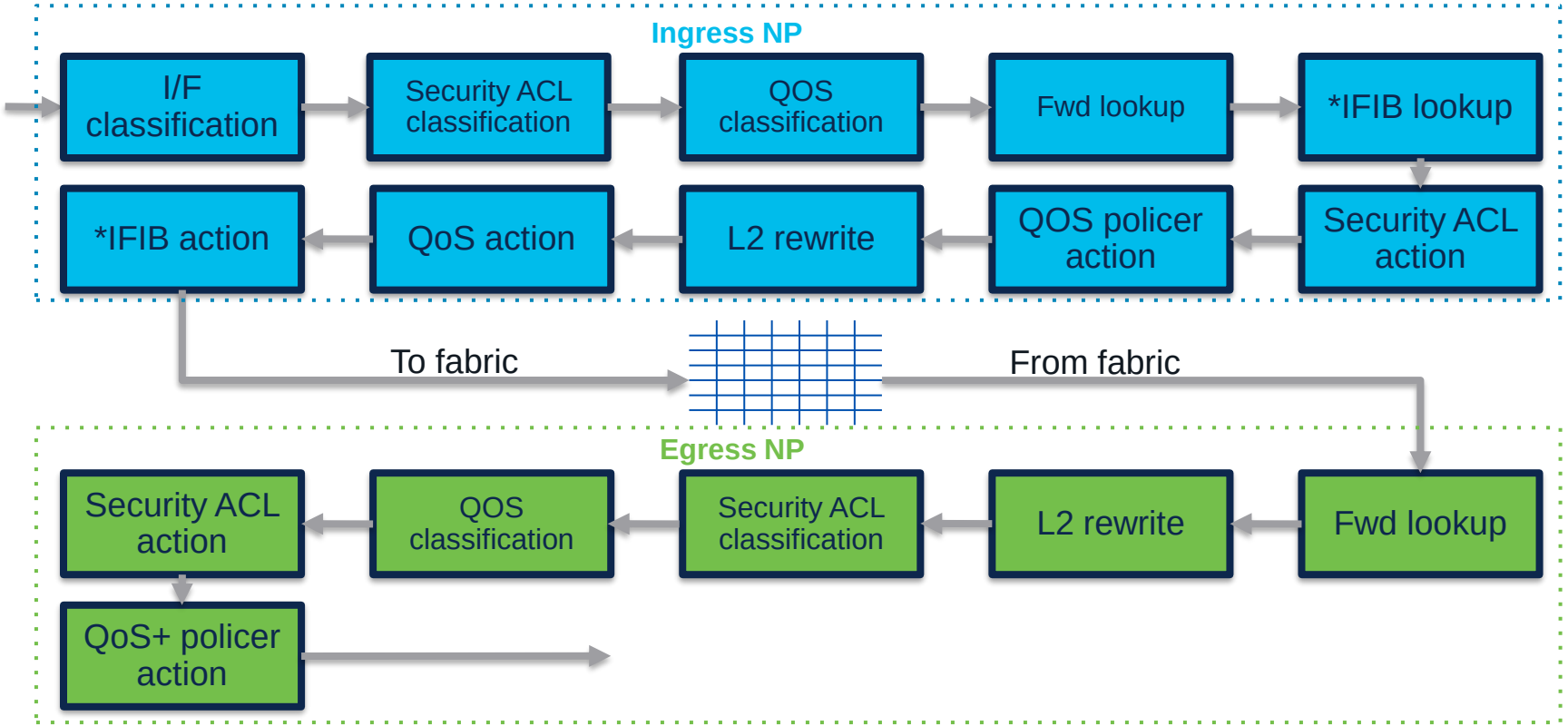



```
RP/0/RSP0/CPU0:rasr9k-1y#show controllers
TenGigE0/4/0/20 stats
Sun Feb 24 14:44:18.899 UTC
Statistics for interface TenGigE0/4/0/20 (cached
values):
Ingress:
    Input total bytes           = 3081227904920
    Input good bytes            = 3081227904920

    Input total packets         = 23220024479
    Input 802.1Q frames         = 0
    Input pause frames          = 0
    Input pkts 64 bytes         = 7143534733
    Input pkts 65-127 bytes     = 2888766549
    Input pkts 128-255 bytes    = 13124923916
    Input pkts 256-511 bytes    = 62799261
    Input pkts 512-1023 bytes   = 0
    Input pkts 1024-1518 bytes  = 0
    Input pkts 1519-Max bytes   = 0

    Input good pkts             = 23220024479
    Input unicast pkts          = 23220023458
    Input multicast pkts        = 62
    Input broadcast pkts        = 959

    Input drop overrun          = 0
.
```

```
Egress:
    Output total bytes          = 1345771624
    Output good bytes           = 1345771624

    Output total packets        = 21895707
    Output 802.1Q frames        = 0
    Output pause frames         = 0
    Output pkts 64 bytes        = 21665536
    Output pkts 65-127 bytes    = 21179
    Output pkts 128-255 bytes   = 168767
    Output pkts 256-511 bytes   = 40225
    Output pkts 512-1023 bytes  = 0
    Output pkts 1024-1518 bytes = 0
    Output pkts 1519-Max bytes  = 0

    Output good pkts            = 21895707
    Output unicast pkts         = 21870499
    Output multicast pkts       = 25195
    Output broadcast pkts       = 13
    Output drop underrun        = 0
    Output drop abort           = 0
    Output drop other           = 0

    Output error other          = 0
```

NP Feature Processing

Order of processing by ingress and egress NP's



Inside NP

Processing pipelines

Inside NP

Reading pipeline counters


```
RP/0/RSP0/CPU0:rasr9000-2w-a#show controllers NP counters np0 location 0/1/CPU0
Wed Nov 27 21:09:07.635 EST

                Node: 0/1/CPU0:
----------------------------------------------------------------

Show global stats counters for NP0, revision v2

Read 64 non-zero NP counters:
Offset  Counter                                        FrameValue   Rate (pps)
--------------------------------------------------------------------------------
  16    MDF_TX_LC_CPU                                     6722114           10
  17    MDF_TX_WIRE                                       1826039            3
  21    MDF_TX_FABRIC                                     1635541            2
  29    PARSE_FAB_RECEIVE_CNT                             1837406            3
  33    PARSE_INTR_RECEIVE_CNT                            5083364            7
  37    PARSE_INJ_RECEIVE_CNT                             1228130            2
 499    RSV_ING_L2_SMAC_MISS                                   60            0
 502    RSV_ING_L2_LEARN                                       60            0
 541    RSV_REFRESH_FROM_NOTIFY_CNT                            62            0
 584    RSV_L2BC_BVI                                            2            0
 604    RESOLVE_REMOTE_RACK_PREP_CNT                      5539915            8
 708    LRN_PERIODIC_AGING_DELETE_ENTRY                        60            0
 774    ARP                                                   119            0
 848    PUNT_ADJ                                                2            0
 852    PUNT_ACL_DENY                                         161            0
 900    PUNT_STATISTICS                                   5083356            7
 902    PUNT_DIAGS_RSP_ACT                                  11419            0
 904    PUNT_DIAGS_RSP_STBY                                 11427            0
```

List of NP counters:
https://supportforums.cisco.com/docs/DOC-26566

NP Counters and Rates

Example: Ingress NP, no drops


```
RP/0/RSP0/CPU0:rasr9000-2w-b#show controllers NP counters np0 location 0/0/CPU0
Mon Dec  9 15:16:34.889 EST
                Node: 0/0/CPU0:
----------------------------------------------------------------

Show global stats counters for NP0, revision v2

Read 59 non-zero NP counters:
Offset  Counter                                        FrameValue   Rate (pps)
--------------------------------------------------------------------------------
  16  MDF_TX_LC_CPU                                     10255120            8
  17  MDF_TX_WIRE                                    6382883323428            1
  21  MDF_TX_FABRIC                                  8903307706961     31250074
  29  PARSE_FAB_RECEIVE_CNT                          6382883151049            0
  33  PARSE_INTR_RECEIVE_CNT                            8653828            8
  37  PARSE_INJ_RECEIVE_CNT                              744943            1
  41  PARSE_ENET_RECEIVE_CNT                         8910925981070     31250074
  45  PARSE_TM_LOOP_RECEIVE_CNT                         8035316            5
  49  PARSE_TOP_LOOP_RECEIVE_CNT                            61            0
  57  PARSE_ING_DISCARD                                2344591            0
 195  PRS_HEALTH_MON                                   8035316            5
 204  INTR_FRAME_TYPE_7                                8653827            8
 214  DBG_PRS_EP_L_PRS_VPLS_PW_IMPOSE                      10            0
 233  PARSE_RSP_INJ_FAB_CNT                             70634            0
 235  PARSE_RSP_INJ_DIAGS_CNT                           55255            0
 236  PARSE_EGR_INJ_PKT_TYP_UNKNOWN                     66847            0
 237  PARSE_EGR_INJ_PKT_TYP_IPV4                         3787            0
 246  PARSE_LC_INJ_FAB_CNT                             101092            0
.
```

To FIA

From Phy

NP Counters and Rates

NP drops, rate and direction


```
RP/0/RSP0/CPU0:rasr9000-2w-b#show controllers NP counters np0 location 0/0/CPU0
Tue Dec 10 14:18:39.195 EST

                Node: 0/0/CPU0:
----------------------------------------------------------------

Show global stats counters for NP0, revision v2

Read 59 non-zero NP counters:
Offset  Counter                                          FrameValue   Rate (pps)
-------------------------------------------------------------------------------
  16  MDF_TX_LC_CPU                                        11004363            9
  17  MDF_TX_WIRE                                      8712222364719     29761820
  21  MDF_TX_FABRIC                                   11063035007386     27714366
  29  PARSE_FAB_RECEIVE_CNT                            8712222113330     29761820
  33  PARSE_INTR_RECEIVE_CNT                              9401470            9
  37  PARSE_INJ_RECEIVE_CNT                                832185            1
  41  PARSE_ENET_RECEIVE_CNT                          11070653296959     27714366
  45  PARSE_TM_LOOP_RECEIVE_CNT                            8437075            5
 359  PARSE_MAC_NOTIFY_RCVD                                    183            0
 367  PARSE_FAST_DISCARD_LOW_PRIORITY_DROP_0           106211394050       883832
 368  PARSE_FAST_DISCARD_LOW_PRIORITY_DROP_1           106210662138       883856
 369  PARSE_FAST_DISCARD_LOW_PRIORITY_DROP_2           106211061617       883943
 370  PARSE_FAST_DISCARD_LOW_PRIORITY_DROP_3           106211474043       883922
 373  DBG_RSV_EP_L_RSV_ING_L3_IFIB                      3707021673            0
 830  PUNT_NO_MATCH                                        4746            0
 831  PUNT_NO_MATCH_EXCD                                 464963896            0
 849  PUNT_ADJ_EXCD                                        273406            0
 852  PUNT_ACL_DENY                                       1479378            0
 853  PUNT_ACL_DENY_EXCD                                1163570900            0
```

To egress

To fabric

From fabric

From interface

NP catching up

NP Counters and Rates

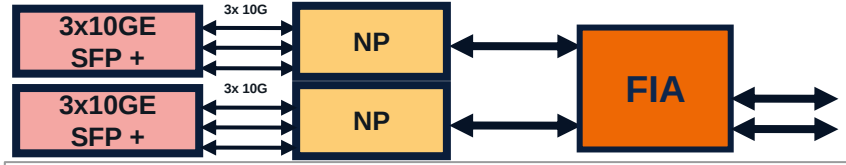
Traffic Manager drops


```
RP/0/RSP0/CPU0:rasr9000-2w-b#show controllers NP tm counters np1 location 0/0/CPU0
Tue Dec 10 14:40:47.210 EST

                Node: 0/0/CPU0:
----------------------------------------------------------------

==== TM Counters (NP 1 TM 0) ====
 TM Counters:
xmt paks: 897837659243,  xmt bytes: 62718673698431
drop paks: 29447137293,  drop_bytes: 2002405351616

RP/0/RSP0/CPU0:rasr9000-2w-b#
RP/0/RSP0/CPU0:rasr9000-2w-b#
RP/0/RSP0/CPU0:rasr9000-2w-b#show controllers NP tm counters np1 location 0/0/CPU0
Tue Dec 10 14:40:49.816 EST

                Node: 0/0/CPU0:
----------------------------------------------------------------

==== TM Counters (NP 1 TM 0) ====
 TM Counters:
xmt paks: 897909308598,  xmt bytes: 62723686013270
drop paks: 29466027670,  drop_bytes: 2003689898884
```

FIA Counters

FIA counts, drops and direction

```
RP/0/RSP0/CPU0:rasr9000-2w-b#show controllers fabric
fia instance 0 stats location 0/0/CPU0
Tue Dec 10 14:49:58.704 EST

 ********** FIA-0 **********
Category: count-0
         From Unicast Xbar[0]              733461306331
         From Unicast Xbar[1]              733460650405
         From Unicast Xbar[2]                         0
         From Unicast Xbar[3]                         0
       From MultiCast Xbar[0]                    233068
       From MultiCast Xbar[1]                         0
       From MultiCast Xbar[2]                         0
       From MultiCast Xbar[3]                         0
           To Unicast Xbar[0]              933450146675
           To Unicast Xbar[1]              932066610046
           To Unicast Xbar[2]                         0
           To Unicast Xbar[3]                         0
         To MultiCast Xbar[0]                    451799
         To MultiCast Xbar[1]                         0
         To MultiCast Xbar[2]                         0
         To MultiCast Xbar[3]                         0
         To Line Interface[0]           8759312354291
         To Line Interface[1]            457138023968
       From Line Interface[0]          11117127781061
       From Line Interface[1]            489302108080
                Ingress drop:             97191712670
                 Egress drop:                       0
                 Total drop:               97191712670
```




```
RP/0/RSP0/CPU0:rasr9000-2w-b#show controllers fabric fia
instance 0 drops ingress location 0/0/CPU0
Tue Dec 10 15:33:37.655 EST
 ********** FIA-0 **********
Category: in_drop-0
              From Spaui Drop-0                          0
                    accpt tbl-0                          0
                      ctl len-0                          0
                    short pkt-0                          0
                   max pkt len-0                         0
                   min pkt len-0                         0
              From Spaui Drop-1                          0
                    accpt tbl-1                          0
                      ctl len-1                          0
                    short pkt-1                          0
                   max pkt len-1                         0
                   min pkt len-1                         0
                        Tail drp             125787328841
                         Vqi drp                         0
               Header parsing drp                        0
                  pw to ni drp                           0
               ni from pw drp                            0
                  sp0 crc err                            0
                sp0 bad align                            0
                 sp0 bad code                            0
               sp0 align fail                            3
                 sp0 prot err                            0
                 sp1 crc err                             0
.
```

Back pressure from egress NP

Agenda

- ✓ System Architecture: System anatomy & health
- ✓ Operating System & Configuration: IOS-XR & configuration models
- ✓ Control, Management, Security: Processing of control & exceptions
- ✓ Transit Packet/Frame Journey: Life of L3/L2 unicast/multicast
- ➢ MPLS Operation: Processing, forwarding & L3/L2 service operation
- ➢ Troubleshooting: Diagnostics, counters, drops, and packet capture

5 MPLS Operation

cisco Live!

MPLS in the Data Plane

Main MPLS Label Usage:

- Forwarding Label
  - The destination for this label is … [FEC]. Usually, a host address of label edge router (LER)
  - A path to a label destination is a label switched path (LSP)
  - Intermediate nodes may not know much about payload or the basis for its forwarding
  - Ultimate destination may not need the label [PHP] --- Penultimate hop popping
- Service Label
  - How to handle this payload [IP, L3VPN VRF, L2VPN, PW, CEoP, control]
  - Significant to edge nodes. The forwarding nodes along the path may not know what it means

Forwarding

Without label

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show cef ipv4 10.101.188.1
.
 remote adjacency to GigabitEthernet0/1/0/1
 Prefix Len 32, traffic index 0, precedence routine (0), priority 1
   via 10.100.11.1, GigabitEthernet0/1/0/1, 4 dependencies, weight 0, class 0 [flags
0x0]
    path-idx 0 [0x721f30e0 0x0]
    next hop 10.100.11.1
    remote adjacency
```

IP nexthop

With label

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show cef ipv4 10.101.188.1
.
 remote adjacency to GigabitEthernet0/1/0/1
 Prefix Len 32, traffic index 0, precedence routine (0), priority 1
   via 10.100.11.1, GigabitEthernet0/1/0/1, 20 dependencies, weight 0, class 0 [flags
0x0]
    path-idx 0 [0x723990b4 0x0]
    next hop 10.100.11.1
    remote adjacency
     local label 111012      labels imposed {101000}
```

MPLS in

MPLS out

Label Operations

```
RP/0/0/CPU0:P101#show cef ipv4 10.101.188.1/32
.
 Prefix Len 32, traffic index 0, precedence n/a, priority 1
   via 10.100.108.1, tunnel-te181, 3 dependencies, weight 0, class 0 [flags 0x0]
    path-idx 0 [0xacc9d674 0x0]
    next hop 10.100.108.1
    local adjacency
     local label 101000      labels imposed {108000}

RP/0/0/CPU0:P101#show mpls forwarding
Thu Jun  6 09:16:22.581 EDT
Local  Outgoing    Prefix            Outgoing      Next Hop        Bytes
Label  Label       or ID             Interface                     Switched
------ ----------- ----------------- ------------  --------------- ------------
101000 108000      10.101.188.1/32   tt181         10.100.108.1    0
101001 102000      11210             Gi0/0/0/0.112 10.100.112.2    0
101005 Pop         10.101.111.1/32   tt1111        10.101.111.1    0
101007 Exp-Null-v4 11211             Gi0/0/0/1     10.100.11.11    0
.
101039 Unlabelled  10.101.124.1/32   Gi0/0/0/0.112 10.100.112.2    0
101040 Unlabelled  10.101.125.1/32   Gi0/0/0/0.112 10.100.112.2    4591105
.
```

Label Push

Swap

Pop top label

Unlabeled: Pop all to bottom of stack

Aggregate Label

- Bottom of stack label not sufficient for forwarding decision
- Payload header has to be used for forwarding lookup

Pop topmost & forward

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show mpls forwarding
Mon Jun 17 21:05:46.166 EDT
Local  Outgoing    Prefix             Outgoing     Next Hop        Bytes
Label  Label       or ID              Interface                    Switched
------ ----------- ------------------ ------------ --------------- ------------
.
111007 101000      10.101.188.1/32    tt1111       10.100.101.1    375048
111014 Pop         PW(10.101.188.1:1) Gi0/1/0/3.1  point2point     314906
111015 Aggregate   CUST-A: Per-VRF Aggr[V]   \
                                      CUST-A                       6320
111016 Unlabelled  172.20.210.0/24[V] Gi0/1/0/3.200 172.20.200.2    0
```

Pop all labels & forward

Pop & lookup

Aggregate Label vs. Non-Aggregate

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show mpls forwarding vrf CUST-A detail
Mon Jun 17 21:31:10.474 EDT
Local  Outgoing    Prefix            Outgoing     Next Hop        Bytes
Label  Label       or ID             Interface                    Switched
------ ----------- ----------------- ------------ --------------- ------------
111015 Aggregate   CUST-A: Per-VRF Aggr[V]   \
                                     CUST-A                       8240
     Updated Jun 17 20:03:20.046
     Path Flags: 0x10 [ ]
     MAC/Encaps: 0/0, MTU: 0
     Label Stack (Top -> Bottom): { }
     Packets Switched: 65

111016 Unlabelled  172.20.210.0/24[V] Gi0/1/0/3.200 172.20.200.2    0
     Updated Jun 17 21:23:42.495
     Version: 47, Priority: 3
     MAC/Encaps: 18/18, MTU: 1500
     Label Stack (Top -> Bottom): { Unlabelled }
     Packets Switched: 0

RP/0/RSP0/CPU0:rasr9000-2w-a#show route vrf CUST-A ipv4 static
Mon Jun 17 21:34:57.549 EDT
S    172.20.210.0/24 [1/0] via 172.20.200.2, 00:11:45
```

Pop & lookup

Pop & forward

Aggregate Label: Example

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show route vrf CUST-A ipv4 connected
Mon Jun 17 21:34:49.647 EDT

C    172.20.200.0/24 is directly connected, 00:56:39, GigabitEthernet0/1/0/3.200

RP/0/RSP0/CPU0:rasr9000-2w-a#show mpls forwarding vrf CUST-A detail
Mon Jun 17 21:31:10.474 EDT
Local  Outgoing    Prefix            Outgoing     Next Hop        Bytes
Label  Label       or ID             Interface                    Switched
------ ----------- ----------------- ------------ --------------- ------------
111015 Aggregate   CUST-A: Per-VRF Aggr[V]   \
                                     CUST-A                       8240
     Updated Jun 17 20:03:20.046
     Path Flags: 0x10 [ ]
     MAC/Encaps: 0/0, MTU: 0
     Label Stack (Top -> Bottom): { }
     Packets Switched: 65

.
```

Pop & lookup

Non-Aggregate Label: Example

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show route vrf CUST-A ipv4 static
Mon Jun 17 21:34:57.549 EDT

S    172.20.210.0/24 [1/0] via 172.20.200.2, 00:11:45


RP/0/RSP0/CPU0:rasr9000-2w-a#show mpls forwarding vrf CUST-A detail
Mon Jun 17 21:31:10.474 EDT

111016 Unlabelled  172.20.210.0/24[V] Gi0/1/0/3.200 172.20.200.2    0
     Updated Jun 17 21:23:42.495
     Version: 47, Priority: 3
     MAC/Encaps: 18/18, MTU: 1500
     Label Stack (Top -> Bottom): { Unlabelled }
     Packets Switched: 0
```

Pop & forward

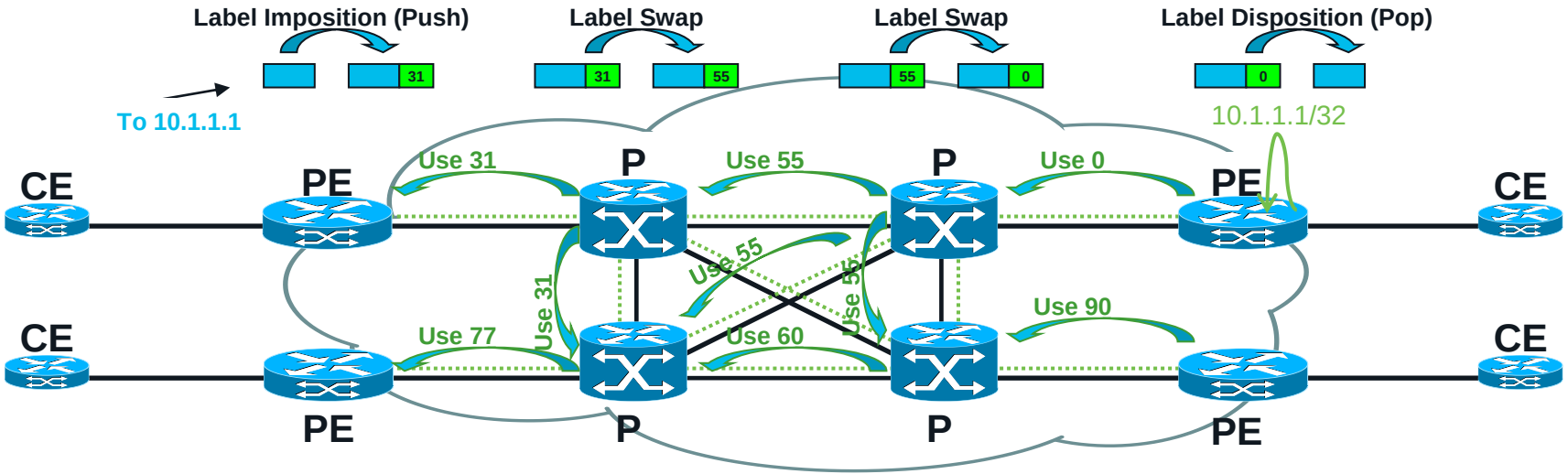
Forwarding Labels

MPLS Datapath Forwarding Characteristics

- Data-plane level operations:
  - **Push** (one or more labels)
    - Service label: 1
    - L2VPN FAT label: 1
    - LDP: 1
    - RSVP: 1
    - FRR: on ASR9k max +1 (if more, TI-LFA creates a backup tunnel)
    - SRTE: on ASR9k max 10
  - **Pop** (one or more)
  - **Swap**

LDP: Label Distribution Protocol

- Binds and advertises labels for all IGP prefixes [Cisco default]
- Multicast hellos for neighbor discovery. TCP 646 for session.
- Hop by hop. No guarantee for end to end LSP.



LDP: Control Plane: Neighbor Sessions

```
RP/0/0/CPU0:P101#show mpls ldp neighbor
Thu Jun  6 10:41:01.283 EDT

Peer LDP Identifier: 10.100.108.1:0
  TCP connection: 10.100.108.1:31207 - 10.100.101.1:646; MD5 on
  Graceful Restart: Yes (Reconnect Timeout: 120 sec, Recovery: 180 sec)
  Session Holdtime: 180 sec
  State: Oper; Msgs sent/rcvd: 27745/27777; Downstream-Unsolicited
  Up time: 2w2d
  LDP Discovery Sources:
    Targeted Hello (10.100.101.1 -> 10.100.108.1, active)
  Addresses bound to this peer:
    10.100.87.8      10.100.108.1     10.100.168.8     10.100.178.8         10.100.188.8

Peer LDP Identifier: 10.101.111.1:0
  TCP connection: 10.101.111.1:35863 - 10.100.101.1:646; MD5 on
  Graceful Restart: Yes (Reconnect Timeout: 120 sec, Recovery: 0 sec)
  Session Holdtime: 180 sec
  State: Oper; Msgs sent/rcvd: 3024/3020; Downstream-Unsolicited
  Up time: 1d19h
  LDP Discovery Sources:
    Targeted Hello (10.100.101.1 -> 10.101.111.1, active)
    GigabitEthernet0/0/0/1
  Addresses bound to this peer:
    10.100.11.11     10.101.111.1     172.16.200.150   172.16.200.151       192.168.2.2
```

My potential Dest prefixes through this neighbor as next hop

LDP: Control Plane: Label Binding

```
RP/0/0/CPU0:P101#show mpls ldp bindings
.
10.101.111.1/32, rev 161
        Local binding: label: 101005
        Remote bindings: (1 peers)
            Peer                 Label
            -----------------    --------
            10.100.108.1:0       108009
10.101.112.1/32, rev 116
        Local binding: label: 101041
        Remote bindings: (2 peers)
            Peer                 Label
            -----------------    --------
            10.100.108.1:0       108038
            10.101.111.1:0       111006
10.101.124.1/32, rev 117
        Local binding: label: 101039
        Remote bindings: (2 peers)
            Peer                 Label
            -----------------    --------
            10.100.108.1:0       108031
            10.101.111.1:0       111007
```

Use the one matching IGP route, if any

```
10.101.125.1/32, rev 118
        Local binding: label: 101040
        Remote bindings: (2 peers)
            Peer                 Label
            -----------------    --------
            10.100.108.1:0       108032
            10.101.111.1:0       111008
10.101.135.1/32, rev 119
        Local binding: label: 101042
        Remote bindings: (2 peers)
            Peer                 Label
            -----------------    --------
            10.100.108.1:0       108033
            10.101.111.1:0       111009
10.101.137.1/32, rev 120
        Local binding: label: 101043
        Remote bindings: (2 peers)
            Peer                 Label
            -----------------    --------
            10.100.108.1:0       108034
            10.101.111.1:0       111010
.
```

LDP: Forwarding: FIB and LFIB

```
RP/0/RSP0/CPU0:rasr9000-2w-b#show cef ipv4 10.101.111.1/32
Wed Mar 19 12:25:01.496 EDT
10.101.111.1/32, version 272, internal 0x4004001 0x0 (ptr 0x7238643c) [1], 0x0 (0x71635290),
0x450 (0x71e26460)
 Updated Mar 19 12:23:44.913
 remote adjacency to GigabitEthernet0/1/0/1
 Prefix Len 32, traffic index 0, precedence n/a, priority 3
   via 10.100.188.8, GigabitEthernet0/1/0/1, 20 dependencies, weight 0, class 0 [flags 0x0]
    path-idx 0 NHID 0x0 [0x719ea954 0x0]
    next hop 10.100.188.8
    remote adjacency
     local label 188017      labels imposed {108002}

RP/0/RSP0/CPU0:rasr9000-2w-b#show mpls forwarding labels 188017 detail
Wed Mar 19 12:25:05.202 EDT
Local  Outgoing    Prefix            Outgoing     Next Hop        Bytes
Label  Label       or ID             Interface                    Switched
------ ----------- ----------------- ------------ --------------- ------------
188017 108002      10.101.111.1/32   Gi0/1/0/1    10.100.188.8    1558
     Updated Mar 19 12:23:44.913
     Version: 272, Priority: 3
     MAC/Encaps: 14/18, MTU: 1386
     Label Stack (Top -> Bottom): { 108002 }
     NHID: 0x5
     Packets Switched: 19
```

IPv4 in

MPLS in

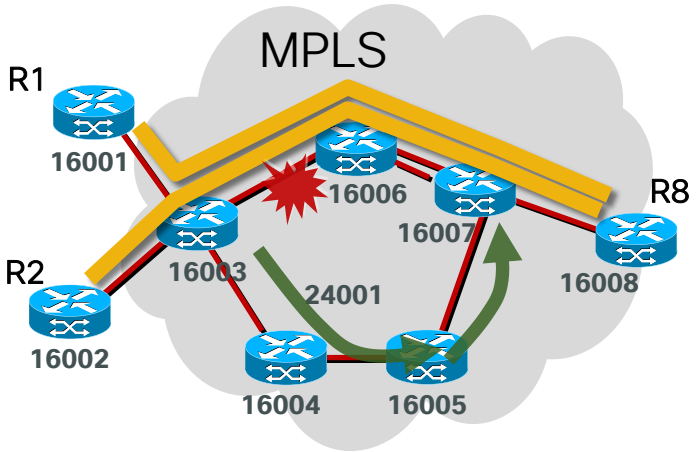
LDP: Forwarding: In the Forwarding Plane

```
RP/0/RSP0/CPU0:rasr9000-2w-b#show cef ipv4
10.101.111.1/32 hardware ingress location 0/1/CPU0
.
 local adjacency 10.100.188.8
 Prefix Len 32, traffic index 0, precedence n/a,
priority 3
   via 10.100.188.8, GigabitEthernet0/1/0/1, 21
dependencies, weight 0, class 0 [flags 0x0]
    path-idx 0 NHID 0x5 [0x8b15d134 0x0]
    next hop 10.100.188.8
    local adjacency
     local label 188017      labels imposed {108002}
.
NR-LDI H/W Result for path 0 [index: 0x34c7 (BE),
common to all NPs]:
.
        output_label: 108002
        label_msb: 0x1a5e               label_lsb: 0x2
                exp: 0x0                        eos: 0x1
.
 RX H/W Result for 1st NP:0 (index: 0x3b (BE)):
.
 if_handle       : 0x3e0
.
TX H/W Result for NP:0 (index: 0x33d3 (BE)):
.
    uidb_index      : 0x900 (LE)
    l3_mtu          : 1386
    adj_stats_index : 0x381f61
    dest_mac        : 0x000c.29f4.90c6
.
```

```
RP/0/RSP0/CPU0:rasr9000-2w-b#show mpls forwarding labels
188017 hardware ingress location 0/1/CPU0
Wed Mar 19 13:01:00.202 EDT
Local  Outgoing Prefix         Outgoing  Next Hop    Bytes
Label  Label    or ID          Interface           Switched
------ ------ --------------- ------------ ------------
188017 108002 10.101.111.1/32 Gi0/1/0/1 10.100.188.8 N/A
.
NR-LDI H/W Result for path 0 [index: 0x34c7 (BE),
common to all NPs]:
.
        output_label: 108002
                label_msb: 0x1a5e
label_lsb: 0x2
                        exp: 0x0
eos: 0x1
.
 RX H/W Result for 1st NP:0 (index: 0x3b (BE)):
.
 if_handle       : 0x3e0
.
TX H/W Result for NP:0 (index: 0x33d3 (BE)):
.
    uidb_index      : 0x900 (LE)
    l3_mtu          : 1386
    adj_stats_index : 0x381f61
    dest_mac        : 0x000c.29f4.90c6
    ether reserved  : 0000000000000000
.
```

Segment Routing Path Control & TiLFA

Segment Routing Traffic Engineering (SR-TE)



- Source Routing
  - Source chooses a path and encodes in the packet header as an ordered list of segments
  - The rest of the network nodes executes the encoded instructions
- SR-TE Policy Path Control
  - Policy label stack with Node-SID, or Adj-SID
  - Each Policy assigned unique Binding-SID
  - ECMP paths load-balance by IGP Nature
- Topology Independent LFA
  - Automated 1:N path protection
  - Local reroute comparable to MPLS TE Link / Node, but no RSVP or LDP
  - IGP algorithm, support microloop avoidance

Segment Routing: Control Plane: LFIB

```
RP/0/0/CPU0:P103#show mpls forwarding
Sat Jan  7 17:54:03.427 UTC
Local  Outgoing    Prefix             Outgoing     Next Hop        Bytes
Label  Label       or ID              Interface                    Switched
------ ----------- ------------------ ------------ --------------- ------------
15101  Pop         SRLB (idx 101)     Gi0/0/0/0    100.1.1.2       0
15102  Pop         SRLB (idx 102)     Gi0/0/0/1    100.1.1.6       0
16002  Pop         SR Pfx (idx 2)     Gi0/0/0/0    100.1.1.2       0
       16002       SR Pfx (idx 2)     Gi0/0/0/1    100.1.1.6       0          (!)
16003  Pop         SR Pfx (idx 3)     Gi0/0/0/1    100.1.1.6       0
       16003       SR Pfx (idx 3)     Gi0/0/0/0    100.1.1.2       0          (!)
16004  16004       SR Pfx (idx 4)     Gi0/0/0/0    100.1.1.2       0
       16004       SR Pfx (idx 4)     Gi0/0/0/1    100.1.1.6       0          (!)
16005  16005       SR Pfx (idx 5)     Gi0/0/0/1    100.1.1.6       0
       16005       SR Pfx (idx 5)     Gi0/0/0/0    100.1.1.2       0          (!)
16006  16006       SR Pfx (idx 6)     Gi0/0/0/0    100.1.1.2       0
       16006       SR Pfx (idx 6)     Gi0/0/0/1    100.1.1.6       0
24000  Pop         SR Adj (idx 0)     Gi0/0/0/1    100.1.1.6       0
24001  Pop         SR Adj (idx 0)     Gi0/0/0/1    100.1.1.6       0
```

SR Prefix or Adjacency labels

backup path

Segment Routing: Forwarding: FIB and LFIB

IPv4 in

```
RP/0/0/CPU0:P103#show cef ipv4 1.1.1.5/32
Sat Jan  7 18:02:21.983 UTC
1.1.1.5/32, version 66, labeled SR, internal 0x1000001 0x81 (ptr 0xa14164b8) [1], 0x0 (0xa13f88c8), 0xa28 (0xa19ce148)
 Updated Jan  7 17:28:03.074
 local adjacency 100.1.1.6
 Prefix Len 32, traffic index 0, precedence n/a, priority 1
   via 100.1.1.2/32, GigabitEthernet0/0/0/0, 11 dependencies, weight 0, class 0, backup (Local-LFA) [flags 0x300]
    path-idx 0 NHID 0x0 [0xa18b8588 0x0]
    next hop 100.1.1.2/32
    local adjacency
     local label 16005      labels imposed {16005}
   via 100.1.1.6/32, GigabitEthernet0/0/0/1, 11 dependencies, weight 0, class 0, protected [flags 0x400]
    path-idx 1 bkup-idx 0 NHID 0x0 [0xa1998690 0x0]
    next hop 100.1.1.6/32
     local label 16005      labels imposed {16005}
```

MPLS in

```
RP/0/0/CPU0: P103#show mpls forwarding labels 16005 detail
Local  Outgoing    Prefix             Outgoing     Next Hop        Bytes
Label  Label       or ID              Interface                    Switched
------ ----------- ------------------ ------------ --------------- ------------
16005  16005       SR Pfx (idx 5)     Gi0/0/0/1    100.1.1.6       0
     Updated: Jan  7 17:25:26.824
     Path Flags: 0x400 [  BKUP-IDX:0 (0xa1998690) ]
     Version: 66, Priority: 1
     Label Stack (Top -> Bottom): { 16005 }
     NHID: 0x0, Encap-ID: N/A, Path idx: 1, Backup path idx: 0, Weight: 0
     MAC/Encaps: 14/18, MTU: 1500
     Outgoing Interface: GigabitEthernet0/0/0/1 (ifhandle 0x00000040)
     Packets Switched: 0
       16005       SR Pfx (idx 5)     Gi0/0/0/0    100.1.1.2       0            (!)
```

Segment Routing: Forwarding: In the Forwarding Plane

```
RP/0/0/CPU0:P103#show cef ipv4 1.1.1.5/32 hardware ingress location 0/0/CPU0
Sat Jan  7 18:16:11.586 UTC
1.1.1.5/32, version 66, labeled SR, internal 0x1000001 0x81 (ptr 0xa14164b8) [1], 0x0 (0xa13f88c8), 0xa28 (0xa19ce148)
 Updated Jan  7 17:28:03.073
 local adjacency 100.1.1.6
 Prefix Len 32, traffic index 0, precedence n/a, priority 1
   via 100.1.1.2/32, GigabitEthernet0/0/0/0, 11 dependencies, weight 0, class 0, backup (Local-LFA) [flags 0x300]
    path-idx 0 NHID 0x0 [0xa18b8588 0x0]
    next hop 100.1.1.2/32
    local adjacency
     local label 16005      labels imposed {16005}
   via 100.1.1.6/32, GigabitEthernet0/0/0/1, 11 dependencies, weight 0, class 0, protected [flags 0x400]
    path-idx 1 bkup-idx 0 NHID 0x0 [0xa1998690 0x0]
    next hop 100.1.1.6/32
```

IPv4 in

```
RP/0/0/CPU0:P103#show mpls forwarding labels 16005 hardware ingress location 0/0/CPU0
Sat Jan  7 18:20:16.329 UTC
Local  Outgoing    Prefix            Outgoing     Next Hop        Bytes
Label  Label       or ID             Interface                    Switched
------ ----------- ----------------- ------------ --------------- ------------
16005  16005       SR Pfx (idx 5)    Gi0/0/0/1    100.1.1.6       N/A
       16005       SR Pfx (idx 5)    Gi0/0/0/0    100.1.1.2       N/A          (!)
```

MPLS in

Use Case: L2VPN Preferred-path SRTE Policy

- Use preferred-path configuration to specify SR-TE Policy used to transport Pseudowire service traffic

```
l2vpn
 pw-class EoMPLS-PWCLASS
  encapsulation mpls
   preferred-path sr-te policy
srte_c_2_ep_1.1.1.4
 xconnect group vpws
  p2p vpws1
   interface Bundle-Ether2.2
   neighbor ipv4 1.1.1.4 pw-id 1
    pw-class EoMPLS-PWCLASS
!
segment-routing
 traffic-eng
  policy POLICY1
   color 2 end-point ipv4 1.1.1.4
   autoroute
    include ipv4 10.10.1.0/24
   binding-sid mpls 999
   candidate-paths
    preference 100
     dynamic
      metric
       type te
```

```
R1# show l2vpn xc pw-id 1 detail
Group vpws, XC vpws1, state is up; Interworking none
  AC: Bundle-Ether2.2, state is up
…
  PW: neighbor 1.1.1.4, PW ID 1, state is up ( established )
    PW class xc-vpls, XC ID 0xa000001f
    Encapsulation MPLS, protocol LDP
    Source address 1.1.1.1
    PW type Ethernet, control word disabled, interworking none
    PW backup disable delay 0 sec
    Sequencing not set
    Preferred path Active : SR TE srte_c_2_ep_1.1.1.4, Statically
configured, fallback disabled
    Tunnel: Up
```

L2VPN Prefer-Path SRTE Forwarding

```
R1# show segment-routing traffic-eng policy name
srte_c_2_ep_1.1.1.4
Color: 2, End-point: 1.1.1.4
  Name: srte_c_2_ep_1.1.1.4
  Status:
    Admin: up  Operational: up for 02:48:30 (since Apr  1
12:22:57.663)
  Candidate-paths:
    Preference: 200 (configuration) (active)
      Name: POLICY1
      Requested BSID: 999
      PCC info:
        Symbolic name: cfg_ POLICY1_discr_200
        PLSP-ID: 4
      Explicit: segment-list SIDLIST1 (valid)
        Weight: 1, Metric Type: TE
          16002 [Prefix-SID, 1.1.1.2]
          24004 [Adjacency-SID, 10.23.0.1 - 10.23.0.2]
          16004 [Prefix-SID, 1.1.1.4]
    Preference: 100 (configuration)
      Requested BSID: 999
      PCC info:
        Symbolic name: cfg_POLICY1_discr_100
        PLSP-ID: 6
```

```
R1# show l2vpn forwarding detail location 0/2/cpu0
Local interface: Bundle-Ether2.2, Xconnect id: 0x1, Status: up
  Segment 1
    AC, Bundle-Ether2.2, Ethernet VLAN mode, status: Bound
    Statistics:
      packets: received 0, sent 0
      bytes: received 0, sent 0
  Segment 2
    MPLS, Tunnel interface: srte_c_2_ep_1.1.1.4, status: Bound,
Active
    Pseudowire label: 24007
    Control word disabled
    Backup PW
      MPLS, Destination address: 1.1.1.5, pw-id: 1, status: Bound
      Pseudowire label: 24008
    Statistics:
      packets: received 0, sent 0
      bytes: received 0, sent 0
```

Use Case: L3VPN Dynamic SRTE Policy

- Egress PE5 populates a vrf prefix 11.11.1.0/24 requires low latency service

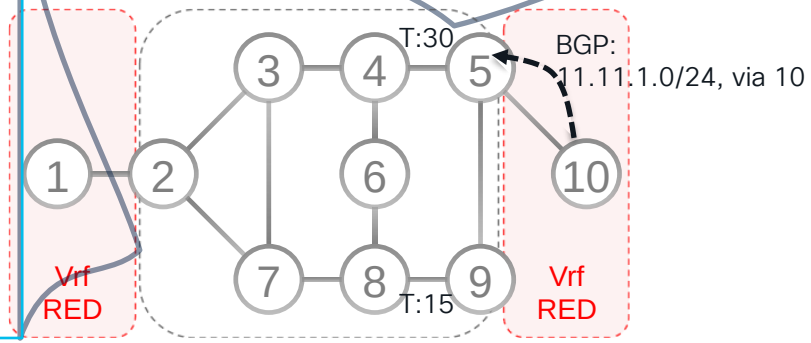
  => PE5 tags with extcommunity “color2”

- Ingress PE2 initiates on-demand policy by color template

  => PE2 finds a path to PE5 with optimized metric latency

```
extcommunity-set opaque color2
  2
end-set
!
route-policy bgp_col
 if destination in (11.11.1.0/24) then
    set extcommunity color color2
  endif
end-policy
!
router bgp 1
 neighbor 1.1.1.2
   remote-as 1
   address-family vpnv4 unicast
    route-policy bgp_col out
```

```
router bgp 1
 neighbor 1.1.1.5
  address-family vpnv4 unicast
 vrf RED
  rd 2:2
  address-family ipv4 unicast
!
segment-routing
 traffic-eng
  on-demand color 2
   dynamic
     metric
     type latency
!
performance-measurement
 interface GigabitEthernet0/1/1/0
  delay-measurement
```



measure link delay to compute path latency
 # show isis database R1 verbose | inc Delay
Link Average Delay: 7 us
Link Min/Max Delay: 7/7 us

L3VPN Work-flow and SRTE Steering

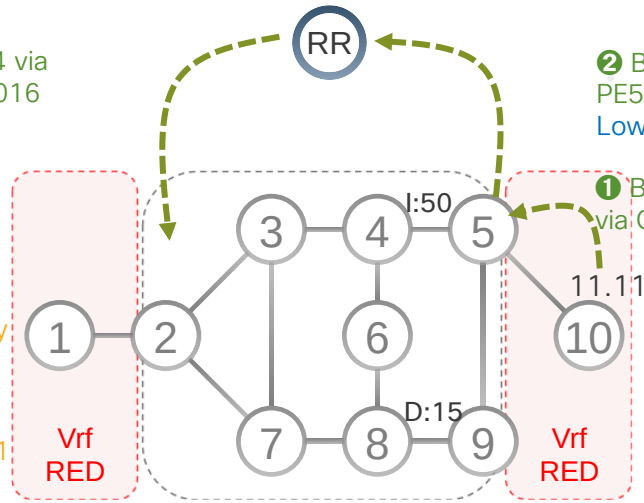
```
R1# show segment-routing traffic-eng policy color 2
Color: 2, End-point: 1.1.1.5
  Name: srte_c_2_ep_1.1.1.5
  Status:
    Admin: up  Operational: up for 00:39:14 (since Mar 31)
  Candidate-paths:
    Preference: 200 (BGP ODN) (active)
      Requested BSID: dynamic
      PCC info:
        Symbolic name: bgp_c_2_ep_1.1.1.5_discr_200
        PLSP-ID: 2
      Dynamic (valid)
        Metric Type: LATENCY,   Path Accumulated Metric: 10
          16005 [Prefix-SID, 1.1.1.5]
    Preference: 100 (BGP ODN)
      Requested BSID: dynamic
      PCC info:
        Symbolic name: bgp_c_2_ep_1.1.1.5_discr_100
        PLSP-ID: 1
      Dynamic
        Metric Type: NONE,   Path Accumulated Metric: 0
  Attributes:
    Binding SID: 24031
    Forward Class: 0
    Steering BGP disabled: no
    IPv6 caps enable: yes
#show cef vrf RED 11.11.1.0/24
Prefix Len 24, traffic index 0, precedence n/a, priority 3
  via local-label 24031, 3 dependencies, recursive
```

❸ BGP: 11.11.1.0/24 via PE5 VPN-LABEL: 24016 Low-Delay (color 2)

❹ PE2 check prefix from PE5 with Low-latency (color 2)

❺ Initiate SRTE policy use template color 2

❻ Policy → SID-list <16005> BSID 24031



❷ BGP: 11.11.1.0/24 via PE5 VPN-LABEL: 24016 Low-Delay (color2)

❶ BGP: 11.11.1.0/24 via CE10

11.11.1.0/24

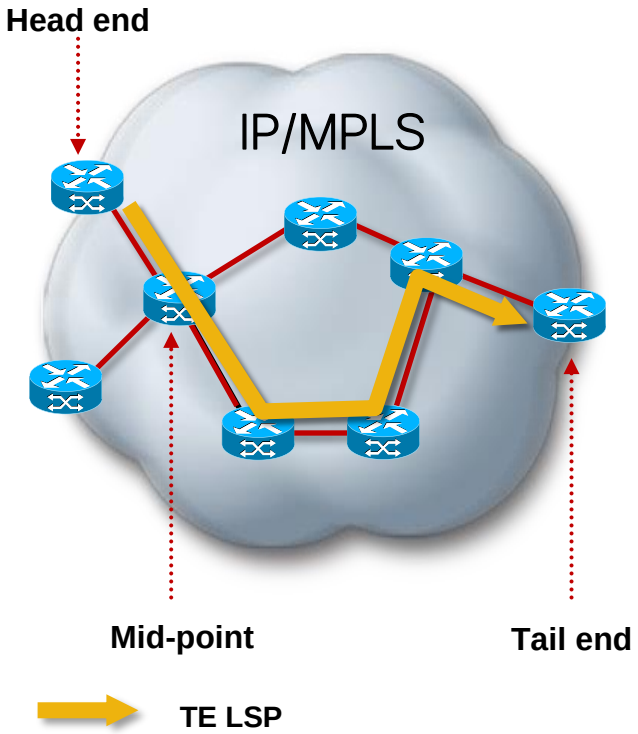
```
R1# show bgp vrf RED 11.11.1.0/24
Paths: (1 available, best #1)
Local
    2.2.2.2 C:2 (bsid:24031) (metric 10) from 2.2.2.2 (2.2.2.2)
      Received Label 24016
...
      Extended community: Color:2 RT:2:2
      SR policy color 2, up, registered, bsid 24031, if-handle 0x02000fe0
      Source AFI: VPNv4 Unicast, Source VRF: RED, Source Route Distinguisher: 2:2
```

L3VPN Dynamic SRTE Forwarding

```
R1# show bgp vpnv4 unicast vrf RED 11.11.1.0/24
BGP routing table entry for 11.11.1.0/24, Route Distinguisher: 1:1
Versions:
  Process           bRIB/RIB  SendTblVer
  Speaker                23          23
    Local Label: 24016
Last Modified: May 31 11:17:21.253 for 00:00:12
Paths: (2 available, best #1)
  Advertised to peers (in unique update groups):
    3.3.3.3
  Path #1: Received by speaker 0
  Advertised to peers (in unique update groups):
    3.3.3.3
  Local
    0.0.0.0 from 0.0.0.0 (1.1.1.1)
      Origin incomplete, metric 0, localpref 100, weight 32768, valid, redistributed, best,
group-best, import-candidate
      Received Path ID 0, Local Path ID 1, version 14
      Extended community: RT:1:1
  Path #2: Received by speaker 0
  Not advertised to any peer
  Local
    3.3.3.3 C:2 (bsid:24031) (metric 20) from 3.3.3.3 (3.3.3.3)
      Received Label 24016
      Origin incomplete, metric 0, localpref 100, valid, internal, import-candidate, imported
      Received Path ID 0, Local Path ID 0, version 0
      Extended community: Color:2 RT:1:1
      SR policy color 2, up, registered, bsid 24031, if-handle 0x00000250
      Source AFI: VPNv4 Unicast, Source VRF: RED, Source Route Distinguisher: 1:1
```

```
R1# show cef vrf RED 11.11.1.0/24 detail
11.11.1.0/24, version 11, internal 0x5000001 0x0 (ptr 0xa1251fa8) [1], 0x0 (0x0),
0x208 (0xa175739c)
 Updated May 31 11:30:35.697
 Prefix Len 24, traffic index 0, precedence n/a, priority 3
  gateway array (0xa1170a0c) reference count 1, flags 0x2038, source rib (7), 0
backups
              [1 type 1 flags 0x48441 (0xa1773898) ext 0x0 (0x0)]
  LW-LDI[type=0, refc=0, ptr=0x0, sh-ldi=0x0]
  gateway array update type-time 1 May 31 11:30:35.697
 LDI Update time May 31 11:30:35.697
  via local-label 24031, 3 dependencies, recursive [flags 0x6000]
    path-idx 0 NHID 0x0 [0xa17cdbb0 0x0]
    recursion-via-label
    next hop VRF - 'default', table - 0xe0000000
    next hop via 24031/0/21
     next hop srte_c_2_ep labels imposed {ImplNull 24016}


    Load distribution: 0 (refcount 1)

    Hash  OK  Interface                 Address
    0     Y   Unknown                   24031/0
```

How RSVP-TE Works



- Link information Distribution*
  - ISIS-TE
  - OSPF-TE
- Path Calculation (CSPF)*
- Path Setup (RSVP-TE)
- Forwarding Traffic down Tunnel*
  - Autoroute announce
  - Static route
  - PBTS / CBTS
  - Forwarding Adjacency
- Attributes / Affinity
- Bandwidth / Loadshare

TE: RSVP

```
RP/0/0/CPU0:P104#show rsvp neighbors
Mon Jun 10 08:38:07.069 EDT
Global Neighbor: 10.100.102.1
  Interface Neighbor   Interface
  -------------------- ------------
  10.100.124.2         GigabitEthernet0/0/0/0.124
Global Neighbor: 10.100.103.1
  Interface Neighbor   Interface
  -------------------- ------------
  10.100.134.3         GigabitEthernet0/0/0/0.134
Global Neighbor: 10.100.105.1
  Interface Neighbor   Interface
  -------------------- ------------
  10.100.145.5         GigabitEthernet0/0/0/0.145
Global Neighbor: 10.100.106.1
  Interface Neighbor   Interface
  -------------------- ------------
  10.100.146.6         GigabitEthernet0/0/0/0.146
Global Neighbor: 10.101.124.1
  Interface Neighbor   Interface
  -------------------- ------------
  10.100.42.24         GigabitEthernet0/0/0/0.1424
```

TE: RSVP

```
RP/0/0/CPU0:P104#show rsvp interface
Mon Jun 10 08:38:12.129 EDT

*: RDM: Default I/F B/W % : 75% [default] (max resv/bc0), 0% [default] (bc1)

Interface    MaxBW (bps)  MaxFlow (bps) Allocated (bps)      MaxSub (bps)
------------ ------------ ------------- -------------------- -------------
Gi0/0/0/0.124         38G           38G           3G (  7%)            0
Gi0/0/0/0.134         38G           38G           2G (  5%)            0
tt10452                0             0            0 (  0%)            0
Gi0/0/0/0.145         38G           38G           0 (  0%)            0
Gi0/0/0/0.146         38G           38G           5G ( 13%)            0
tt10454                0             0            0 (  0%)            0
Gi0/0/0/0.1424       9500M         9500M          0 (  0%)            0
tt10456                0             0            0 (  0%)            0
tt10457                0             0            0 (  0%)            0
tt10459                0             0            0 (  0%)            0
tt10460                0             0            0 (  0%)            0
```

FRR backup tunnels (no BW) at headend

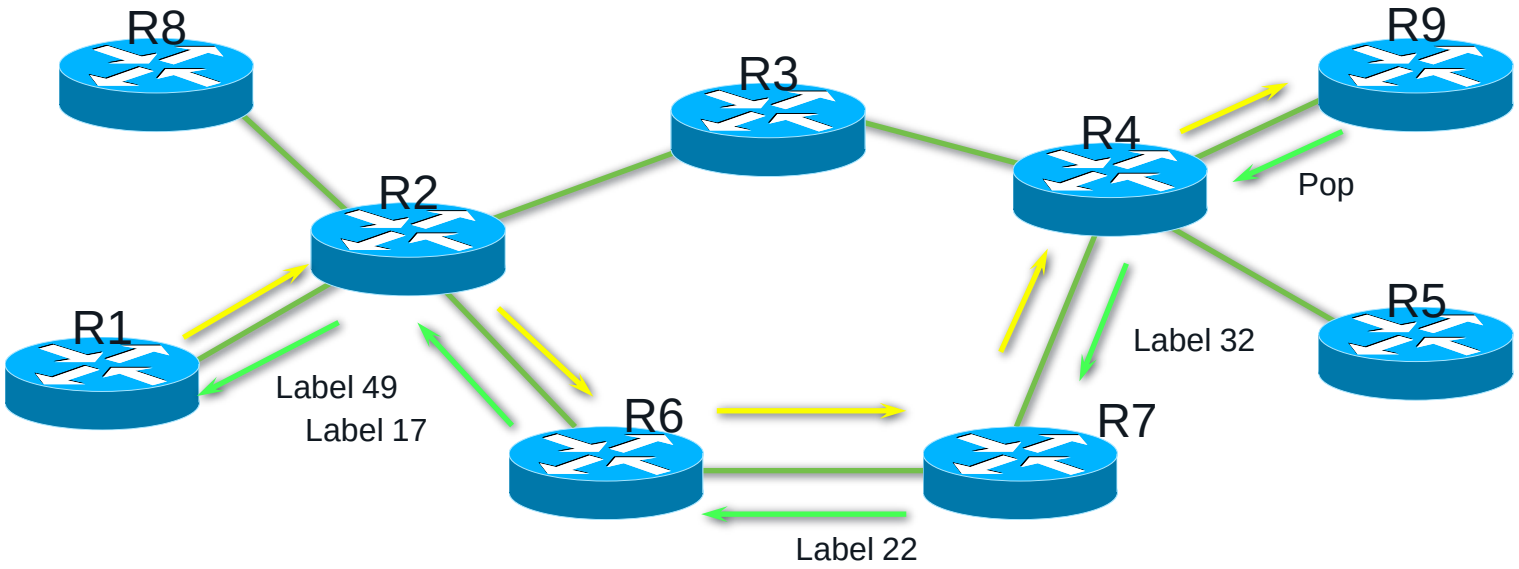
TE: Links

```
RP/0/0/CPU0:P104#show mpls traffic-eng link-management summary
Mon Jun 10 11:47:53.059 EDT

  System Information::
      Links Count          : 5 (Maximum Links Supported 500)
      Flooding System      : enabled
      IGP Areas Count      : 1

  IGP Areas
  ----------

  IGP Area[1]:: IS-IS ISIS level 2
      Flooding Protocol    : IS-IS
      Flooding Status      : flooded
      Periodic Flooding    : enabled (every 180 seconds)
      Flooded Links        : 5
      IGP System ID        : 0101.0010.4001
      MPLS TE Router ID    : 10.100.104.1
      IGP Neighbors        : 5
```

TE: Topology View [At Headend]

```
RP/0/0/CPU0:PE135#show mpls traffic-eng topology summary
Mon Jun 10 11:50:35.198 EDT
My_System_id: 0101.0113.5001.00 (IS-IS ISIS level-2)
My_BC_Model_Type: RDM

Signalling error holddown: 10 sec Global Link Generation 5292918

IS-IS ISIS  level 2
  Local System Id:          0101.0113.5001
  TE router ID configured:  10.101.135.1
               in use:      10.101.135.1

    IGP Id: 0101.0010.8001.00, MPLS TE Id: 10.100.108.1 Router Node
        4 links
.
    IGP Id: 0101.0111.2001.00, MPLS TE Id: 10.101.112.1 Router Node (Overloaded)
        2 links
.
    IGP Id: 0101.0010.8001.05, Network Node
        2 links
.
    Total: 43 nodes (16 router, 27 network), 108 links

Grand Total: 43 nodes (16 router, 27 network)  108 links
```

TE: The Label Exchange

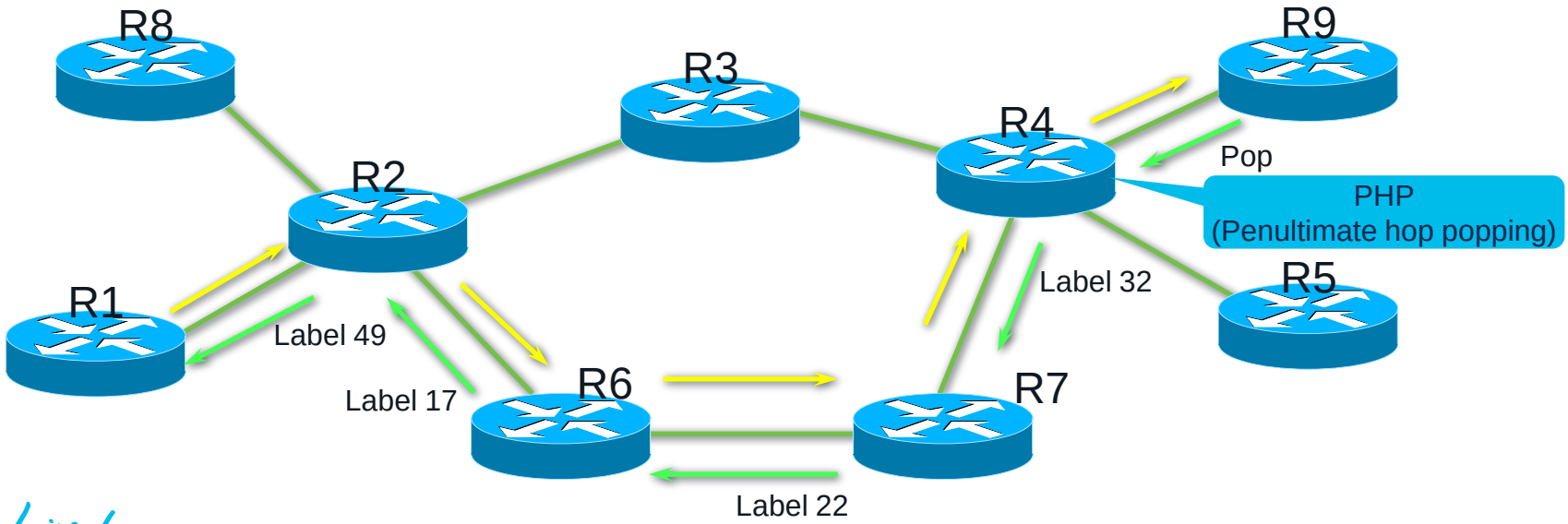


Setup: Path (ERO = R1->R2->R6->R7->R4->R9)

Reply: Resv communicates labels & reserves bandwidth on each link

TE: The Path [The Unidirectional Tunnel]

- CSPF calculation & signaling initiated by headend based on its MPLS TE topology view.
- Creates a P2P [or P2MP, MP2MP] unidirectional tunnel.
- Signaling and periodic refreshing of state done using RSVP.



TE: Control Plane View

```
RP/0/0/CPU0:P104#show mpls traffic-eng tunnels tabular
Mon Jun 10 14:48:32.737 EDT

            Tunnel    LSP      Destination         Source    Tun     FRR  LSP  Path
              Name     ID         Address        Address  State  State Role  Prot
----------------- ----- --------------- --------------- ------ ------ ---- -----
  *tunnel-te10452     7    10.100.108.1    10.100.104.1     up  Inact Head Inact
  *tunnel-te10454     2    10.101.124.1    10.100.104.1     up  Inact Head Inact
 .
  *tunnel-te10460     2    10.101.125.1    10.100.104.1     up  Inact Head Inact
        P101_t181    25    10.100.108.1    10.100.101.1     up  Inact  Mid
        P108_t811    25    10.100.101.1    10.100.108.1     up  Inact  Mid
autob_P101_t10150    24    10.101.135.1    10.100.101.1     up  Inact  Mid
autob_P101_t10152    25    10.101.137.1    10.100.101.1     up  Inact  Mid
.
autob_P108_t10868    13    10.100.104.1    10.100.108.1     up  Inact Tail
autom_PE112_t1121    16    10.101.188.1    10.101.112.1     up  Ready  Mid
autom_PE124_t1241    16    10.101.188.1    10.101.124.1     up  Ready  Mid
autob_PE124_t1245    27    10.100.103.1    10.101.124.1     up  Inact  Mid
.
autom_rasr9000-2w    11    10.101.135.1    10.101.188.1     up  Ready  Mid
autom_rasr9000-2w    11    10.101.137.1    10.101.188.1     up  Ready  Mid
* = automatically created backup tunnel
+ = automatically created mesh tunnel
```

FRR enabled but inactive

TE: Control Plane View

```
RP/0/0/CPU0:P104#show mpls traffic-eng tunnels brief
Mon Jun 10 15:03:38.965 EDT

                      TUNNEL NAME          DESTINATION      STATUS  STATE
                  *tunnel-te10452        10.100.108.1          up  up
                  *tunnel-te10454        10.101.124.1          up  up
                  *tunnel-te10456        10.100.101.1          up  up
                  *tunnel-te10457        10.101.135.1          up  up
                  *tunnel-te10459        10.101.137.1          up  up
                  *tunnel-te10460        10.101.125.1          up  up
                        P101_t181        10.100.108.1          up  up
                        P108_t811        10.100.101.1          up  up
               autob_P101_t10150_        10.101.135.1          up  up
               autob_P101_t10152_        10.101.137.1          up  up
.
               autom_rasr9000-2w-        10.101.112.1          up  up
               autom_rasr9000-2w-        10.101.124.1          up  up
               autom_rasr9000-2w-        10.101.125.1          up  up
               autom_rasr9000-2w-        10.101.135.1          up  up
               autom_rasr9000-2w-        10.101.137.1          up  up
* = automatically created backup tunnel
Displayed 6 (of 6) heads, 33 (of 33) midpoints, 6 (of 6) tails
Displayed 6 up, 0 down, 0 recovering, 0 recovered heads
```

TE: Control Plane View: Tunnel Headend

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show mpls traffic-eng tunnels 11106 detail
Tue Jun 11 00:04:29.172 EDT


Name: tunnel-te11106  Destination: 10.101.112.1 (auto-tunnel mesh)
  Status:
    Admin:    up Oper:   up   Path:  valid   Signalling: connected

    path option 10,  type dynamic  (Basis for Setup, path weight 22230)
    G-PID: 0x0800 (derived from egress interface properties)
    Bandwidth Requested: 1000000 kbps  CT0
    Creation Time: Fri May 31 16:37:30 2013 (1w3d ago)
  Config Parameters:
    Bandwidth:  1000000 kbps (CT0) Priority:  3  3 Affinity: 0x0/0xffff
    Metric Type: TE (default)
.
 Current LSP Info:
    Instance: 2, Signaling Area: IS-IS ISIS level-2
    Uptime: 1w3d (since Fri May 31 16:37:30 EDT 2013)
    Outgoing Interface: GigabitEthernet0/1/0/1, Outgoing Label: 101048
    Router-IDs: local      10.101.111.1
                downstream 10.100.101.1
    Soft Preemption: None
.
```

TE: Control Plane View: Tunnel Midpoint

```
RP/0/0/CPU0:P104#show mpls traffic-eng tunnels 10152
Mon Jun 10 15:01:13.715 EDT


LSP Tunnel 10.100.101.1 10152 [25] is signalled, connection is up
  Tunnel Name: autob_P101_t10152_Gi0_0_0_0.113_10.100.103.1 Tunnel Role: Mid
  InLabel: GigabitEthernet0/0/0/0.124, 104054
  OutLabel: GigabitEthernet0/0/0/0.145, 105016
  Signalling Info:
    Src 10.100.101.1 Dst 10.101.137.1, Tun ID 10152, Tun Inst 25, Ext ID 10.100.101.1
    Router-IDs: upstream   10.100.102.1
                local      10.100.104.1
                downstream 10.100.105.1
    Bandwidth: 0 kbps (CT0) Priority:  6  6 DSTE-class: no match
    Soft Preemption: None
    Path Info:
      Incoming Address: 10.100.124.4
      Incoming:
      Explicit Route:
        Strict, 10.100.124.4
        Strict, 10.100.145.4
        Strict, 10.100.145.5
        Strict, 10.100.157.5
        .
```

TE: Forwarding: Headend Forwarding Methods

- Auto-route announce [IGP]
- Forwarding adjacency [IGP]
- Policy-based (& class-based) routing
- Static routes
- Pseudo-wire tunnel selection

TE: Forwarding: Headend

```
RP/0/0/CPU0:PE135#show mpls forwarding tunnels
Tue Jun 11 07:44:26.151 EDT
Tunnel    Outgoing    Outgoing      Next Hop        Bytes
Name      Label       Interface                     Switched
--------- ----------- ------------ --------------- ------------
tt13501   103049      Gi0/0/0/0.1335 10.100.35.3     0
tt13502   103051      Gi0/0/0/0.1335 10.100.35.3     0
tt13503   103048      Gi0/0/0/0.1335 10.100.35.3     105560
tt13504   103047      Gi0/0/0/0.1335 10.100.35.3     2600
tt13505   103037      Gi0/0/0/0.1335 10.100.35.3     0
.

RP/0/0/CPU0:PE135#show mpls forwarding tunnels 13503 detail
Tue Jun 11 07:45:18.917 EDT
Tunnel    Outgoing    Outgoing      Next Hop        Bytes
Name      Label       Interface                     Switched
--------- ----------- ------------ --------------- ------------
tt13503   103048      Gi0/0/0/0.1335 10.100.35.3     105560
     Updated May 31 07:34:51.047
     Version: 401, Priority: 2
     MAC/Encaps: 18/22, MTU: 4456
     Label Stack (Top -> Bottom): { 103048 }
     Local Label: 135001
     Packets Switched: 1015
```

TE: Hardware Forwarding: Headend

```
RP/0/RSP0/CPU0:rasr9000-2w-b#show cef ipv4 10.101.137.1/32 hardware ingress location 0/1/CPU0
Tue Jun 11 02:36:41.139 EDT
10.101.137.1/32, version 183, internal 0x4000001 (ptr 0x885cd0d8) [1], 0x0 (0x87741040), 0x0
(0x0)
 Updated Jun 11 02:19:01.772
 Prefix Len 32, traffic index 0, precedence routine (0), priority 1
   via 10.101.137.1, tunnel-te18804, 3 dependencies, weight 0, class 0 [flags 0x0]
    path-idx 0 [0x8a77f2d8 0x0]
    next hop 10.101.137.1
    local adjacency
LEAF - HAL pd context :
 sub-type : IPV4, ecd_marked:0, has_collapsed_ldi:0, collapse_bwalk_required:0, ecdv2_marked:0
Leaf H/W Result:

    Physical Result: 0x11a00200 (LE)

    Raw Data0: 0x11850000 b9020000 00000000 00000000
    Raw Data1: 0x00000000 00000000 00200000 0000a2ff
    leaf_resolve_control_byte0
                    reserved: 0                              match: 1                    valid: 1
                 ifib_lookup: 0
              txadj_internal: 0
                      rec_fs: 0
.
```

TE: Hardware Forwarding: Headend – Cont.

```
TE-NH H/W Result for 1st NP:0 (index: 0x5 (BE)):

   Raw Data0: 0x5100002d e6311a5e b170115f 0000008a
   Raw Data1: 0x02000000 16000000 00000000 00000000
   cb0
                      spare: 0          default_action: 1
          backup_indication: 0                   match: 1
                       rsvd: 0                   valid: 1
   cb1
                   spare_cb: 0              tp_path_ss: 0
   te_nh_incomplete: 0
         tunnel_over_tunnel: 0                   spare: 0
   cb2
                      spare: 0        te_nh_incomplete: 0
                   spare_cb: 0
   TE_local_label:
                      label: 188003
                        exp: 0                     eos: 1
   TE_tunnel_label:
                      label: 108011
                        exp: 0                     eos: 1
        te_nh_stats_ptr: 0x70115f
   merge_point_label:
.
```

TE: Hardware Forwarding: Headend – Cont.

```
 RX H/W Result for 1st NP:0 (index: 0x16 (BE)):

    Raw Data0: 0x91000000 00000088 06000200 00000000
    adj_resolve_control_byte0
                        match: 1                    valid: 1
                   iptunl_adj: 0
                  remote_rack: 0

    adj_resolve_control_byte1
                    adj_down: 0                mgscp_en: 0
      rx_lag_hash_en: 0
          rx_lag_adj: 0

    adj_resolve_control_byte2
                  rx_lag_adj: 0            rx_adj_null0: 0
                 rp_destined: 0                 rx_punt: 0
                     rx_drop: 0
    sfp/vqi           : 0x88
    if_handle         : 0x6000200
.
RP/0/RSP0/CPU0:rasr9000-2w-b#show controllers pm location 0/1/CPU0 | begin 6000200
Tue Jun 11 05:19:29.503 EDT
Ifname(2): GigabitEthernet0_1_0_1, ifh: 0x6000200 :
.
```

Egress interface

TE: Forwarding: Midpoint

```
RP/0/0/CPU0:P104#show mpls forwarding
Tue Jun 11 07:04:03.137 EDT
Local  Outgoing    Prefix            Outgoing     Next Hop        Bytes
Label  Label       or ID             Interface                    Switched
------ ----------- ----------------- ------------ --------------- ------------
104003 Exp-Null-v4 10254             Gi0/0/0/0.1424 10.100.42.24    0
104004 Exp-Null-v4 10357             Gi0/0/0/0.1424 10.100.42.24    0
104005 Exp-Null-v4 12450             Gi0/0/0/0.134 10.100.134.3    0
104010 106032      10352             Gi0/0/0/0.146 10.100.146.6    0
.
104028 Exp-Null-v4 10750             Gi0/0/0/0.134 10.100.134.3    0
104042 106003      181               Gi0/0/0/0.146 10.100.146.6    10747212
104044 106004      12410             Gi0/0/0/0.146 10.100.146.6    0
104046 Exp-Null-v4 10260             Gi0/0/0/0.145 10.100.145.5    0
104047 105058      10257             Gi0/0/0/0.145 10.100.145.5    0
104048 105007      10259             Gi0/0/0/0.145 10.100.145.5    0
104049 106017      11210             Gi0/0/0/0.146 10.100.146.6    0
104050 Exp-Null-v4 10153             Gi0/0/0/0.145 10.100.145.5    684820
104052 105059      10150             Gi0/0/0/0.145 10.100.145.5    1108
104054 105016      10152             Gi0/0/0/0.145 10.100.145.5    1392
104059 Exp-Null-v4 10196             Gi0/0/0/0.1424 10.100.42.24    0
104061 102021      18801             Gi0/0/0/0.124 10.100.124.2    0
104062 102044      811               Gi0/0/0/0.124 10.100.124.2    1057172746
.
```

Tunnel ID

TE: Hardware Forwarding: Midpoint

```
RP/0/0/CPU0:P104#show mpls forwarding labels 104062 hardware ingress location 0/0/CPU0
Tue Jun 11 11:31:38.647 EDT
Local  Outgoing    Prefix            Outgoing     Next Hop        Bytes
Label  Label       or ID             Interface                    Switched
------ ----------- ----------------- ------------ --------------- ------------
104062 102044      811               Gi0/0/0/0.124 10.100.124.2     N/A
.
```

Labels in Labels: LDP in TE

```
RP/0/RSP0/CPU0:rasr9000-2w-a#traceroute 10.101.178.1 source 10.101.111.1
Tue Jun 11 20:49:15.907 EDT

Type escape sequence to abort.
Tracing the route to 10.101.178.1

 1  10.100.11.1 [MPLS: Label 101055 Ex                    2 msec
 2  10.100.11.1 [MPLS: Label 101055 Exp [illegible]  2 msec  3 msec  2 msec
 3  10.100.113.3 [MPLS: Label 103097 Exp 0] 1 msec  2 msec  2 msec
 4  10.100.135.5 [MPLS: Label 105060 Exp 0] 2 msec  2 msec  2 msec
 5  10.100.157.7 [MPLS: Label 107068 Exp 0] 2 msec  2 msec  2 msec
 6  10.100.78.78 2 msec  *  1 msec

RP/0/RSP0/CPU0:rasr9000-2w-a#traceroute 10.101.178.1 source 10.101.111.1
Tue Jun 11 20:56:32.972 EDT

Type escape sequence to abort.
Tracing the route to 10.101.178.1

 1  10.100.11.1 [MPLS: Labels 101055/0 Exp 0] 3 msec  3 msec  2 msec
 2  10.100.113.3 [MPLS: Labels 103097/0 Exp 0] 1 msec  2 msec  1 msec
 3  10.100.135.5 [MPLS: Labels 105060/0 Exp 0] 1 msec  2 msec  2 msec
 4  10.100.157.7 [MPLS: Labels 107068/0 Exp 0] 2 msec  2 msec  2 msec
 5  10.100.78.78 2 msec  *  3 msec
```

TE label

TE tunnel

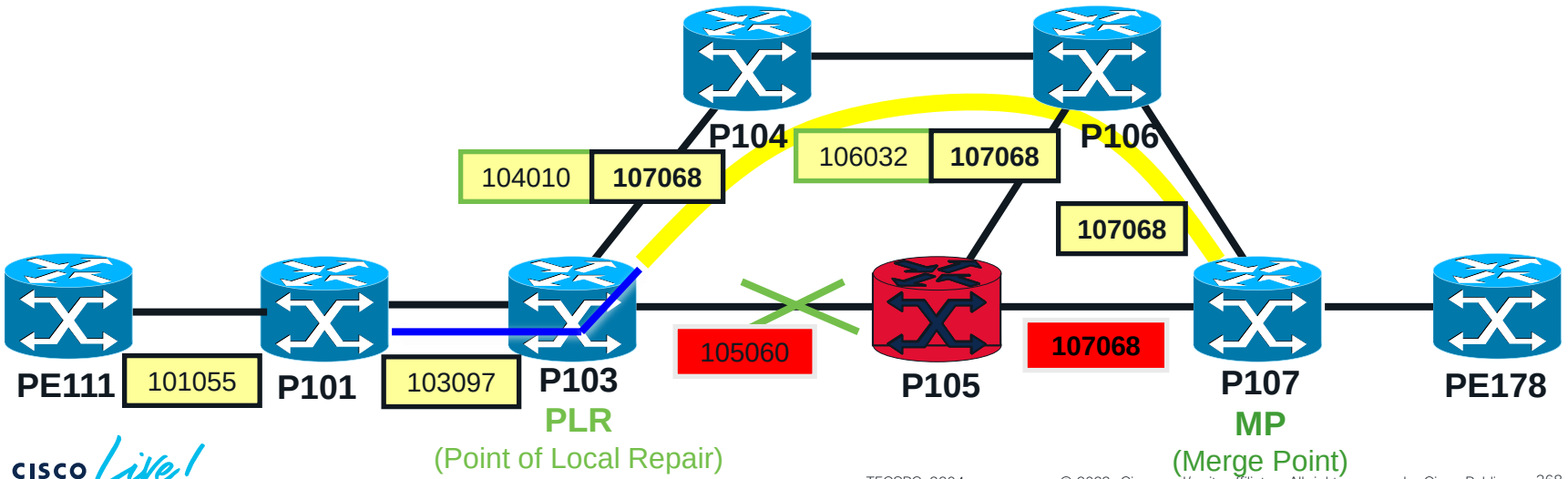
LDP explicit null

LDP in TE tunnel

Labels in Labels: LDP in TE with NNHOP FRR

```
RP/0/RSP0/CPU0:rasr9000-2w-a#traceroute 10.101.178.1 source 10.101.111.1
.
 1  10.100.11.1 [MPLS: Labels 101055/0 Exp 0] 3 msec  3 msec  3 msec
 2  10.100.113.3 [MPLS: Labels 103097/0 Exp 0] 3 msec  3 msec  2 msec
 3  10.100.134.4 [MPLS: Labels 104010/107068/0 Exp 0] 3 msec  2 msec  3 msec
 4  10.100.146.6 [MPLS: Labels 106032/107068/0 Exp 0] 4 msec  4 msec  2 msec
 5  10.100.167.7 [MPLS: Labels 0/107068/0 Exp 0] 3 msec  3 msec  3 msec
 6  10.100.78.78 3 msec  *  2 msec
```

LDP in TE tunnel and FRR active = 3 labels



Labels in Labels: LDP in TE at Headend

```
RP/0/RSP0/CPU0:PE111#show cef ipv4 10.101.178.1/32 hardware egress location 0/1/CPU0
Tue Jun 11 21:41:12.866 EDT
10.101.178.1/32, version 285, internal 0x4004001 (ptr 0x8854bcf8) [1], 0x0 (0x876e74a0), 0x450
(0x89cb6110)
 Updated Jun 11 21:12:35.330
 Prefix Len 32, traffic index 0, precedence routine (0), priority 1
   via 10.101.178.1, tunnel-te11111, 3 dependencies, weight 0, class 0 [flags 0x0]
    path-idx 0 [0x8a7a27d8 0x0]
    next hop 10.101.178.1
    local adjacency
     local label 111011      labels imposed {0}
.
 TE-NH H/W Result for 1st NP:0 (index: 0x34 (BE)):

    Raw Data0: 0x5100001b 1ac118ab f138145f 0000008a
    Raw Data1: 0x02000000 16000000 00000000 00000000
.
    TE_tunnel_label:
                        label: 101055
                          exp: 0                    eos: 1
         te_nh_stats_ptr: 0x38145f
    merge_point_label:
                        label: 0
.
```

LDP label

TE label [topmost]

Labels in Labels: TE NNHOP PLR

```
RP/0/0/CPU0:P103#show mpls forwarding labels 103097 hardware ingress detail
Tue Jun 11 12:51:07.075 EDT
Local  Outgoing    Prefix            Outgoing     Next Hop        Bytes
Label  Label       or ID             Interface                    Switched
------ ----------- ----------------- ------------ --------------- ------------
103097 105060      11111             Gi0/0/0/0.135 10.100.135.5    N/A
     Updated Jun 11 12:17:41.262
     Path Flags: 0x400 [ BKUP-IDX:1 (0xacde6f2c) ]
     Version: 1598, Priority: 2
     MAC/Encaps: 18/22, MTU: 4456
     Label Stack (Top -> Bottom): { 105052 }
     Packets Switched: 0
.
RP/0/0/CPU0:P103#show mpls forwarding labels 103097 hardware ingress detail
Tue Jun 11 12:51:31.414 EDT
Local  Outgoing    Prefix            Outgoing     Next Hop        Bytes
Label  Label       or ID             Interface                    Switched
------ ----------- ----------------- ------------ --------------- ------------
103097 107068      11111             tt10352      10.100.135.5    N/A
     Updated Jun 11 12:51:26.135
     Version: 1675, Priority: 2
     MAC/Encaps: 18/26, MTU: 4456
     Label Stack (Top -> Bottom): { 104010 107068 }
.
```

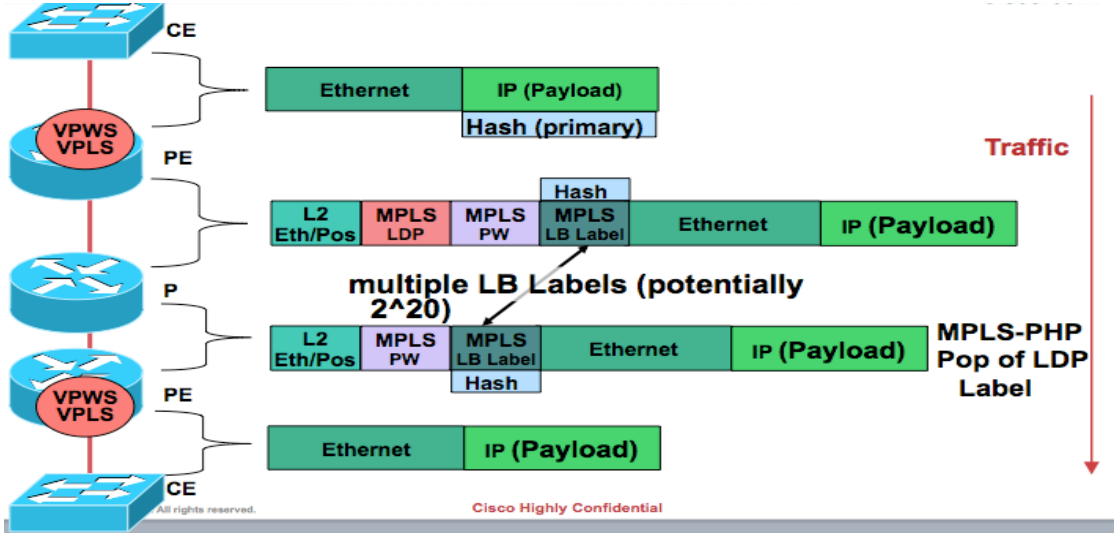
FRR Ready

FRR Active

FRR backup tunnel

Labels in Labels: More Forwarding Labels

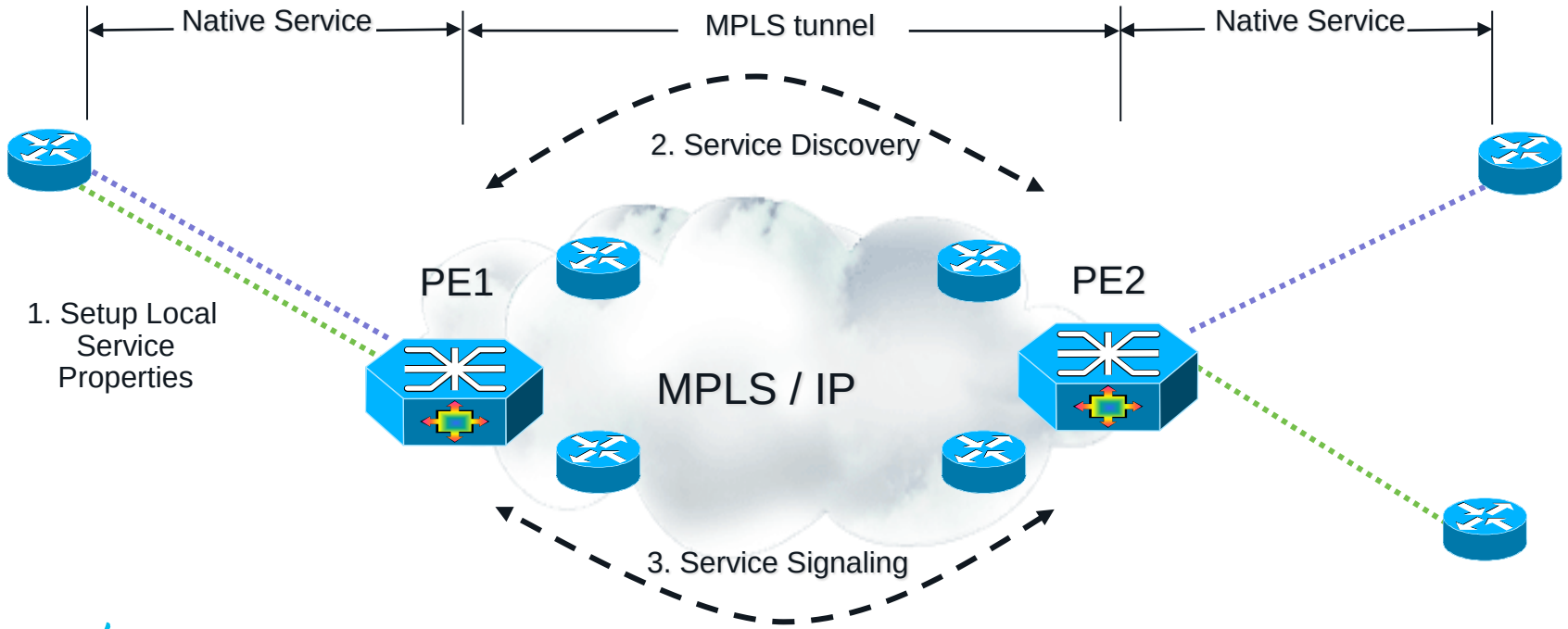
- Flow Label
  - Flow Aware Transport Pseudo-Wire [FAT PW] for VPWS and VPLS
  - http://www.cisco.com/en/US/partner/docs/routers/asr9000/software/asr9k_r4.3/lxvpn/configuration/guide/lesc43p2mps.html#wp1339194
  - Used for forwarding hashing, but it is at bottom of stack.



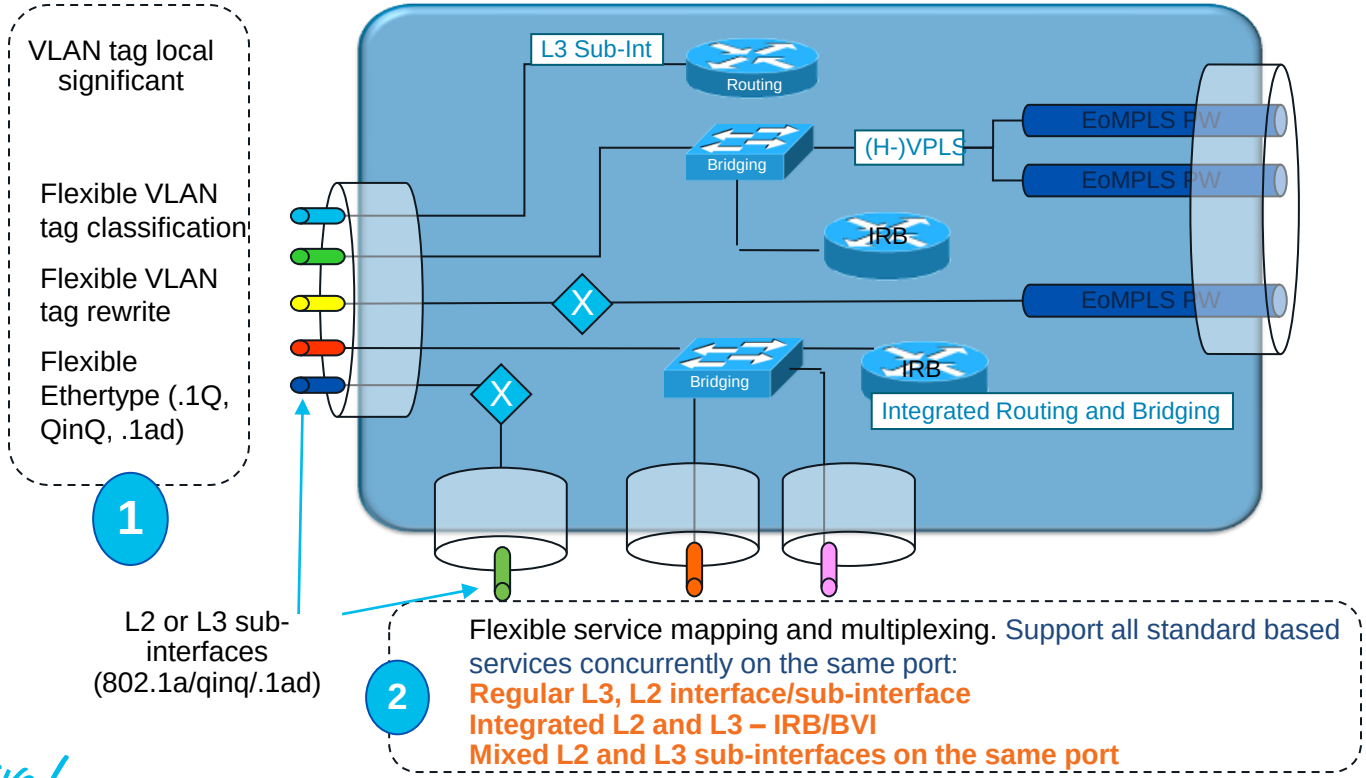
Service Labels

The Service Exchange

- Discovery: Manual or BGP
- Signaling: LDP or BGP



Service Attachment Points



LDP Signaling: PW Example

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show l2vpn xconnect detail
.
  PW: neighbor 10.101.188.1, PW ID 1, state is up ( established )
    PW class ONE, XC ID 0xc0000001
    Encapsulation MPLS, protocol LDP
    Source address 10.101.111.1
    PW type Ethernet, control word enabled, interworking none
    PW backup disable delay 0 sec
    Sequencing not set
    Load Balance Hashing: pw-label
    Flow Label flags configured (Tx=1,Rx=1), negotiated (Tx=1,Rx=1)

    PW Status TLV in use
      MPLS          Local                          Remote
      ------------ ------------------------------ ------------------------------
      Label         111014                         188014
      Group ID      0x6000180                      0x6000180
      Interface     GigabitEthernet0/1/0/3.1       GigabitEthernet0/1/0/3.1
      MTU           1504                           1504
      Control word  enabled                        enabled
      PW type       Ethernet                       Ethernet
      VCCV CV type  0x2                            0x2
                    (LSP ping verification)        (LSP ping verification)
.
```

Advertised

Received

Forwarding: AC to PW

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show l2vpn forwarding interface g0/1/0/3.1 hardware ingress detail
location 0/1/CPU0
Mon Jun 17 23:18:54.890 EDT
Local interface: GigabitEthernet0/1/0/3.1, Xconnect id: 0xc40001, Status: up
  Segment 1
    AC, GigabitEthernet0/1/0/3.1, status: Bound
    Statistics:
      packets: received 2809, sent 2810
      bytes: received 330634, sent 386882
      packets dropped: PLU 0, tail 0
      bytes dropped: PLU 0, tail 0
  Segment 2
    MPLS, Destination address: 10.101.188.1, pw-id: 1, status: Bound
    Pseudowire label: 188014    Control word enabled
    Load-Balance-Type: pw-label
    Flow Label flag: Tx=1
    Statistics:
      packets: received 2810, sent 2809
      bytes: received 386882, sent 330634
      packets dropped: PLU 0, tail 0, out of order 0
      bytes dropped: PLU 0, tail 0, out of order 0
  Platform AC context:
  Ingress AC: AToM, State: Bound
.
```

Forwarding: AC to PW – Cont.

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show l2vpn forwarding interface g0/1/0/3.1 hardware ingress detail
location 0/1/CPU0
.
  Platform AC context:
  Ingress AC: AToM, State: Bound
    Flags: Remote is PW
.

  Platform PW context:
  Ingress PW: AToM, State: Bound
  XID: 0xc0008000, bridge: 0, MAC limit: 0, l2vpn ldi index: 0x0002, vc label: 188014,
  nr_ldi_hash: 0x68, r_ldi_hash: 0xb3, lag_hash: 0xf4, SHG: None
    Flags: Control Word, Flow Label imposition
  NP0
    Xconnect ID: 0xc0008000, NP: 0
      Type: Pseudowire (with control word)
      Flags: Learn enable, Type 5, Local replication, Flow Label imposition
      VC label hash, nR-LDI Hash: 0x68, R-LDI Hash: 0xb6, LAG Hash: 0xf4,
      VC output label: 0x2de6e (188014), LDI: 0x0002, stats ptr: 0x00000000
      Split Horizon Group: None
.
```

Forwarding: PW to AC

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show mpls forwarding labels 111014 hardware ingress detail location
0/1/CPU0
Mon Jun 17 23:58:30.490 EDT
Local  Outgoing    Prefix            Outgoing     Next Hop        Bytes
Label  Label       or ID             Interface                    Switched
------ ----------- ----------------- ------------ --------------- ------------
111014 Pop         PW(10.101.188.1:1) Gi0/1/0/3.1  point2point     N/A
     Updated Jun 15 01:29:06.149
     Path Flags: 0x8 [ ]
     PW Flow Label: Enabled
     MAC/Encaps: 0/0, MTU: 0
     Label Stack (Top -> Bottom): { }
     Packets Switched: 0

LEAF - HAL pd context :
 sub-type : MPLS_VPWS, ecd_marked:0, has_collapsed_ldi:0, collapse_bwalk_required:0, ecdv2_mar
Leaf H/W Result:

Raw Data0: 0x51009400 01004004 00000000 00000000
Raw Data1: 0x00000000 00000000 00002013 5f000000
cb0
                vpn_special: 0          vc_label_vpws: 1
              vc_label_vpls: 0                   match: 1
.
```

BGP Signaling: L3 IPv4 VPN

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show bgp vpnv4 unicast labels
Tue Jun 18 01:25:27.965 EDT
BGP router identifier 10.101.111.1, local AS number 65001
BGP generic scan interval 60 secs
BGP table state: Active
Table ID: 0x0   RD version: 348768
BGP main routing table version 40
BGP scan interval 60 secs

Status codes: s suppressed, d damped, h history, * valid, > best
              i - internal, r RIB-failure, S stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network            Next Hop        Rcvd Label      Local Label
Route Distinguisher: 65001:1 (default for vrf CUST-A)
*> 172.20.200.0/24    0.0.0.0         nolabel         111015
*>i172.20.201.0/24    10.101.188.1    188015          nolabel
* i                   10.101.188.1    188015          nolabel
*> 172.20.210.0/24    172.20.200.2    nolabel         111016
*>i172.20.211.0/24    10.101.188.1    188016          nolabel
* i                   10.101.188.1    188016          nolabel

Processed 4 prefixes, 6 paths
```

Advertised

Received

Forwarding: IPv4 to VPNv4

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show cef vrf CUST-A ipv4 172.20.211.0/24 hardware ingress location
0/1/CPU0
Tue Jun 18 01:45:27.771 EDT
172.20.211.0/24, version 50, internal 0x14004001 (ptr 0x87935564) [1], 0x0 (0x0), 0x410
(0x89c84170)
 Updated Jun 18 01:25:21.070
 Prefix Len 24, traffic index 0, precedence routine (0), priority 3
   via 10.101.188.1, 5 dependencies, recursive [flags 0x6010]
    path-idx 0 [0x89d4cb84 0x0]
    next hop VRF - 'default', table - 0xe0000000
    next hop 10.101.188.1 via 111007/0/21
     next hop 10.100.101.1/32 tt1111       labels imposed {101000 188016}
LEAF - HAL pd context :
 sub-type : IPV4, ecd_marked:0, has_collapsed_ldi:0, collapse_bwalk_required:0, ecdv2_marked:0
Leaf H/W Result:

Physical Result: 0x11ba0200 (LE)

Raw Data0: 0x51924000 2de70100 00000000 00000000
Raw Data1: 0x0b000000 00000000 00180000 0000a2ff
leaf_resolve_control_byte0
                    reserved: 0                          match: 1                    valid: 1
                 ifib_lookup: 0
.
```

LDP label

VPNv4 label

Forwarding: VPNv4 to IPv4

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show mpls forwarding labels 111016 detail hardware ingress location
0/1/CPU0
Tue Jun 18 02:08:02.870 EDT
Local  Outgoing    Prefix            Outgoing     Next Hop        Bytes
Label  Label       or ID             Interface                    Switched
------ ----------- ----------------- ------------ --------------- ------------
111016 Unlabelled  172.20.210.0/24[V] Gi0/1/0/3.200 172.20.200.2    N/A
     Updated Jun 17 21:23:42.087
     Version: 47, Priority: 3
     MAC/Encaps: 18/18, MTU: 1500
     Label Stack (Top -> Bottom): { Unlabelled }
     Packets Switched: 0

LEAF - HAL pd context :
 sub-type : MPLS_VPN, ecd_marked:0, has_collapsed_ldi:0, collapse_bwalk_required:0, ecdv2_mark
Leaf H/W Result:

Raw Data0: 0x11020900 00000000 00000000 00000000
Raw Data1: 0x99000000 00000000 10000000 0000a2ff
cb0
                vpn_special: 0          vc_label_vpws: 0
              vc_label_vpls: 0                  match: 1
                       rsvd: 0                  valid: 1


.
```

Advertised local label

Forwarding: VPNv4 Aggregate to IPv4

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show mpls forwarding labels 111015 detail hardware ingress location
0/1/CPU0
Tue Jun 18 02:06:14.191 EDT
Local  Outgoing    Prefix            Outgoing     Next Hop        Bytes
Label  Label       or ID             Interface                    Switched
------ ----------- ----------------- ------------ --------------- ------------
111015 Aggregate   CUST-A: Per-VRF Aggr[V]   \
                                     CUST-A                       N/A
     Updated Jun 17 20:05:10.525
     Path Flags: 0x10 [ ]
     MAC/Encaps: 0/0, MTU: 0
     Label Stack (Top -> Bottom): { }
     Packets Switched: 0

LEAF - HAL pd context :
 sub-type : MPLS_DEAG, ecd_marked:0, has_collapsed_ldi:0, collapse_bwalk_required:0, ecdv2_mar
Leaf H/W Result:

Raw Data0: 0x910008ff 00000000 00000000 00000000
Raw Data1: 0x00000000 00000000 10000015 5f000000
cb0
               vpn_special: 1          vc_label_vpws: 0
             vc_label_vpls: 0                   match: 1
.
```

Advertised local label

Forwarding: Load Sharing To Core Bundle

```
RP/0/RSP0/CPU0:rasr9000-2w-b#show cef vrf TRAFFIC ipv4 10.10.6.6 hardware ingress location 0/0/CPU0
Tue Dec 10 10:14:29.843 EST
10.10.6.0/24, version 41, internal 0x14004001 (ptr 0x8dd2b964) [1], 0x0 (0x0), 0x410 (0x90d929b0)
 Updated Dec  5 08:06:31.568
 Prefix Len 24, traffic index 0, precedence routine (0), priority 3
   via 10.101.111.1, 7 dependencies, recursive [flags 0x6010]
    path-idx 0 [0x90e5ab08 0x0]
    next hop VRF - 'default', table - 0xe0000000
    next hop 10.101.111.1 via 188002/0/21
     next hop 0.0.0.0/32 tt180        labels imposed {ImplNull 111018}
LEAF - HAL pd context :
 sub-type : IPV4, ecd_marked:0, has_collapsed_ldi:0, collapse_bwalk_required:0, ecdv2_marked:0
Leaf H/W Result:

    Physical Result: 0x117a0300 (LE)
.

    Other fields:
                   leaf_ptr: 0xc4bc05(LE)            bgp_next_hop: 0xa656f01
                   urpf_ptr: 0
NextHopPrefix:label:eos=188002:0

Please use show cef  or show mpls forwarding command again
with nexthop prefix specified for nexthop hardware details
```

Forwarding: Load Sharing To Core Bundle

```
RP/0/RSP0/CPU0:rasr9000-2w-b#show cef ipv4 10.101.111.1 hardware ingress detail location 0/0/CPU0 |
include if_handle
Tue Dec 10 11:02:58.582 EST
    if_handle       : 0x2d320

RP/0/RSP0/CPU0:rasr9000-2w-b#show controllers pm location 0/0/CPU0 | include 2d320
Tue Dec 10 11:03:07.283 EST
Ifname(2): Bundle-Ether1, ifh: 0x2d320 :
parent_bundle_ifh  0x2d320

RP/0/RSP0/CPU0:rasr9000-2w-b#show bundle Bundle-Ether 1
Tue Dec 10 11:03:14.257 EST

Bundle-Ether1
  Status:                                  Up
  Local links <active/standby/configured>:   5 / 0 / 5
  Local bandwidth <effective/available>:     50000000 (50000000) kbps
  MAC address (source):                    10f3.110b.161b (Chassis pool)
.
  Port                Device           State        Port ID         B/W, kbps
  ------------------  ---------------  -----------  --------------  ----------
  Te0/0/0/5           Local            Active       0x8000, 0x0001    10000000      Link is Active
  Te0/1/1/0           Local            Active       0x8000, 0x0003    10000000      Link is Active
  Te0/1/1/1           Local            Active       0x8000, 0x0004    10000000      Link is Active
  Te0/1/1/2           Local            Active       0x8000, 0x0002    10000000      Link is Active
  Te0/1/1/3           Local            Active       0x8000, 0x0005    10000000      Link is Active
```

Forwarding: Load Sharing To Core Bundle

```
RP/0/RSP0/CPU0:rasr9000-2w-b#bundle-hash Bundle-Ether 1 location 0/0/CPU0
Tue Dec 10 11:17:15.224 EST
Calculate Bundle-Hash for L2 or L3 or sub-int based: 2/3/4 [3]:
Enter traffic type (1.IPv4-inbound, 2.MPLS-inbound, 3:IPv6-inbound): [1]:
Single SA/DA pair or range: S/R [S]:
Enter source IPv4 address [255.255.255.255]: 10.10.3.3
Enter destination IPv4 address [255.255.255.255]: 10.10.6.6
Compute destination address set for all members? [y/n]: n
Enter L4 protocol ID. (Enter 0 to skip L4 data) [0]:
Invalid protocol. L4 data skipped.
Link hashed [hash_val:3] to is TenGigE0/1/1/3 ICL () LON 4 ifh 0x6000680

Another? [y]: n

RP/0/RSP0/CPU0:rasr9000-2w-b#
```

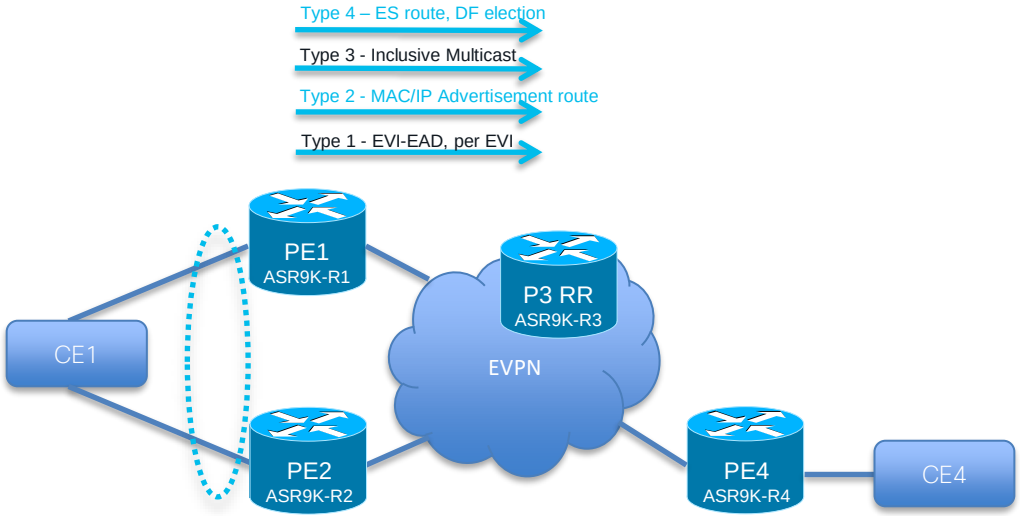
Forwarding: Load Sharing To Multiple CE's

```
RP/0/RSP0/CPU0:rasr9000-2w-b#show cef vrf TRAFFIC ipv4 10.10.3.3 hardware ingress location 0/1/CPU0
Tue Dec 10 12:13:20.504 EST
10.10.3.0/24, version 33, internal 0x4000001 (ptr 0x877d8564) [1], 0x0 (0x0), 0x0 (0x0)
 Updated Dec  5 08:06:32.256
 Prefix Len 24, traffic index 0, precedence routine (0), priority 3
   via 192.2.1.2, 3 dependencies, recursive [flags 0x0]
    path-idx 0 [0x877d8964 0x0]
    next hop 192.2.1.2 via 192.2.1.2/32
   via 192.3.1.2, 3 dependencies, recursive [flags 0x0]
    path-idx 1 [0x877d8f64 0x0]
    next hop 192.3.1.2 via 192.3.1.2/32
   via 192.4.1.2, 3 dependencies, recursive [flags 0x0]
    path-idx 2 [0x877d87e4 0x0]
    next hop 192.4.1.2 via 192.4.1.2/32
.
NextHopPrefix:192.2.1.2/32

Please use show cef  or show mpls forwarding command again
with nexthop prefix specified for nexthop hardware details
NextHopPrefix:192.3.1.2/32

Please use show cef  or show mpls forwarding command again
with nexthop prefix specified for nexthop hardware details
NextHopPrefix:192.4.1.2/32

Please use show cef  or show mpls forwarding command again
with nexthop prefix specified for nexthop hardware details
```

Forwarding: Load Sharing To Multiple CE's

```
RP/0/RSP0/CPU0:rasr9000-2w-b#show cef vrf TRAFFIC ipv4 exact-route 10.10.6.6 10.10.3.3 hardware
ingress location 0/0/CPU0
Tue Dec 10 12:17:36.699 EST
10.10.3.0/24, version 33, internal 0x4000001 (ptr 0x8dd31064) [1], 0x0 (0x0), 0x0 (0x0)
 Updated Dec  5 08:06:31.504
 local adjacency 192.2.1.2
 Prefix Len 24, traffic index 0, precedence routine (0), priority 3
   via TenGigE0/0/0/0
   via 192.2.1.2, 3 dependencies, recursive [flags 0x0]
    path-idx 0 [0x8dd29564 0x0]
    next hop 192.2.1.2 via 192.2.1.2/32

RP/0/RSP0/CPU0:rasr9000-2w-b#show cef adjacency tenGigE 0/0/0/0 192.2.1.2 hardware egress detail
location 0/0/CPU0
Tue Dec 10 12:23:14.902 EST
Display protocol is ipv4
Interface    Address                                          Type    Refcount

Te0/0/0/0    Prefix: 192.2.1.2/32                             local   5
             Adjacency: PT:0x8a7742e8 192.2.1.2/32
             Interface: Te0/0/0/0
             MAC: 02.c0.00.00.f3.10.02.01.90.61.05.11.00.00
             Interface Type: 0x1e, Base Flags: 0x1 (0x91c7ad58)
             Nhinfo PT: 0x91c7ad58, Idb PT: 0x8d18a318, If Handle: 0x40000c0
             Dependent adj type: remote (0x90fd7c70)
.
```

Use Case: EVPN Bridge-domain VPLS



EVPN biggest advantage over the conventional VPLS solutions is that MAC learning done via control plane, without sending BUM traffic in the core

On both PE1 and PE2:

```
interface GigabitEthernet0/0/0/0 l2transport
!
l2vpn
 bridge group 100
  bridge-domain 100
   interface GigabitEthernet0/0/0/0
 evi 100
!
evpn
 evi 100
  advertise-mac
 interface GigabitEthernet0/0/0/0
  ethernet-segment
   identifier type 0 36.37.00.00.00.00.00.11.00
!
router bgp 65000
 bgp router-id 1.1.1.1
 address-family l2vpn evpn
 neighbor-group rrC
  remote-as 65000
  update-source Loopback0
  address-family l2vpn evpn
 neighbor 3.3.3.3
  use neighbor-group rrC
```

EVPN single-active Ethernet Segment

```
RP/0/0/CPU0:PE1# show evpn ethernet-segment detail
Ethernet Segment Id      Interface                         Nexthops
------------------------ --------------------------------- --------------------
0036.3700.0000.0000.1100 Gi0/0/0/0                         1.1.1.1
                                                           2.2.2.2
  ES to BGP Gates   : Ready
  ES to L2FIB Gates : Ready
  Main port         :
     Interface name : GigabitEthernet0/0/0/0
     Interface MAC  : fa16.3e3d.fe16
     IfHandle       : 0x00000020
     State          : Up
     Redundancy     : Not Defined
  ESI type          : 0
     Value          : 36.3700.0000.0000.1100
  ES Import RT      : 3637.0000.0000 (from ESI)
  Source MAC        : 0000.0000.0000 (N/A)
  Topology          :
     Operational    : MH, Single-active
     Configured     : Single-active (AApS) (default)
  Service Carving   : Auto-selection
  Peering Details   :
     1.1.1.1 [MOD:P:00]
     2.2.2.2 [MOD:P:00]
  Service Carving Results:
     Forwarders     : 1
     Permanent      : 0
     Elected        : 1
     Not Elected    : 0
```

EVPN default Single-Active per Service (AApS)

EVPN MH elected forwarder

EVPN Ethernet Segment Advertisement

```
RP/0/0/CPU0:PE1# show evpn evi vpn-id 100 detail

VPN-ID     Encap  Bridge Domain                 Type
---------- ------ ----------------------------- -------------------
100        MPLS   100                           EVPN
   Stitching: Regular
   Unicast Label  : 24008
   Multicast Label: 24120
   Flow Label: N
   Control-Word: Enabled
   E-Tree: Root
   Forward-class: 0
   Advertise MACs: Yes
   Advertise BVI MACs: No
   Aliasing: Enabled
   UUF: Enabled
   Re-origination: Enabled
   Multicast source connected: No
   BGP Implicit Import: Enabled
   VRF Name:
…
   RD Config: none
   RD Auto  : (auto) 1.1.1.1:100
   RT Auto  : 65000:100
   Route Targets in Use           Type
   ------------------------------ --------------------
   65000:100                      Import
   65000:100                      Export
```

EVPN EVI 100 advertise-mac

EVPN RD-Auto = router-id + ESI

EVPN RT-Auto = BGP-AS + ESI

BGP L2VPN EVPN Flooding

```
RP/0/0/CPU0:PE1# show bgp l2vpn evpn

Status codes: s suppressed, d damped, h history, * valid, > best
              i - internal, r RIB-failure, S stale, N Nexthop-discard
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network            Next Hop            Metric LocPrf Weight Path
Route Distinguisher: 1.1.1.1:0 (default for vrf ES:GLOBAL)
*> [1][1.1.1.1:1][0036.3700.0000.0000.1100][4294967295]/184
                      0.0.0.0                                  0 i
*> [4][0036.3700.0000.0000.1100][32][1.1.1.1]/128
                      0.0.0.0                                  0 i
*>i[4][0036.3700.0000.0000.1100][32][2.2.2.2]/128
                      2.2.2.2                       100        0 i
Route Distinguisher: 1.1.1.1:100 (default for vrf 100)
*> [1][0036.3700.0000.0000.1100][0]/120
                      0.0.0.0                                  0 i
* i                   2.2.2.2                       100        0 i
*>i[1][0036.3700.0000.0000.1100][4294967295]/120
                      2.2.2.2                       100        0 i
*> [2][0][48][1022.1122.1122][0]/104
                      0.0.0.0                                  0 i
*>i[2][0][48][1022.2211.2211][0]/104
                      4.4.4.4                       100        0 i
*> [3][0][32][1.1.1.1]/80
                      0.0.0.0                                  0 i
*>i[3][0][32][2.2.2.2]/80
                      2.2.2.2                       100        0 i
*>i[3][0][32][4.4.4.4]/80
                      4.4.4.4                       100        0 i
```

PE1 learn and install MAC 1022.2211.2211 advertised from 4.4.4.4

BGP L2VPN EVPN MAC Learning Details

```
RP/0/0/CPU0:PE1# show bgp l2vpn evpn rd 1.1.1.1:100 [2][0][48][1022.2211.2211][0]/104

BGP routing table entry for [2][0][48][1022.2211.2211][0]/104, Route Distinguisher: 1.1.1.1:100
Versions:
  Process           bRIB/RIB  SendTblVer
  Speaker                 92          92
Last Modified: Dec 13 17:19:07.294 for 00:02:29
Paths: (1 available, best #1)
  Not advertised to any peer
  Path #1: Received by speaker 0
  Not advertised to any peer
  Local
    4.4.4.4 (metric 3) from 3.3.3.3 (4.4.4.4)
      Received Label 24007
      Origin IGP, localpref 100, valid, internal, best, group-best, import-candidate, imported, rib-
install
      Received Path ID 0, Local Path ID 1, version 92
      Extended community: SoO:4.4.4.4:100 RT:65000:100
      Originator: 4.4.4.4, Cluster list: 3.3.3.3
      EVPN ESI: 0000.0000.0000.0000.0000
      Source AFI: L2VPN EVPN, Source VRF: default, Source Route Distinguisher: 4.4.4.4:100
```

EVPN received label 24007

EVPN source RT: 65000:100 & RD: 4.4.4.4:100

EVPN Data Plane Local Forwarding:

```
RP/0/0/CPU0:PE1# show evpn evi vpn-id 100 mac detail
VPN-ID     Encap  MAC address    IP address                        Nexthop                          Label
---------- ------ -------------- --------------------------------- -------------------------------- ------
100        MPLS   1022.2211.2211 ::                                4.4.4.4                          24007
   Ethernet Tag                          : 0
   Multi-paths Resolved                  : True
   Multi-paths Internal label            : 24010

RP/0/0/CPU0:PE1#sh mpls forwarding labels 24010
Local  Outgoing    Prefix             Outgoing     Next Hop        Bytes
Label  Label       or ID              Interface                    Switched
------ ----------- ------------------ ------------ --------------- ------------
24010  24007       EVPN:100                        4.4.4.4         0

RP/0/0/CPU0:PE1#show l2vpn forwarding bridge-domain 100:100 hardware ingress detail loc 0/0/CPU0
Bridge-domain name: 100:100, id: 0, state: up
 MAC learning: enabled
 Software MAC learning: disabled
 MAC port down flush: enabled
 Flooding:
   Broadcast & Multicast: enabled
   Unknown unicast: enabled
 MAC aging time: 300 s, Type: inactivity
 MAC limit: 4000, Action: none, Notification: syslog
 MAC limit reached: no, threshold: 75%
 MAC Secure: disabled, Logging: disabled
 DHCPv4 snooping: profile not known on this node
 Dynamic ARP Inspection: disabled, Logging: disabled
 ...
```

EVPN install MAC 1022.2211.2211 from 4.4.4.4 with label 24010

Lots of information: All EFP's, all VPLS's, all MAC's!

EVPN Data Plane Remote Forwarding:

```
RP/0/0/CPU0:PE1# show evpn evi vpn-id 100 mac detail
VPN-ID     Encap  MAC address    IP address                        Nexthop                          Label
---------- ------ -------------- --------------------------------- -------------------------------- ------
100        MPLS   1022.1122.1122 ::                                1.1.1.1                          24008
   Ethernet Tag                          : 0
   Multi-paths Resolved                  : True
   Multi-paths Internal label            : 24009

RP/0/0/CPU0:PE1#sh mpls forwarding labels 24009
Local  Outgoing    Prefix             Outgoing     Next Hop        Bytes
Label  Label       or ID              Interface                    Switched
------ ----------- ------------------ ------------ --------------- ------------
24009  24008       EVPN:100                        1.1.1.1         0
       24008       EVPN:100                        2.2.2.2         0           (!)

RP/0/0/CPU0:PE4#show l2vpn forwarding bridge-domain 100:100 hardware ingress detail location 0/0/CPU0
Bridge-domain name: 100:100, id: 0, state: up
 MAC learning: enabled
 Software MAC learning: disabled
 MAC port down flush: enabled
 Flooding:
   Broadcast & Multicast: enabled
   Unknown unicast: enabled
 MAC aging time: 300 s, Type: inactivity
 MAC limit: 4000, Action: none, Notification: syslog
 MAC limit reached: no, threshold: 75%
 MAC Secure: disabled, Logging: disabled
 DHCPv4 snooping: profile not known on this node
...
```

EVPN install MAC 1022.1122.1122 from 1.1.1.1 with label 24009

EVPN MH backup

BGP Signaling: VPLS Bridge

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show bgp l2vpn vpls
Tue Jun 18 18:59:28.339 EDT
BGP router identifier 10.101.111.1, local AS number 65001
BGP generic scan interval 60 secs
BGP table state: Active
Table ID: 0x0   RD version: 412008
BGP main routing table version 35
BGP scan interval 60 secs

Status codes: s suppressed, d damped, h history, * valid, > best
              i - internal, r RIB-failure, S stale
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network            Next Hop        Rcvd Label      Local Label
Route Distinguisher: 65001:100 (default for vrf BRIDGES:BR-A)
*> 111:110/32         0.0.0.0         nolabel         111030
*> 111:130/32         0.0.0.0         nolabel         111090
*> 111:180/32         0.0.0.0         nolabel         111060
.
*>i188:110/32         10.101.188.1    188060          nolabel
* i                   10.101.188.1    188060          nolabel
*>i188:130/32         10.101.188.1    188090          nolabel
* i                   10.101.188.1    188090          nolabel
.
Processed 13 prefixes, 23 paths
```

Advertised

Received

Forwarding: VPLS Bridge: EFP to VFI

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show l2vpn forwarding bridge-domain BRIDGES:BR-A hardware ingress
detail location 0/1/CPU0
Tue Jun 18 21:18:34.152 EDT

Bridge-domain name: BRIDGES:BR-A, id: 0, state: up
 MAC learning: enabled
 MAC port down flush: enabled
 Flooding:
   Broadcast & Multicast: enabled
   Unknown unicast: enabled
 MAC aging time: 300 s, Type: inactivity
 MAC limit: 4000, Action: none, Notification: syslog
 MAC limit reached: no
 MAC Secure: disabled, Logging: disabled
 DHCPv4 snooping: profile not known on this node
 Dynamic ARP Inspection: disabled, Logging: disabled
 IP Source Guard: disabled, Logging: disabled
 IGMP snooping: disabled, flooding: enabled
 Bridge MTU: 1500 bytes
 Number of bridge ports: 5
 Number of MAC addresses: 2
 Multi-spanning tree instance: 0
  Platform bridge context:
.
```

Lots of information: All EFP's, all PW's, all labels!

Forwarding: VPLS Bridge: EFP to VFI

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show l2vpn forwarding bridge-domain BRIDGES:BR-A mac-address
hardware ingress detail location 0/1/CPU0
Tue Jun 18 21:10:27.472 EDT
 To Resynchronize MAC table from the Network Processors, use the command...
    l2vpn resynchronize forwarding mac-address-table location <r/s/i>

Mac Address    Type    Learned from/Filtered on    LC learned Resync Age          Mapped to
------------------------------------------------------------------------------------------
0022.9088.2ac0 dynamic Gi0/1/0/3.300              0/1/CPU0   0d 0h 0m 5s         N/A
0022.55e6.ae20 dynamic (10.101.188.1, 300)        0/1/CPU0   0d 0h 0m 4s         N/A
.
```

Dest MAC

PW

Forwarding: VPLS Bridge: EFP to VFI

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show l2vpn forwarding neighbor 10.101.188.1 300 hardware egress
detail location 0/1/CPU0
Tue Jun 18 21:21:47.126 EDT
Xconnect id: 0xc0000007, Status: up
  Segment 1
    MPLS, Destination address: 10.101.188.1, pw-id: 300, status: Bound
    Pseudowire label: 188061    Control word disabled
    Statistics:
      packets: received 376162, sent 376005
      bytes: received 51910302, sent 44368536
      packets dropped: PLU 0, tail 0, out of order 0
      bytes dropped: PLU 0, tail 0, out of order 0
  Segment 2
    Bridge id: 0, Split horizon group id: 1
    Storm control: disabled
    MAC learning: enabled
    MAC port down flush: enabled
    Flooding:
      Broadcast & Multicast: enabled
      Unknown unicast: enabled
    MAC aging time: 300 s, Type: inactivity
    MAC limit: 4000, Action: none, Notification: syslog
    MAC limit reached: no
.
```

PW label

Forwarding: VPLS Bridge: VFI to EFP

```
RP/0/RSP0/CPU0:rasr9000-2w-b#show mpls forwarding
Tue Jun 18 06:37:43.199 EDT
Local  Outgoing    Prefix             Outgoing     Next Hop        Bytes
Label  Label       or ID              Interface                    Switched
------ ----------- ------------------ ------------ --------------- ------------
188000 Pop         10.101.112.1/32    tt18800      10.101.112.1    0
188001 Pop         10.101.124.1/32    tt18801      10.101.124.1    0
188002 Pop         10.101.125.1/32    tt18802      10.101.125.1    300
188003 Pop         10.101.135.1/32    tt18803      10.101.135.1    0
188004 Pop         10.101.137.1/32    tt18804      10.101.137.1    300
188005 Pop         10.101.178.1/32    tt18805      10.101.178.1    0
188013 108009      10.101.111.1/32    tt8881       10.100.108.1    219343578
188014 Pop         PW(10.101.111.1:1) Gi0/1/0/3.1  point2point     319362
188015 Aggregate   CUST-A: Per-VRF Aggr[V]   \
                                      CUST-A                       49800
188016 Unlabelled  172.20.211.0/24[V] Gi0/1/0/3.200 172.20.201.2    0
188061 Pop         PW(10.101.111.1:300)   \
                                      BD=0         point2point     100172000
188075 Pop         PW(10.101.125.1:300)   \
                                      BD=0         point2point     0
188097 Pop         PW(10.101.137.1:300)   \
                                      BD=0         point2point     0
```

Pop label

Forwarding: VPLS Bridge: VFI to EFP

```
RP/0/RSP0/CPU0:rasr9000-2w-b#show l2vpn forwarding bridge-domain BRIDGES:BR-A mac-address
hardware egress location 0/1/CPU0
Tue Jun 18 06:44:04.464 EDT
 To Resynchronize MAC table from the Network Processors, use the command...
    l2vpn resynchronize forwarding mac-address-table location <r/s/i>

Mac Address    Type    Learned from/Filtered on    LC learned Resync Age          Mapped to
--------------------------------------------------------------------------------
0022.55e6.ae20 dynamic Gi0/1/0/3.300              0/1/CPU0   0d 0h 0m 4s         N/A
0022.9088.2ac0 dynamic (10.101.111.1, 300)        0/1/CPU0   0d 0h 0m 5s         N/A
.
```

Dest MAC

Agenda

- ✓ System Architecture: System anatomy & health
- ✓ Operating System & Configuration: IOS-XR & configuration models
- ✓ Control, Management, Security: Processing of control & exceptions
- ✓ Transit Packet/Frame Journey: Life of L3/L2 unicast/multicast
- ✓ MPLS Operation: Processing, forwarding & L3/L2 service operation
- ➢ Troubleshooting: Diagnostics, counters, drops, and packet capture

6 Troubleshooting

System Diagnostics

Background Diagnostics

RSP default diagnostics

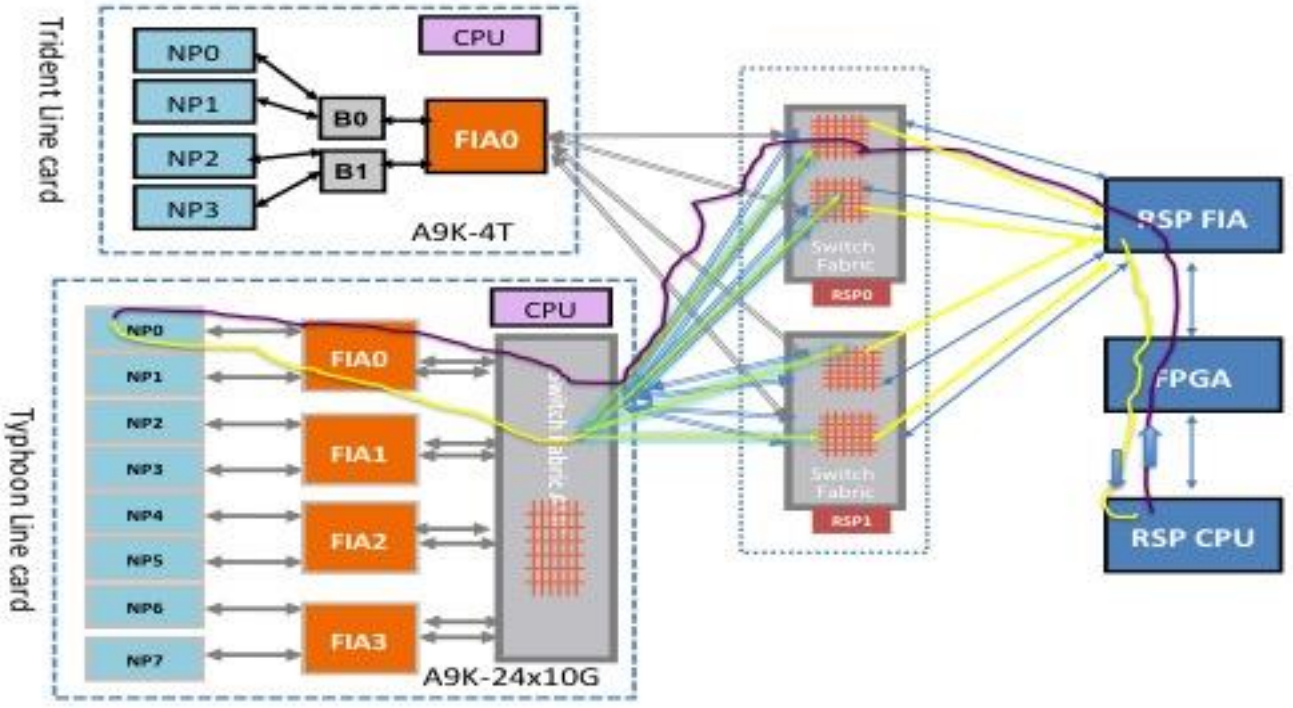
```
RP/0/RSP0/CPU0:rasr9000-2w-a#admin show diagnostic content location 0/RSP0/CPU0
Wed Dec 11 19:44:32.957 EST

RP 0/RSP0/CPU0:

  Diagnostics test suite attributes:
    M/C/* - Minimal bootup level test / Complete bootup level test / NA
    B/O/* - Basic ondemand test / not Ondemand test / NA
    P/V/* - Per port test / Per device test / NA
    D/N/* - Disruptive test / Non-disruptive test / NA
      S/* - Only applicable to standby unit / NA
      X/* - Not a health monitoring test / NA
      F/* - Fixed monitoring interval test / NA
      E/* - Always enabled monitoring test / NA
      A/I - Monitoring is active / Monitoring is inactive

                                                         Test Interval     Thre-
  ID   Test Name                            Attributes      (day hh:mm:ss.ms shold)
  ==== ==================================== ============ ================= =====
    1) CPUCtrlScratchRegister -----------> ***N****A       000 00:01:00.000 1
    2) ClkCtrlScratchRegister -----------> ***N****A       000 00:01:00.000 1
    3) ZenJfScratchRegister -------------> ***N****A       000 00:01:00.000 1
    4) FabSwitchIdRegister --------------> *B*N****A       000 00:01:00.000 1
    5) SrspStandbyEobcHeartbeat ---------> *B*NS***A       000 00:00:05.000 3
    6) SrspActiveEobcHeartbeat ----------> *B*NS***A       000 00:00:05.000 3
    7) FabricLoopback -------------------> MB*N****A       000 00:01:00.000 3
    8) PuntFabricDataPath ---------------> *B*N****A       000 00:01:00.000 3
```

Every minute or every 5 seconds

Error threshold (consecutive)

Background Diagnostics

Test example: PuntFabricDataPath



- Looping the path between RP CPU and each NP
- Active RP: unicast
- Standby RP: mcast

Background Diagnostics

LC default diagnostics

```
RP/0/RSP0/CPU0:rasr9000-2w-a#admin show diagnostic content location 0/0/CPU0
Wed Dec 11 20:32:08.842 EST

A9K-24x10GE-SE 0/0/CPU0:
  Diagnostics test suite attributes:
    M/C/* - Minimal bootup level test / Complete bootup level test / NA
    B/O/* - Basic ondemand test / not Ondemand test / NA
    P/V/* - Per port test / Per device test / NA
    D/N/* - Disruptive test / Non-disruptive test / NA
      S/* - Only applicable to standby unit / NA
      X/* - Not a health monitoring test / NA
      F/* - Fixed monitoring interval test / NA
      E/* - Always enabled monitoring test / NA
      A/I - Monitoring is active / Monitoring is inactive

                                                        Test Interval    Thre-
  ID   Test Name                                 Attributes     (day hh:mm:ss.ms shold)
  ==== ======================================== ============ ================= =====
    1) CPUCtrlScratchRegister ----------> *B*N****A       000 00:01:00.000 1
    2) PHYCtrlScratchRegister ----------> *B*N****A       000 00:01:00.000 1
    3) PortCtrlScratchRegister ---------> *B*N****A       000 00:01:00.000 1
    4) FIAScratchRegister --------------> *B*N****A       000 00:01:00.000 1
    5) LcEobcHeartbeat -----------------> *B*N****A       000 00:00:05.000 3
    6) NPULoopback ---------------------> *B*N****A       000 00:01:00.000 3
```

Background Diagnostics

Reading the results

```
RP/0/RSP0/CPU0:rasr9000-2w-a#admin show diagnostic result location
0/RSP1/CPU0 detail
Wed Dec 11 20:35:26.998 EST

Current bootup diagnostic level for RP 0/RSP1/CPU0: minimal
RP 0/RSP1/CPU0:
  Overall diagnostic result: PASS
  Diagnostic level at card bootup: minimal
  Test results: (. = Pass, F = Fail, U = Untested)

_____________________________________________________________________

  1  ) CPUCtrlScratchRegister ----------> .

          Error code ------------------> 0 (DIAG_SUCCESS)
          Total run count -------------> 31553
          Last test execution time ----> Wed Dec 11 20:35:08 2013
          First test failure time -----> n/a
          Last test failure time ------> n/a
          Last test pass time ---------> Wed Dec 11 20:35:08 2013
          Total failure count ---------> 0
          Consecutive failure count ---> 0

_____________________________________________________________________

  2  ) ClkCtrlScratchRegister ----------> .

          Error code ------------------> 0 (DIAG_SUCCESS)
          Total run count -------------> 31553
          Last test execution time ----> Wed Dec 11 20:35:08 2013
          First test failure time -----> n/a
          Last test failure time ------> n/a
          Last test pass time ---------> Wed Dec 11 20:35:08 2013
          Total failure count ---------> 0
          Consecutive failure count ---> 0
.
```

```
 6  ) SrspActiveEobcHeartbeat ---------> .

          Error code ------------------> 0 (DIAG_SUCCESS)
          Total run count -------------> 378621
          Last test execution time ----> Wed Dec 11 20:35:25 2013
          First test failure time -----> n/a
          Last test failure time ------> n/a
          Last test pass time ---------> Wed Dec 11 20:35:25 2013
          Total failure count ---------> 0
          Consecutive failure count ---> 0

_____________________________________________________________________

  7  ) FabricLoopback ------------------> .

          Error code ------------------> 0 (DIAG_SUCCESS)
          Total run count -------------> 31552
          Last test execution time ----> Wed Dec 11 20:35:08 2013
          First test failure time -----> n/a
          Last test failure time ------> n/a
          Last test pass time ---------> Wed Dec 11 20:35:08 2013
          Total failure count ---------> 0
          Consecutive failure count ---> 0

_____________________________________________________________________

  8  ) PuntFabricDataPath --------------> .

          Error code ------------------> 0 (DIAG_SUCCESS)
          Total run count -------------> 31552
          Last test execution time ----> Wed Dec 11 20:35:08 2013
          First test failure time -----> n/a
          Last test failure time ------> n/a
          Last test pass time ---------> Wed Dec 11 20:35:08 2013
          Total failure count ---------> 0
          Consecutive failure count ---> 0

_____________________________________________________________________
```

Background Diagnostics

Errors and clears

- Set: threshold exceeded
  - Path to LC 2 NP 0 failed in this example
- Clear: test previously failing, now passed
  - Indication of “transient” fault. Keep watching
- “show pfm location all” shows platform errors reported

```
RP/0/RSP0/CPU0:Feb 5 05:05:44.051 :
pfm_node_rp[354]:%PLATFORM-DIAGS-3-PUNT_FABRIC_DATA_PATH_FAILED :
Set|online_diag_rsp[237686]|System Punt/Fabric/data Path
Test(0x2000004)|failure threshold is 3, (slot, NP)failed:
(0/2/CPU0, 0)
```

```
RP/0/RSP0/CPU0:Feb 5 05:05:46.051 :
pfm_node_rp[354]:%PLATFORM-DIAGS-3-PUNT_FABRIC_DATA_PATH_FAILED :
Clear|online_diag_rsp[237686]|System Punt/Fabric/data Path
Test(0x2000004)|failure threshold is 3, (slot, NP)failed:
(0/2/CPU0, 0)
```

On Board Failure Logging [OBFL]

Logging errors, temperature, voltage locally on NVRAM

```
RP/0/RSP0/CPU0:rasr9000-2w-b#admin show logging onboard ?
  all           All Application
  cbc           CBC OBFL Commands
  detail        Onboard logging detail information
  diagnostic    Online Diagnostic Application
  environment   Environment Application
  error         Syslog Application
  location      locations to filter on
  raw           Onboard logging raw information
  summary       Onboard logging summary information
  temperature   Temperature Application
  trace         Debug traces for OBFL(cisco-support)
  uptime        Uptime Application
  verbose       Display internal debugging information
  voltage       Voltage Application
  |             Output Modifiers
  <cr>
RP/0/RSP0/CPU0:rasr9000-2w-b#admin show logging onboard error location 0/1/CPU0 | utility tail count 15
Tue Jan 21 23:32:30.047 EST
11/15/2013 19:27:50 sev:1 0/1/CPU0 pfm_node_lc[290]: %PLATFORM-CROSSBAR-1-SERDES_ERROR_LNK0 :
Set|fab_xbar[172110]|Crossbar Switch(0x1017010)|Slot_0_XBAR_1
12/19/2013 17:28:35 sev:1 0/1/CPU0 pfm_node_lc[291]: %PLATFORM-CROSSBAR-1-SERDES_ERROR_LNK0 :
Set|fab_xbar[172110]|Crossbar Switch(0x1017010)|Slot_0_XBAR_1
```

EOBC Switch. (32-bit XR)

Switch links and interfaces

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show controllers
backplane ethernet detail location 0/RSP0/CPU0
Wed Dec 11 21:53:1.24 EST
GigabitEthernet0_RSP0_CPU0 is up
  Active RSP is RSP 0

  Hardware is Gigabit Ethernet, H/W address is
564b.4700.0001
  Internet address is 127.0.1.0
  MTU 1514 bytes
  Encapsulation ASR9KIES (ASR9K Internal Ethernet
Server)
  Mode : Full Duplex,  Rate : 1Gb/s
    317863661 packets input, 4291854728 bytes, 0 total
input drops
    0 packets discarded (0 bytes) in garbage
collection
    300 packets discarded (83692 bytes) in recv
processing
    0 incomplete frames discarded
    0 packets discarded due to bad headers
    0 packets waiting for clients
    1 packets waiting on Rx
    Received 379557 broadcast packets, 17762716
multicast packets 0 dropped flood packets
    Input errors: 0 CRC, 0 overrun, 0 alignment, 0
length, 0 collision
    301873561 packets output, 2396666126 bytes, 0
total output drops
    Output 0 broadcast packets, 37045337 multicast
```

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show controllers backplane
ethernet detail location 0/0/CPU0
Wed Dec 11 21:53:00.797 EST
GigabitEthernet0_0_CPU0 is up
  Active RSP is RSP 0

  Hardware is Gigabit Ethernet, H/W address is
564b.4700.0821
  Internet address is 127.0.1.2
  MTU 1514 bytes
  Encapsulation ASR9KIES (ASR9K Internal Ethernet Server)
  Mode : Full Duplex,  Rate : 1Gb/s
    107127056 packets input, 707649689 bytes, 0 total
input drops
    12 packets discarded (792 bytes) in garbage
collection
    29433394 packets discarded (936722920 bytes) in recv
processing
    0 incomplete frames discarded
    0 packets discarded due to bad headers
    0 packets waiting for clients
    1 packets waiting on Rx
    Received 379530 broadcast packets, 53601523 multicast
packets 0 dropped flood packets
    Input errors: 0 CRC, 0 overrun, 0 alignment, 0
length, 0 collision
    67306802 packets output, 22632683 bytes, 0 total
output drops
    Output 1 broadcast packets, 1179989 multicast
.
```

EOBC Switch. (64-bit XR)

Switch links and interfaces

```
sysadmin-vm:0_RSP0# show controller switch summary location 0/RP0/RP-SW
Sun Jan  8  15:34:49.602 UTC+00:00
Rack  Card  Switch  Rack Serial Number
-------------------------------------
0     RP0   RP-SW

      Phys   Admin  Port      Protocol  Forward
Port  State  State  Speed     State     State       Connects To
-------------------------------------------------------------------------
4     Up     Up     10-Gbps   -         Forwarding  eth1 TPA
16    Up     Up     10-Gbps   -         Forwarding  CHA 6 (RP0 Punt)
17    Up     Up     10-Gbps   -         Forwarding  CHA 5 (RP0 Punt)
18    Up     Up     10-Gbps   -         Forwarding  CHA 7 (RP0 Punt)
19    Up     Up     10-Gbps   -         Forwarding  CHA 4 (RP0 Punt)
24    Up     Up     10-Gbps   -         Forwarding  RP0 CPU N1 P1
25    Up     Up     10-Gbps   -         Forwarding  RP0 CPU N1 P0
34    Up     Up     1-Gbps    -         Forwarding  (RP0 Ctrl)
35    Up     Up     1-Gbps    -         Forwarding  (RP1 Ctrl)
36    Up     Up     10-Gbps   -         Forwarding  CHA 0 (RP0 Punt)
37    Up     Up     10-Gbps   -         Forwarding  CHA 1 (RP0 Punt)
38    Up     Up     10-Gbps   -         Forwarding  CHA 2 (RP0 Punt)
39    Up     Up     10-Gbps   -         Forwarding  CHA 3 (RP0 Punt)
40    Up     Up     10-Gbps   -         Forwarding  LC1
42    Down   Up     10-Gbps   -         -           LC0
48    Up     Up     10-Gbps   -         Forwarding  LC3
49    Up     Up     10-Gbps   -         Forwarding  LC2
50    Up     Up     10-Gbps   -         Forwarding  LC5
51    Up     Up     10-Gbps   -         Forwarding  LC4
52    Down   Up     10-Gbps   -         -           LC6
53    Up     Up     10-Gbps   -         Forwarding  LC7
54    Up     Up     10-Gbps   -         Forwarding  RP1 Card (RP0 Ctrl)
55    Up     Up     10-Gbps   -         Forwarding  RP1 Card (RP1 Ctrl)

sysadmin-vm:0_RSP0#
```

LC & NP Resources

L3 forwarding resources

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show cef resource hardware ingress detail location 0/0/CPU0
Wed Dec 11 22:27:57.319 EST
CEF resource availability summary state: GREEN
CEF will work normally
  ipv4 shared memory resource:
        CurrMode GREEN, CurrAvail 1609412608 bytes, MaxAvail 1683308544 bytes
  ipv6 shared memory resource:
        CurrMode GREEN, CurrAvail 1609412608 bytes, MaxAvail 1683308544 bytes
  mpls shared memory resource:
        CurrMode GREEN, CurrAvail 1609412608 bytes, MaxAvail 1683308544 bytes
  common shared memory resource:
        CurrMode GREEN, CurrAvail 1609412608 bytes, MaxAvail 1683308544 bytes
  DATA_TYPE_TABLE_SET hardware resource: GREEN
  DATA_TYPE_TABLE hardware resource: GREEN
  DATA_TYPE_IDB hardware resource: GREEN
  DATA_TYPE_IDB_EXT hardware resource: GREEN
  DATA_TYPE_LEAF hardware resource: GREEN
  DATA_TYPE_LOADINFO hardware resource: GREEN
  DATA_TYPE_PATH_LIST hardware resource: GREEN
  DATA_TYPE_NHINFO hardware resource: GREEN
  DATA_TYPE_LABEL_INFO hardware resource: GREEN
  DATA_TYPE_FRR_NHINFO hardware resource: GREEN
  DATA_TYPE_ECD hardware resource: GREEN
.
```

LC & NP Resources

L2 service resources

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show l2vpn forwarding resource hardware ingress detail location
0/0/CPU0
Wed Dec 11 22:26:07.070 EST
L2VPN forwarding resource availability summary state: GREEN
  shared memory resource:
        CurrMode GREEN, CurrAvail 1609412608 bytes, MaxAvail 1650212864 bytes
  AC hardware resource: GREEN
  MPLS hardware resource: GREEN
  PBB PORT hardware resource: GREEN
  NHOP hardware resource: GREEN
  L2TP hardware resource: GREEN
  L2TP_SESSION hardware resource: GREEN
  VFI hardware resource: GREEN
  BRIDGE hardware resource: GREEN
  BRIDGE SHG hardware resource: GREEN
  BRIDGE PORT hardware resource: GREEN
  BRIDGE MAC hardware resource: GREEN
  MSTI MAIN PORT hardware resource: GREEN
  BRIDGE MAIN PORT hardware resource: GREEN
  MCAST TABLE hardware resource: GREEN
  MCAST LEAF hardware resource: GREEN
  MCAST XID hardware resource: GREEN
  PBB BMAC SA hardware resource: GREEN
.
```

LC & NP Resources

L2 service resources - continued

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show l2vpn forwarding resource hardware ingress detail location 0/0/CPU0
  Platform resource:
                    Hardware Resource Summary - NP0
                ------------------------------------------
                    Set       Modify       Clear        Get      Get Info    Delete All      In use
                    T/F         T/F          T/F        T/F         T/F          T/F          T/Max
            XID    24/0       117/0         8/0        0/0         0/0          0/0         16/262144
         BRIDGE     1/0        15/0         0/0        0/0         0/0          0/0          1/65536
     STATIC MAC     1/0         0/0         0/0        0/0         0/0          0/0          1/2097152
    BRIDGE PORT     0/0         0/0         0/0        0/0         0/0          0/0          0/393216
           UIDB    39/0         0/0         3/0        0/0         0/0          0/0          -/-
  VPLS PW STATS     7/0         -/-         4/0        -/-         0/0          -/-          -/-
           ISID     0/0         0/0         0/0        0/0         0/0          0/0          0/65536
           L2TP     0/0         0/0         0/0        0/0         0/0          0/0          0/131072
           DHCP     0/0         0/0         0/0        0/0         0/0          0/0          0/131072
      TOTAL MAC     -/-         -/-         -/-        -/-         -/-          -/-          1/2097152
                ----------- ----------- ----------- ----------- ----------- -----------
          Total    72/0       132/0        15/0        0/0         0/0          0/0

                    Hardware Performance Summary
                ------------------------------------------
            XID     < 1 ms     < 1 ms     < 1 ms     < 1 ms     < 1 ms     < 1 ms
         BRIDGE   000.001 s    < 1 ms     < 1 ms     < 1 ms     < 1 ms     < 1 ms
     STATIC MAC     < 1 ms     < 1 ms     < 1 ms     < 1 ms     < 1 ms     < 1 ms
```

TCP Resources

TCP connections states

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show tcp brief
Fri Dec 13 22:29:33.189 EST
   PCB      VRF-ID     Recv-Q Send-Q Local Address          Foreign Address        State
0x10174ad8 0x60000000      0      0  :::179                :::0                   LISTEN
0x101745f8 0x60000001      0      0  :::179                :::0                   LISTEN
0x10174798 0x60000002      0      0  :::179                :::0                   LISTEN
0x10174938 0x60000003      0      0  :::179                :::0                   LISTEN
0x101cce2c 0x6000000b      0      0  :::179                :::0                   LISTEN
0x1016fefc 0x00000000      0      0  :::179                :::0                   LISTEN
0x101ac7ac 0x60000000      0      0  10.101.111.1:179      10.100.101.1:20100     ESTAB
0x1017bbf4 0x60000000      0      0  10.101.111.1:646      10.101.188.1:30687     ESTAB
0x10182b38 0x60000000      0      0  10.101.111.1:179      10.100.103.1:59214     ESTAB
0x1002e004 0x6000000d      0      0  10.100.111.1:17514    10.100.111.100:13680   ESTAB
0x10161e18 0x60000000      0      0  0.0.0.0:23            0.0.0.0:0              LISTEN
0x101cc968 0x00000000      0      0  0.0.0.0:23            0.0.0.0:0              LISTEN
0x101594b4 0x60000000      0      0  0.0.0.0:646           0.0.0.0:0              LISTEN
0x1016f7f0 0x60000000      0      0  0.0.0.0:179           0.0.0.0:0              LISTEN
0x10165d74 0x60000001      0      0  0.0.0.0:179           0.0.0.0:0              LISTEN
0x1016f4b0 0x60000002      0      0  0.0.0.0:179           0.0.0.0:0              LISTEN
0x1016f650 0x60000003      0      0  0.0.0.0:179           0.0.0.0:0              LISTEN
0x101ade54 0x6000000b      0      0  0.0.0.0:179           0.0.0.0:0              LISTEN
0x1016276c 0x00000000      0      0  0.0.0.0:179           0.0.0.0:0              LISTEN
0x1015e304 0x00000000      0      0  0.0.0.0:0             0.0.0.0:0              CLOSED
```

TCP Resources

TCP connections parameters

```
RP/0/RSP0/CPU0:rasr9000-2w-a#show tcp detail pcb 0x10182b38
Wed Dec 11 22:47:18.708 EST

------------------------------------------------------------
Connection state is ESTAB, I/O status: 0, socket status: 0
Established at Thu Dec  5 04:13:28 2013

PCB 0x10182b38, SO 0x10182968, TCPCB 0x101cc2b4, vrfid
0x60000000,
Pak Prio: Medium, TOS: 192, TTL: 255, Hash index: 717
Local host: 10.101.111.1, Local port: 179 (Local App PID:
287053)
Foreign host: 10.100.103.1, Foreign port: 59214

Current send queue size in bytes: 0 (max 24576)
Current receive queue size in bytes: 0 (max 32768)  mis-
ordered: 0 bytes
Current receive queue size in packets: 0 (max 0)

Timer          Starts    Wakeups         Next(msec)
Retrans          9789          0                 0
SendWnd             0          0                 0
TimeWait            0          0                 0
AckHold          9817       9605                 0
KeepAlive           1          0                 0
PmtuAger            0          0                 0
GiveUp              0          0                 0
Throttle            0          0                 0

   iss: 2038437204  snduna: 2038628753  sndnxt: 2038628753
sndmax: 2038628753  sndwnd: 31856       sndcwnd: 3648
   irs: 2495655735  rcvnxt: 2495854116  rcvwnd: 31894
rcvadv: 2495886010
```

```
SRTT: 217 ms,  RTTO: 300 ms,  RTV: 11 ms,  KRTT: 0 ms
minRTT: 1 ms,  maxRTT: 289 ms

ACK hold time: 200 ms, Keepalive time: 0 sec, SYN waittime: 30
sec
Giveup time: 0 ms, Retransmission retries: 0, Retransmit
forever: FALSE
Connect retries remaining: 0, connect retry interval: 0 secs

State flags: none
Feature flags: MD5, Win Scale, Nagle
Request flags: Win Scale

Datagrams (in bytes): MSS 1216, peer MSS 1216, min MSS 1240,
max MSS 1240

Window scales: rcv 0, snd 0, request rcv 0, request snd 0
Timestamp option: recent 0, recent age 0, last ACK sent 0
Sack blocks {start, end}: none
Sack holes {start, end, dups, rxmit}: none

Socket options: SO_REUSEADDR, SO_REUSEPORT, SO_NBIO
Socket states: SS_ISCONNECTED, SS_PRIV
Socket receive buffer states: SB_DEL_WAKEUP
Socket send buffer states: SB_DEL_WAKEUP
Socket receive buffer: Low/High watermark 1/32768
Socket send buffer   : Low/High watermark 2048/24576, Notify
threshold 0

PDU information:
 #PDU's in buffer: 0
FIB Lookup Cache:  IFH: 0x134e0  PD ctx: size: 8  data: 0x0
0xb1494a74
  Num Labels: 0  Label Stack:
```

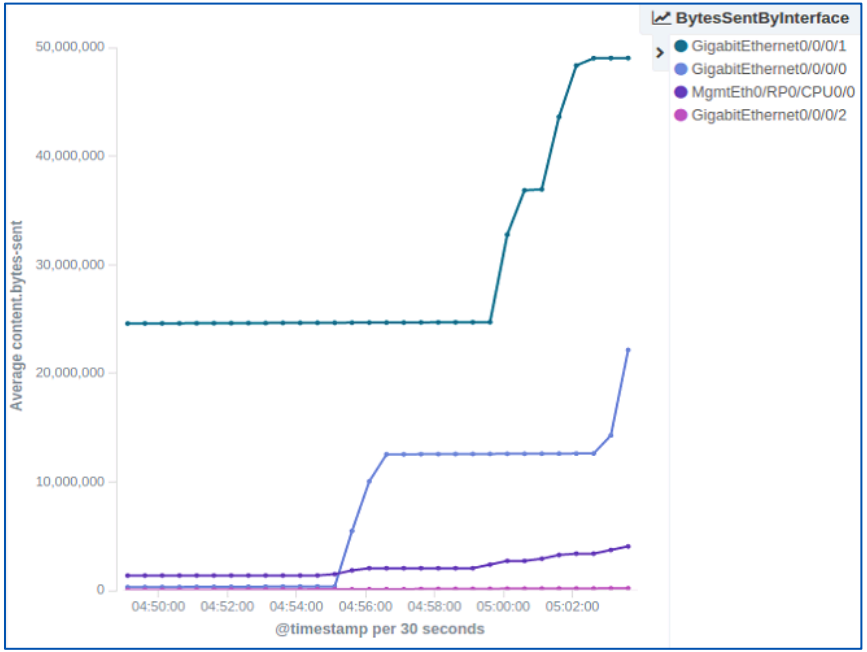
Troubleshooting: Forwarding Path

Telemetry

Streaming YANG model telemetry (push subscription model)

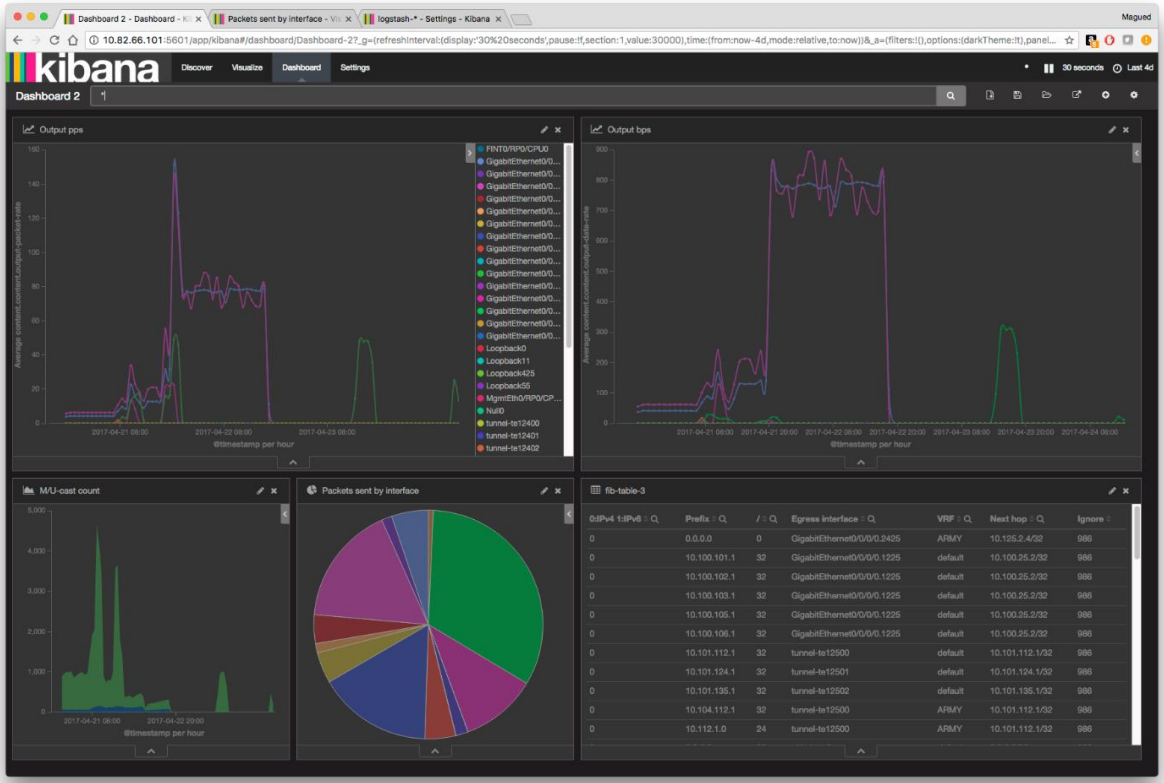
```
RP/0/RP0/CPU0:PE125#show running-config telemetry
model-driven
Tue Feb 14 13:06:18.261 UTC
telemetry model-driven
 destination-group COLL1
  address family ipv4 192.168.30.101 port 2103
   encoding self-describing-gpb
   protocol tcp
  !
 !
 sensor-group YD1
  sensor-path Cisco-IOS-XR-infra-statsd-oper:infra-
statistics/interfaces/interface/latest/generic-counters
 !
 subscription SUB1
  sensor-group-id YD1 sample-interval 60000
  destination-id COLL1
 !
!
```

YANG sub-tree



Telemetry

Build your own dashboard



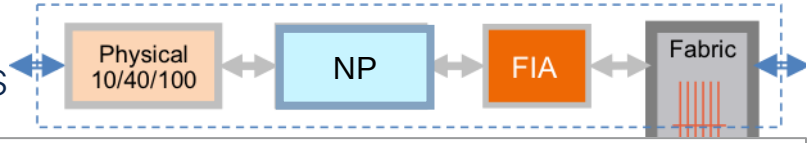
Monitor Interface

See interface stats in almost real time

```
RP/0/RSP0/CPU0:rasr9000-2w-b#monitor interface tenGigE 0/1/1/1

rasr9000-2w-b        Monitor Time: 00:00:22          SysUptime: 501:59:18

TenGigE0/1/1/1 is up, line protocol is up
Encapsulation ARPA

Traffic Stats:(2 second rates)                                 Delta
  Input  Packets:              2495245669613                  14890408
  Input  pps:                          7441113
  Input  Bytes:              164703177204108                 982758522
  Input  Kbps (rate):                  3928857                   ( 39%)
  Output Packets:              3017277633655                  13261227
  Output pps:                          6626897
  Output Bytes:              205177835436607                 901762428
  Output Kbps (rate):                  3605031                   ( 36%)

Errors Stats:
  Input  Total:                              1                       0
  Input  CRC:                                0                       0
  Input  Frame:                              0                       0
  Input  Overrun:                            0                       0
  Output Total:                              0                       0
  Output Underrun:                           0                       0

Quit='q', Freeze='f', Thaw='t', Clear='c', Interface='i',
Next='n', Prev='p'

Brief='b', Detail='d', Protocol(IPv4/IPv6)='r'
```

The Physical

Checking on port physical: SFP/XFP, levels




```
RP/0/RSP0/CPU0:rasr9000-2w-b#show controllers TenGigE 0/0/0/0 phy
Mon Dec  9 13:53:37.848 EST

SFP EEPROM  port: 0
        Xcvr Type: SFP
        Xcvr Code: SFP-10G-SR
        Encoding: 64B66B
        Bit Rate: 10300 Mbps
        Link Reach 50u fiber: 80 meter
        Link Reach 62.5u fiber: 20 meter
        Vendor Name: CISCO-FINISAR
        Vendor OUI: 00.90.65
        Vendor Part Number: FTLX8571D3BCL-C2 (rev.: A   )
        Laser wavelength: 850 nm (fraction: 0.00 nm)
        Optional SFP Signal: Rate Sel, LOS
        Vendor Serial Number: FNS164018G7
        Date Code (yy/mm/dd): 12/10/06  lot code:

        Thresholds:                    Alarm High          Warning High          Warning Low           Alarm Low
              Temperature:            +75.000 C             +70.000 C             +0.000 C             -5.000 C
                  Voltage:           3.630 Volt            3.465 Volt            3.135 Volt           2.970 Volt
                     Bias:          11.800 mAmps          10.800 mAmps           5.000 mAmps          4.000 mAmps
           Transmit Power:  1.479 mW (1.70 dBm)   0.741 mW (-1.30 dBm)  0.186 mW (-7.30 dBm)  0.074 mW (-11.30 dBm)
            Receive Power:  1.585 mW (2.00 dBm)   0.794 mW (-1.00 dBm)  0.102 mW (-9.90 dBm)  0.041 mW (-13.90 dBm)
        Temperature: 26.684
        Voltage: 3.301 Volt
        Tx Bias: 7.612 mAmps
        Tx Power:  0.613 mW (-2.13 dBm)
        Rx Power:  0.567 mW (-2.46 dBm)
        Oper. Status/Control:
```

The Physical

Reading the controller counters: In, out, invalid, unicast, mcast, frame sizes

```
RP/0/RSP0/CPU0:rasr9k-1y#show controllers TenGigE0/4/0/20
stats
Sun Feb 24 14:44:18.899 UTC
Statistics for interface TenGigE0/4/0/20 (cached values):
Ingress:
    Input total bytes          = 3081227904920
    Input good bytes           = 3081227904920

    Input total packets        = 23220024479
    Input 802.1Q frames        = 0
    Input pause frames         = 0
    Input pkts 64 bytes        = 7143534733
    Input pkts 65-127 bytes    = 2888766549
    Input pkts 128-255 bytes   = 13124923916
    Input pkts 256-511 bytes   = 62799261
    Input pkts 512-1023 bytes  = 0
    Input pkts 1024-1518 bytes = 0
    Input pkts 1519-Max bytes  = 0

    Input good pkts            = 23220024479
    Input unicast pkts         = 23220023458
    Input multicast pkts       = 62
    Input broadcast pkts       = 959

    Input drop overrun         = 0
.
```

```
Egress:
    Output total bytes         = 1345771624
    Output good bytes          = 1345771624

    Output total packets       = 21895707
    Output 802.1Q frames       = 0
    Output pause frames        = 0
    Output pkts 64 bytes       = 21665536
    Output pkts 65-127 bytes   = 21179
    Output pkts 128-255 bytes  = 168767
    Output pkts 256-511 bytes  = 40225
    Output pkts 512-1023 bytes = 0
    Output pkts 1024-1518 bytes = 0
    Output pkts 1519-Max bytes = 0

    Output good pkts           = 21895707
    Output unicast pkts        = 21870499
    Output multicast pkts      = 25195
    Output broadcast pkts      = 13
  Output drop underrun         = 0
    Output drop abort          = 0
    Output drop other          = 0

    Output error other         = 0
```

Interface Programming in Hardware

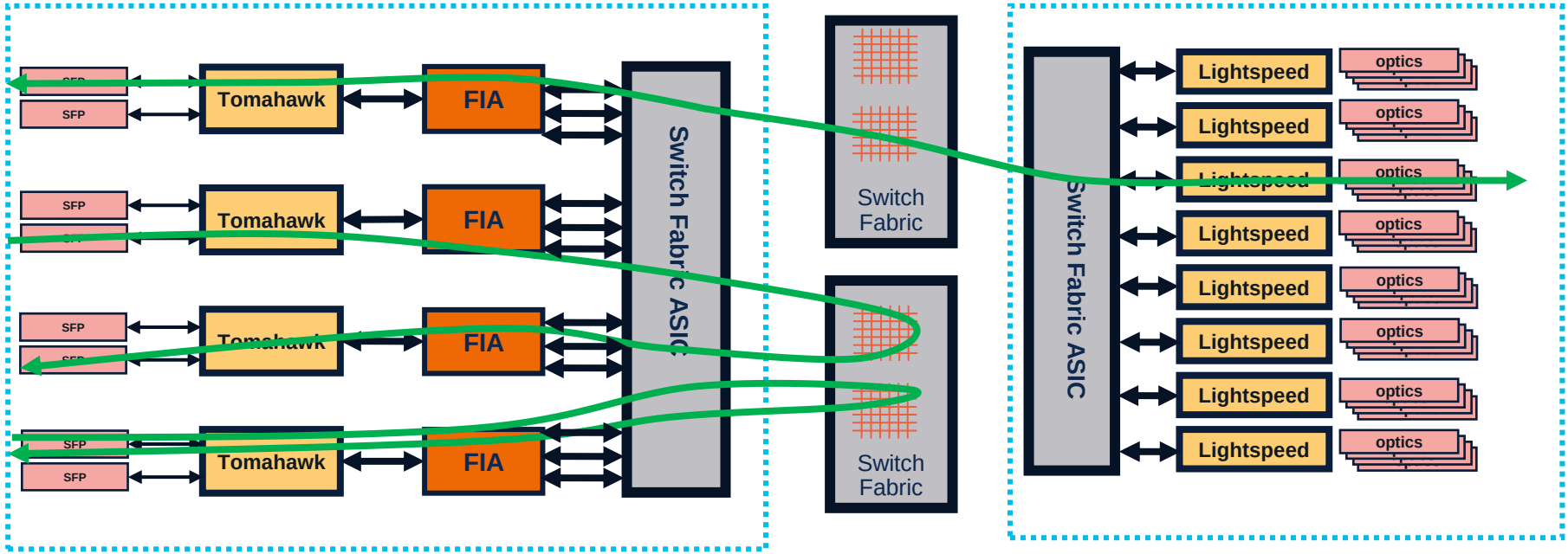
Example L3 VLAN sub-interface

```
RP/0/RSP0/CPU0:rasr9000-2w-b#show ethernet hardware
interface TenGigE 0/0/0/2.200 location 0/0/CPU0
Tue Jan 21 21:45:18.351 EST
------------------------------------------------------------------------------
Physical port                 : 2
Interface name                : TenGigE0/0/0/2.200
Ifhandle                      : 0x040012C0
Parent Ifhandle               : 0x04000140
Status                        : Init|Mem|Alloc|TShm|Pgm
TCAM entries                  : 1
TCAM entry type               : Single tag exact
Channel ID                    : 0
PI policy validity            : 0x0
NP port                       : 4
NP map (previous)             : 0x0 (0x0)
idb pointer                   : 0x5002b570
Admin mode                    : 1 (Up)
Interface state               : 1 (Up)
Interface type                : 3 (L3 Sub-if over
Physical)
tunn_ovrd_mode                : QnQ Child /w no parent
tunneling ethertype set
Ingress UIDB index            : 29
Egress UIDB index             : 29
------------------------------------------------------------------
TCAM key status: 0x404    index: 0

TCAM 0 address: 0x23880
TCAM 1 address: 0x0
TCAM 2 address: 0x0
.
```

```
.
........ TCAM entry 0 (uncompressed logical)
..............................
Port Number   :  mask=0xFFFF  value=0x0004(4)

Validity Bits:
    validity1:  mask=1       value=1
    validity2:  mask=1       value=0
    validity3:  mask=0       value=0
   isid_valid:  mask=0       value=0

Tag 1         :
    ethertype:  mask=0xFFFF  value=0x8100
    VLAN id   :  mask=0x0FFF  value=0x00C8(200)

Tag 2         :
    ethertype:  mask=0x0000  value=0x0000
    VLAN id   :  mask=0x0000  value=0x0000(0)

Source MAC    :
        mask :  0000.0000.0000
       value :  0000.0000.0000

........ TCAM entry 0 (2nd gen physical)
..................................

TCAM mask:
 FC  FF  00  00  FF  FF  FF  FF  FF  FF  FF  FF  00  F0
FF  00  00  BB  BB  BB
TCAM value:
 01  00  00  81  00  00  00  00  00  00  00  00  C8  00
00  04  00  00  00  00
```

Unicast Transit Frame Path

Physical > NP > FIA > Fabric > [FIA > NP] > Physical

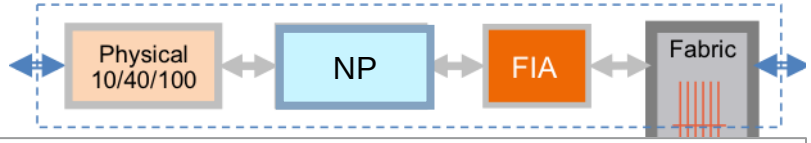


A9K-8X100GE

A99-32X100GE

The Internal Path

Mapping the port to NP and FIA




```
! Example: Path from GigabitEthernet0/0/1/0 192.3.1.2 TO TenGigE0/4/0/20.6 192.6.1.2
RP/0/RSP0/CPU0:rasr9k-1y#show controllers NP ports all location 0/0/CPU0
Fri Feb 22 15:57:32.307 UTC
                Node: 0/0/CPU0:
----------------------------------------------------------------
NP Bridge Fia                          Ports
-- ------ --- --------------------------------------------------
0  --     0   TenGigE0/0/0/0, TenGigE0/0/0/1, TenGigE0/0/0/2, TenGigE0/0/0/3
1  --     1   GigabitEthernet0/0/1/0 - GigabitEthernet0/0/1/19

RP/0/RSP0/CPU0:rasr9k-1y#show controllers NP ports all location 0/4/CPU0
Fri Feb 22 15:55:22.370 UTC
                Node: 0/4/CPU0:
----------------------------------------------------------------
NP Bridge Fia                          Ports
-- ------ --- --------------------------------------------------
0  --     0   TenGigE0/4/0/0, TenGigE0/4/0/1, TenGigE0/4/0/2
1  --     0   TenGigE0/4/0/3, TenGigE0/4/0/4, TenGigE0/4/0/5
2  --     1   TenGigE0/4/0/6, TenGigE0/4/0/7, TenGigE0/4/0/8
3  --     1   TenGigE0/4/0/9, TenGigE0/4/0/10, TenGigE0/4/0/11
4  --     2   TenGigE0/4/0/12, TenGigE0/4/0/13, TenGigE0/4/0/14
5  --     2   TenGigE0/4/0/15, TenGigE0/4/0/16, TenGigE0/4/0/17
6  --     3   TenGigE0/4/0/18, TenGigE0/4/0/19, TenGigE0/4/0/20
7  --     3   TenGigE0/4/0/21, TenGigE0/4/0/22, TenGigE0/4/0/23
```

Map the port to NP and FIA

Inside a Network Processor

Reading pipeline counters


```
RP/0/RSP0/CPU0:rasr9000-2w-a#show controllers NP counters np0 location 0/1/CPU0
Wed Nov 27 21:09:07.635 EST

                Node: 0/1/CPU0:
----------------------------------------------------------------

Show global stats counters for NP0, revision v2

Read 64 non-zero NP counters:
Offset  Counter                                         FrameValue   Rate (pps)
-------------------------------------------------------------------------------
  16   MDF_TX_LC_CPU                                       6722114           10
  17   MDF_TX_WIRE                                         1826039            3
  21   MDF_TX_FABRIC                                       1635541            2
  29   PARSE_FAB_RECEIVE_CNT                               1837406            3
  33   PARSE_INTR_RECEIVE_CNT                              5083364            7
  37   PARSE_INJ_RECEIVE_CNT                               1228130            2
.
 499   RSV_ING_L2_SMAC_MISS                                     60            0
 502   RSV_ING_L2_LEARN                                         60            0
 541   RSV_REFRESH_FROM_NOTIFY_CNT                              62            0
 584   RSV_L2BC_BVI                                              2            0
 604   RESOLVE_REMOTE_RACK_PREP_CNT                        5539915            8
.
 708   LRN_PERIODIC_AGING_DELETE_ENTRY                          60            0
.
 774   ARP                                                     119            0
.
 848   PUNT_ADJ                                                  2            0
 852   PUNT_ACL_DENY                                           161            0
 900   PUNT_STATISTICS                                     5083356            7
 902   PUNT_DIAGS_RSP_ACT                                    11419            0
 904   PUNT_DIAGS_RSP_STBY                                   11427            0
```

Description of NP counters:
`show controllers np descriptions location <location>`

NP Counters and Rates

Example: Ingress NP, no drops


```
RP/0/RSP0/CPU0:rasr9000-2w-b#show controllers NP counters np0 location 0/0/CPU0
Mon Dec  9 15:16:34.889 EST
                Node: 0/0/CPU0:
----------------------------------------------------------------

Show global stats counters for NP0, revision v2

Read 59 non-zero NP counters:
Offset  Counter                                        FrameValue   Rate (pps)
--------------------------------------------------------------------------------
  16  MDF_TX_LC_CPU                                       10255120            8
  17  MDF_TX_WIRE                                   6382883323428            1
  21  MDF_TX_FABRIC                                 8903307706961     31250074
  29  PARSE_FAB_RECEIVE_CNT                         6382883151049            0
  33  PARSE_INTR_RECEIVE_CNT                              8653828            8
  37  PARSE_INJ_RECEIVE_CNT                                744943            1
  41  PARSE_ENET_RECEIVE_CNT                        8910925981070     31250074
  45  PARSE_TM_LOOP_RECEIVE_CNT                           8035316            5
  49  PARSE_TOP_LOOP_RECEIVE_CNT                              61            0
  57  PARSE_ING_DISCARD                                  2344591            0
 195  PRS_HEALTH_MON                                     8035316            5
 204  INTR_FRAME_TYPE_7                                  8653827            8
 214  DBG_PRS_EP_L_PRS_VPLS_PW_IMPOSE                         10            0
 233  PARSE_RSP_INJ_FAB_CNT                              70634            0
 235  PARSE_RSP_INJ_DIAGS_CNT                            55255            0
 236  PARSE_EGR_INJ_PKT_TYP_UNKNOWN                      66847            0
 237  PARSE_EGR_INJ_PKT_TYP_IPV4                          3787            0
 246  PARSE_LC_INJ_FAB_CNT                              101092            0
.
```

To FIA

From Phy

NP Counters and Rates

NP drops, rates and direction


```
RP/0/RSP0/CPU0:rasr9000-2w-b#show controllers NP counters np0 location 0/0/CPU0
Tue Dec 10 14:18:39.195 EST

                Node: 0/0/CPU0:
----------------------------------------------------------------

Show global stats counters for NP0, revision v2

Read 59 non-zero NP counters:
Offset  Counter                                          FrameValue   Rate (pps)
--------------------------------------------------------------------------------
  16  MDF_TX_LC_CPU                                        11004363            9
  17  MDF_TX_WIRE                                      8712222364719     29761820
  21  MDF_TX_FABRIC                                   11063035007386     27714366
  29  PARSE_FAB_RECEIVE_CNT                            8712222113330     29761820
  33  PARSE_INTR_RECEIVE_CNT                              9401470            9
  37  PARSE_INJ_RECEIVE_CNT                                832185            1
  41  PARSE_ENET_RECEIVE_CNT                          11070653296959     27714366
  45  PARSE_TM_LOOP_RECEIVE_CNT                            8437075            5
 359  PARSE_MAC_NOTIFY_RCVD                                   183            0
 367  PARSE_FAST_DISCARD_LOW_PRIORITY_DROP_0           106211394050       883832
 368  PARSE_FAST_DISCARD_LOW_PRIORITY_DROP_1           106210662138       883856
 369  PARSE_FAST_DISCARD_LOW_PRIORITY_DROP_2           106211061617       883943
 370  PARSE_FAST_DISCARD_LOW_PRIORITY_DROP_3           106211474043       883922
 373  DBG_RSV_EP_L_RSV_ING_L3_IFIB                      3707021673            0
 830  PUNT_NO_MATCH                                        4746            0
 831  PUNT_NO_MATCH_EXCD                                 464963896            0
 849  PUNT_ADJ_EXCD                                        273406            0
 852  PUNT_ACL_DENY                                       1479378            0
 853  PUNT_ACL_DENY_EXCD                                1163570900            0
```

To egress

To fabric

From fabric

From interface

Typhoon NP catching up

NP Counters and Rates – Lightspeed (4th/5th gen)

NP drops, rates and direction

```
RP/0/RSP0/CPU0:rasr9000-2w-b#show controllers NP counters np2 location 0/4/CPU0
Tue Feb  1 12:12:49.750 CET

                Node: 0/4/CPU0:
----------------------------------------------------------------

Show global stats counters for NP2, revision v0

Offset  Counter                                          FrameValue   Rate (pps)
-------------------------------------------------------------------------------
 149  IPv4 QoS TCAM lookup miss                                  33            0
 159  IPv6 LL hash lookup miss on egress                          2            0
 162  IPv6 QoS TCAM lookup miss                               67632            0
 214  L2 egress LAG not local                                  1454            0
 384  Egress UIDB in down state                                  32            0
 430  IPv4 ACL deny                                             698            0
 456  IPv4 multicast fail RPF drop                               69            0
 457  IPv4 multicast fail RPF punt                           244747546           50
 691  CLNS multicast from fabric pre-route                    22633108            4
 692  Egress intf inject from fabric                             1818            0
 694  IPv4 from fabric                                            194            0
 695  IPv4 from fabric pre-route                                254175            0
 698  IPv4 multicast from fabric pre-route                    280087974           53
 700  IPv6 from fabric                                            839            0
 702  IPv6 link-local from fabric pre-route                     153029            0
 704  IPv6 multicast from fabric pre-route                       57151            0
 709  Inject to port                                          46809633            9
 710  MPLS from fabric                                         1883153            0
 713  Pre-route punt request                                    187982            0
1417  IPv4 disabled in UIDB                                         10            0
```

SW counters

NP Counters and Rates – Lightspeed (contd.)

NP drops, rates and direction

1467	MPLS leaf with no control flags set	134	0
1469	MPLS receive adjacency	84	0
1502	ARP	12348	0
1523	Diags	88210	0
1571	IPv4 options	5	0
1580	IPv4 multicast do all but forward	244747546	50
1586	ICMP generation needed	9	0
1598	TTL exceeded	688	0
1610	IFIB	28700797	6
1619	ACL denied packets punted for ICMP generation	698	0
1694	Diags RSP active	87548	0
1697	Diags RSP standby	87548	0
HW	Received from Line	879697268863	170139
HW	Transmit to Fabric	879450616886	170089
HW	Received from Fabric	2509265709824	512903
HW	Transmit to Line	2509287364175	512907
HW	Host Inject Received	30860357	6
HW	Host Punt Transmit	255817121	52
HW	Local Loopback Received at iGTR	286146849	58
HW	Local Loopback Transmit by iGTR	286146849	58
HW	Local Loopback Received at Egress	286146849	58
HW	Transmit to TM from eGTR	2509551824765	512961
HW	Transmit to L2	2509543181275	512959
HW	Received from Service Loopback	8643490	2
HW	Transmit to Service Loopback	8643490	2
HW	Internal generated by PDMA	124341747260	23474

SW counters

HW counters

NP Counters and Rates

NP drops


```
RP/0/RSP0/CPU0:ASR9006-2w-a.PE2#show drops np np0 location 0/1/CPU0
Thu Jul  7 16:53:34.665 EDT

                Node: 0/1/CPU0:
----------------------------------------------------------------------

NP 0 Drops:
----------------------------------------------------------------------
RSV_DROP_IN_L3_NOT_MYMAC                                        136912
MODIFY_PUNT_REASON_MISS_DROP                                    2
PARSE_EGR_INJ_PKT_TYP_UNKNOWN                                   4042
PARSE_DROP_IN_UIDB_TCAM_MISS                                    60081
PARSE_DROP_IN_UIDB_DOWN                                         15
PARSE_DROP_IPV4_MCAST_NOT_ENABLED                               331791
UNKNOWN_L2_ON_L3_DISCARD                                        341153
----------------------------------------------------------------------
RP/0/RSP0/CPU0:ASR9006-2w-a.PE2#
```

NP Counters and Rates

Per (sub)interface NP drop counters


```
RP/0/RSP0/CPU0:ASR9006-2w-a.PE2#monitor np interface TenGigE 0/0/0/1 count 3 time 10 location 0/0/CPU0
Thu Jul  7 04:49:13.840 EDT
 Monitor NP counters of TenGigE0_0_0_1 for 30 sec

              ****  Thu Jul  7 04:49:24 2016 ****

Monitor 0 non-zero NP0 counter:  TenGigE0_0_0_1
Offset  Counter                                        FrameValue   Rate (pps)
-------------------------------------------------------------------------------
1171  MDF_PUNT_POLICE_DROP                          7924962277743           21
(Count 1 of 3)

              ****  Thu Jul  7 04:49:34 2016 ****

Monitor 0 non-zero NP0 counter:  TenGigE0_0_0_1
Offset  Counter                                        FrameValue   Rate (pps)
-------------------------------------------------------------------------------
1171  MDF_PUNT_POLICE_DROP                          7924962277933           19
(Count 2 of 3)

              ****  Thu Jul  7 04:49:44 2016 ****

Monitor 0 non-zero NP0 counter:  TenGigE0_0_0_1
Offset  Counter                                        FrameValue   Rate (pps)
-------------------------------------------------------------------------------
1171  MDF_PUNT_POLICE_DROP                          7924962278163           23
(Count 3 of 3)
RP/0/RSP0/CPU0:ASR9006-2w-a.PE2#
```

Non-intrusive

Total per interface

Rate since last read
[10 seconds]

NP Counters and Rates

Decoding dropped frames


```
RP/0/RSP0/CPU0:ASR9006-2w-a.PE2#show controllers np capture np0 location 0/0/CPU0
Thu Jul  7 05:38:27.686 EDT

NP0 capture buffer has seen 8 packets - displaying 8

Sun Jul 03 20:51:59.414 : PARSE_DROP_IN_UIDB_DOWN
 From TenGigE0_0_0_1: 64 byte packet on NP0
0000: ff ff ff ff ff ff 10 f3 11 36 6a 04 08 06 00 01
0010: 08 00 06 04 00 02 10 f3 11 36 6a 04 0a 01 02 01
0020: ff ff ff ff ff ff 0a 01 02 01 00 00 00 00 00 00
0030: 00 00 00 00 00 00 00 00 00 00 00 00

Sun Jul 03 20:51:59.410 : PARSE_DROP_IN_UIDB_DOWN
 From TenGigE0_0_0_1: 253 byte packet on NP0
0000: 01 00 0c cc cc cc 10 f3 11 36 6a 04 00 eb aa aa
0010: 03 00 00 0c 20 00 02 b4 de 09 00 01 00 1c 41 53
0020: 52 39 30 30 31 2d 53 2d 32 59 2d 41 2e 63 69 73
0030: 63 6f 2e 63 6f 6d 00 03 00 12 54 65 6e 47 69 67
0040: 45 30 2f 30 2f 32 2f 30 00 02 00 11 00 00 00 01
0050: 01 01 cc 00 04 0a 01 02 01 00 04 00 08 00 00 00
0060: 01 00 05 00 5b 43 69 73 63 6f 20 49 4f 53 20 58
0070: 52 20 53 6f 66 74 77 61 72 65 2c 20 56 65 72 73
0080: 69 6f 6e 20 35 2e 33 2e 33 5b 44 65 66 61 75 6c
0090: 74 5d 0a 43 6f 70 79 72 69 67 68 74 20 28 63 29
00a0: 20 32 30 31 36 20 62 79 20 43 69 73 63 6f 20 53
00b0: 79 73 74 65 6d 73 2c 20 49 6e 63 2e 00 06 00 16
00c0: 63 69 73 63 6f 20 41 53 52 39 4b 20 53 65 72 69
00d0: 65 73 00 0a 00 06 00 00 00 0b 00 05 01 00 14 00
00e0: 1c 41 53 52 39 30 30 31 6d 39 f5 78 be fd 07 00
.
```

Filter out drops of no interest:

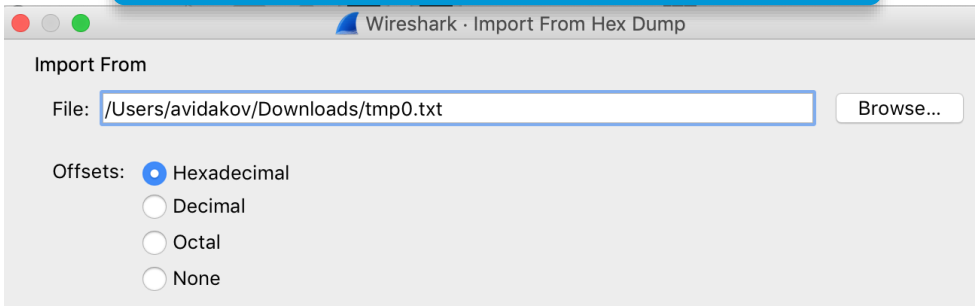
```
sh controllers np capture np1 filter …
```

NP Counters and Rates

Decoding dropped frames


```
▶ Frame 1: 160 bytes on wire (1280 bits), 160 bytes captured (1280 bits)
▶ IEEE 802.3 Ethernet
▶ Logical-Link Control
▼ Cisco Discovery Protocol
    Version: 2
    TTL: 180 seconds
  ▶ Checksum: 0xde09 [incorrect, should be 0xe54e]
  ▼ Device ID: ASR9001-S-2Y-A.cisco.com
      Type: Device ID (0x0001)
      Length: 28
      Device ID: ASR9001-S-2Y-A.cisco.com
  ▶ Port ID: TenGigE0/0/2/0
  ▼ Addresses
      Type: Addresses (0x0002)
      Length: 17
      Number of addresses: 1
    ▶ IP address: 10.1.2.1
  ▶ Capabilities
  ▼ Software Version
      Type: Software version (0x0005)
      Length: 91
      Software version: Cisco IOS XR Software, Version 5.3.3[Default]
```

Decode using Wireshark 'Import From Hex Dump'



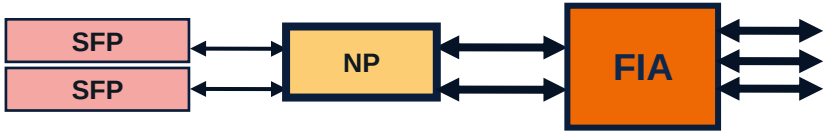
NP Counters and Rates

Traffic Manager drops


```
RP/0/RSP0/CPU0:rasr9000-2w-b#show controllers NP tm counters np1 location 0/0/CPU0
Tue Dec 10 14:40:47.210 EST

                Node: 0/0/CPU0:
----------------------------------------------------------------

==== TM Counters (NP 1 TM 0) ====
 TM Counters:
xmt paks: 897837659243,  xmt bytes: 62718673698431
drop paks: 29447137293,  drop_bytes: 2002405351616

RP/0/RSP0/CPU0:rasr9000-2w-b#
RP/0/RSP0/CPU0:rasr9000-2w-b#
RP/0/RSP0/CPU0:rasr9000-2w-b#show controllers NP tm counters np1 location 0/0/CPU0
Tue Dec 10 14:40:49.816 EST

                Node: 0/0/CPU0:
----------------------------------------------------------------

==== TM Counters (NP 1 TM 0) ====
 TM Counters:
xmt paks: 897909308598,  xmt bytes: 62723686013270
drop paks: 29466027670,  drop_bytes: 2003689898884
```

FIA Counters

FIA counts, drops and direction




```
RP/0/RSP0/CPU0:rasr9000-2w-b#show controllers fabric
fia instance 0 stats location 0/0/CPU0
Tue Dec 10 14:49:58.704 EST

 ********** FIA-0 **********
Category: count-0
         From Unicast Xbar[0]              733461306331
         From Unicast Xbar[1]              733460650405
           From Unicast Xbar[2]                       0
           From Unicast Xbar[3]                       0
       From MultiCast Xbar[0]                    233068
       From MultiCast Xbar[1]                         0
       From MultiCast Xbar[2]                         0
       From MultiCast Xbar[3]                         0
           To Unicast Xbar[0]              933450146675
           To Unicast Xbar[1]              932066610046
           To Unicast Xbar[2]                         0
           To Unicast Xbar[3]                         0
         To MultiCast Xbar[0]                    451799
         To MultiCast Xbar[1]                         0
         To MultiCast Xbar[2]                         0
         To MultiCast Xbar[3]                         0
         To Line Interface[0]            8759312354291
         To Line Interface[1]             457138023968
       From Line Interface[0]           11117127781061
       From Line Interface[1]             489302108080
                Ingress drop:              97191712670
                 Egress drop:                        0
                 Total drop:                97191712670
```

```
RP/0/RSP0/CPU0:rasr9000-2w-b#show controllers fabric fia
instance 0 drops ingress location 0/0/CPU0
Tue Dec 10 15:33:37.655 EST
 ********** FIA-0 **********
Category: in_drop-0
             From Spaui Drop-0                          0
                  accpt tbl-0                           0
                    ctl len-0                           0
                  short pkt-0                           0
                max pkt len-0                           0
                min pkt len-0                           0
             From Spaui Drop-1                          0
                  accpt tbl-1                           0
                    ctl len-1                           0
                  short pkt-1                           0
                max pkt len-1                           0
                min pkt len-1                           0
                     Tail drp               125787328841
                      Vqi drp                           0
          Header parsing drp                            0
                pw to ni drp                            0
             ni from pw drp                             0
                sp0 crc err                             0
              sp0 bad align                             0
                sp0 bad code                            0
             sp0 align fail                             3
               sp0 prot err                             0
                sp1 crc err                             0
              sp1 bad align                             0
.
```

Back pressure from egress NP

Line Card Drops

All drops for all reasons on a line card

```
RP/0/RSP0/CPU0:ASR9006-2w-a.PE2#show drops all
location 0/1/CPU0
Thu Jul  7 17:04:12.291 EDT


=====================================
Checking for drops on 0/1/CPU0
=====================================


show arp traffic:
[arp:ARP] IP Packet drop count for node 0/1/CPU0: 1


show cef drops:
[cef:0/1/CPU0] Discard drops        packets : 15


show controllers fabric fia drops ingress:
[fabric:FIA-0] sp0 crc err: 9
[fabric:FIA-0] sp0 align fail: 3
[fabric:FIA-0] sp1 align fail: 3
[fabric:FIA-1] sp0 align fail: 3
[fabric:FIA-1] sp1 crc err: 14
[fabric:FIA-1] sp1 bad code: 14
[fabric:FIA-1] sp1 align fail: 3
[fabric:FIA-1] sp1 prot err: 1
```

```
show netio drops:
[netio:Interface: GigabitEthernet0/1/0/10]
/pkg/lib/libipv4_fib_switch.dll: 4043
[netio:Interface: GigabitEthernet0/1/0/10]
packet_null_action: 305


show controller np counters:
[np:NP0] RSV_DROP_IN_L3_NOT_MYMAC: 137015
[np:NP0] MODIFY_PUNT_REASON_MISS_DROP: 2
[np:NP0] PARSE_DROP_IN_UIDB_TCAM_MISS: 60198
[np:NP0] PARSE_DROP_IN_UIDB_DOWN: 15
[np:NP0] PARSE_DROP_IPV4_MCAST_NOT_ENABLED: 332431
[np:NP0] UNKNOWN_L2_ON_L3_DISCARD: 341810
[np:NP1] MODIFY_PUNT_REASON_MISS_DROP: 3


show spp node-counters:
[spp:port4/classify] Dropped due to unknown SID: 164862
[spp:port4/classify] Invalid: logged n dropped: 1
```

Line Card Drops

Modifying the " show drops all" template

Example: include drops from " show controllers np fast-drop" command in the output of the " show drops all" command

Step 1: Copy the /pkg/etc/packet_drops.list file to /disk0a:/usr/

```
run
cd /pkg/etc
cp packet_drops.list /disk0a:/usr/
exit
```

Step 2: Edit the file offline or on the router using the 'vim' editor in the shell.

```
run
vim /disk0a:/usr/packet_drops.list
exit
```

Step 3: Add this sequence to the end of the '#NP' section of the packet_drops.list file:

```
[commandstart]
cmd_name = show controller np fast-drop
cmd_exec = prm_np_show fast-drop -s $location
module = np
group = ^.*Show NP EFD stats counters for (NP\d),
default_group = ERROR!! - Group not found!!
drop_regex1 = ^\s*(\S+Priority[0-9]\S)\s+(\d+)
[commandend]
```

Step 4: Verify the command works as expected by using the 'self-test' option:

```
sh drops all self-test location <location> | b fast-drop
```

You should see in the output everything except the lines that are matching the drop_regex1 pattern. Compare this to the output of:

```
sh controllers np fast-drop all location <location>
```

Troubleshooting: Packet Tracing

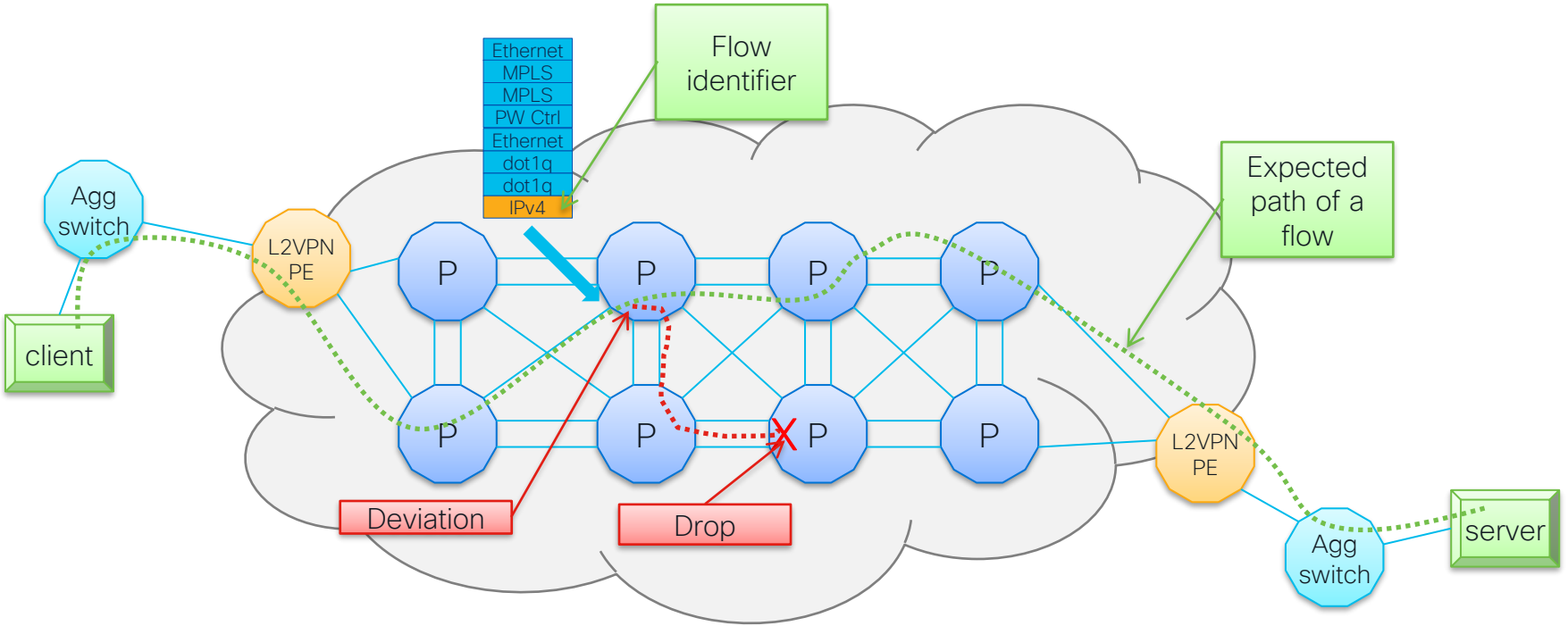
ASR9k: Embedded Packet Tracer Purpose

- Validate service provisioning by tracing the matching flow through the system
- Perform in-depth triaging of packet forwarding issues in data-path and punt-path (inject-path in planning)
- Learn the platform and XR platform independent forwarding infrastructure

ASR9k: Embedded Packet Tracer Key Features

- User-mode intuitive CLI for user interaction (no configuration required).
- Very flexible condition specification to specify a flow of interest.
- Offline Web App to easily derive conditions for a arbitrary header stack.
- Trace packets of a flow through the system.
- Embedded Packet Tracer does not change packet disposition

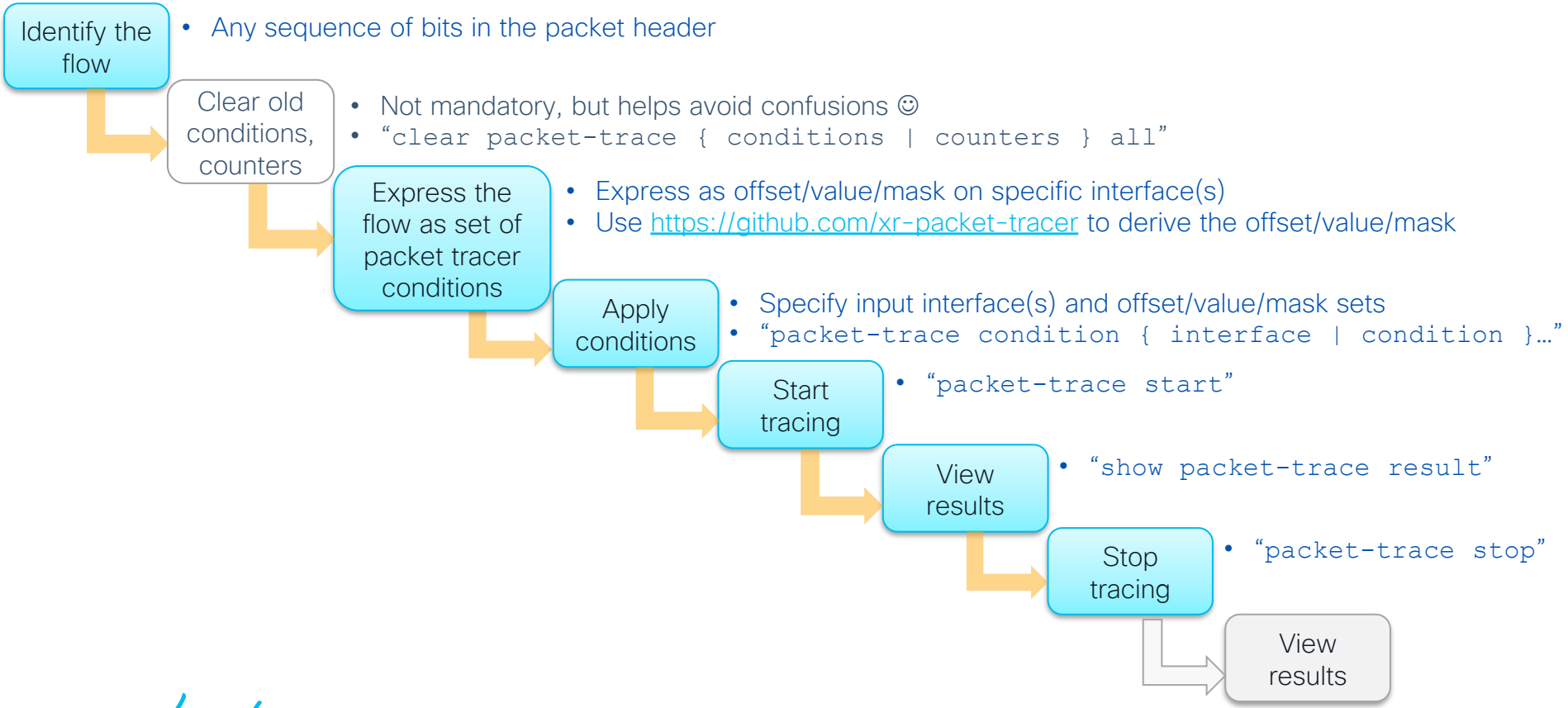
Service Verification/Troubleshooting In L2 VPN



Supported Platforms/Releases

- XR Release 7.1.2:
  - Mark packets on ASR 9000 Tomahawk and Lightspeed+ NP
  - Count traced packets on ASR 9000 Tomahawk and Lightspeed+ NP
- XR Release 7.5.2:
  - Count traced packets on punt path:
    - SPP
    - NetIO libraires
    - UDP, TCP

User Interaction



Embedded Packet Tracer CLI

Command	Description
packet-trace condition interface *<interface>*	Specify interface on which marking should be enabled
packet-trace condition *<id>* **offset** *<offset>* **value** *<value>* **mask** *<mask>*	Specify the conditions as a set of offset/value/mask triplets
packet-trace start	Start packet marking
packet-trace stop	Stop packet marking
show packet-trace status	Display the status of packet tracer: • Conditions buffered by packet trace master process • Tracing state (active/inactive), based on the execution of “start” and “stop” commands
show packet-trace status [**detail**]	Same as above, plus for every location: • List every counting module and any errors it reported • List every marking module, active conditions and any errors it reported
show packet-trace result	Display all non-zero counters
clear packet-trace conditions all	Clear all conditions (only allowed when tracing stops)
clear packet-trace counters all	Clear all counters
show packet-trace description [**detail**]	Display all supported counters

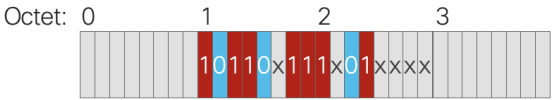
Embedded Packet Trace Condition

- Specify interfaces where condition is applied
  - Condition applies to physical interface
  - To trace on sub-interface, calculate in the encapsulation in the offset
- Specify the offset/value/mask sets
  - ASR9k PD supports up to 3x 4-octet value/mask sets

```
packet-trace condition interface Hu0/5/0/1
packet-trace condition interface Hu0/5/0/3
packet-trace condition 1 offset 14 value 0x7e4 mask 0xfff
packet-trace condition 2 offset 30 value 0xc0a80003 mask 0xffffffff
packet-trace condition 3 offset 34 value 0xc0a80002 mask 0xffffffff
```

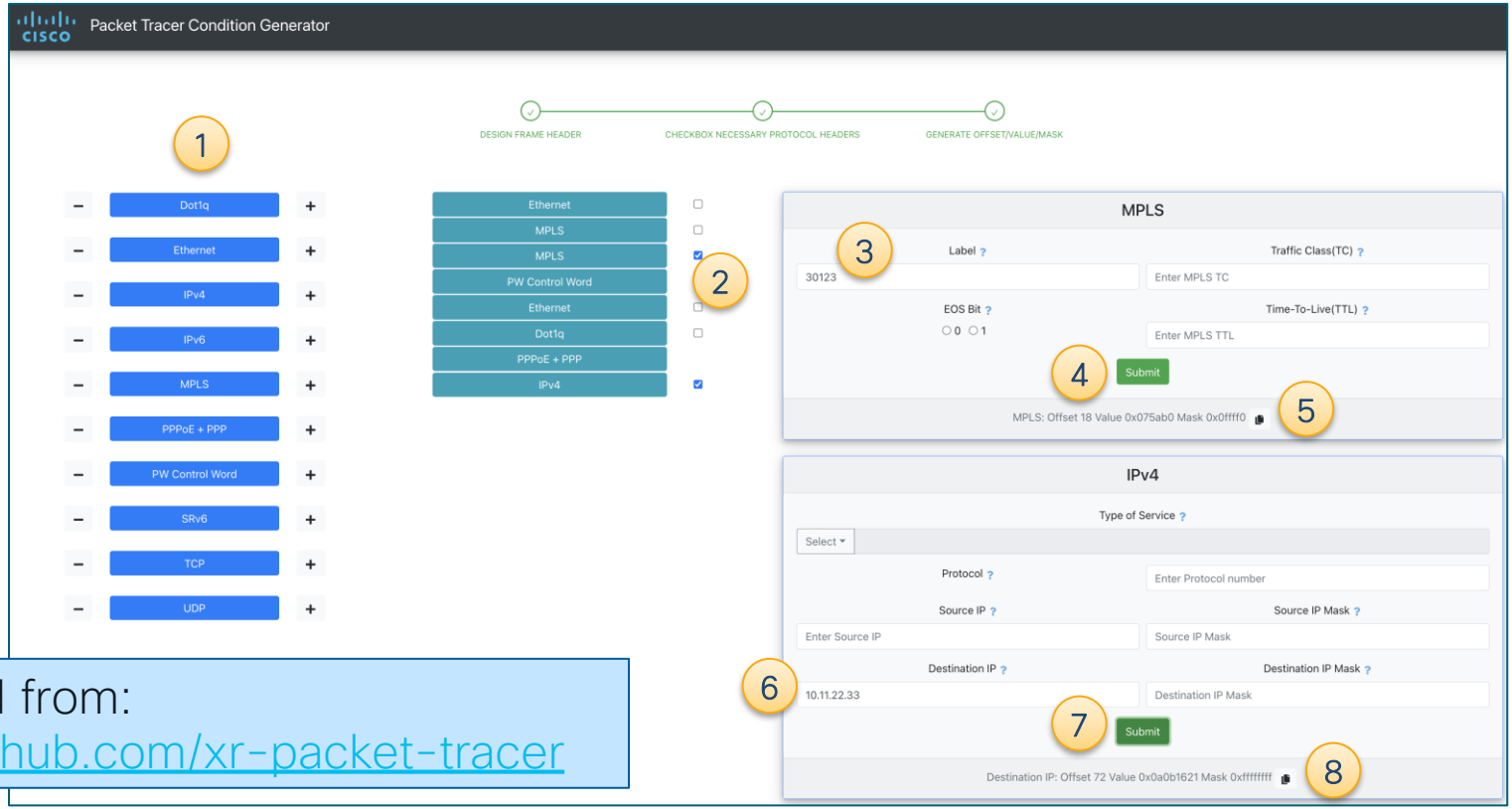
Packet Trace Condition

- “Condition” is expressed in offset/value/mask triplets
  - Offset is expressed in octets from the very start of the Ethernet frame
  - Value defines the expected values on positions we care about starting from the offset
  - You can set 0 or 1 into positions you don’t care about
  - Mask defines which bits we care and which we don’t care about starting from the offset




```
Pattern starts with octet 1          ➔ offset: 1
Pattern value: b1011001110010000
                     b   3   9   0   ➔ value : 0xb390
Pattern mask : b1111101110110000
                     f   b   b   0   ➔ mask  : 0xfbb0
```

Packet Trace Condition Generator Web App



Download from:
https://github.com/xr-packet-tracer

Clear Conditions and Counters

- Packet trace counters can be cleared at any time

```
clear packet-trace counters all
```

- Packet trace conditions can be cleared when packet tracing is not active

```
clear packet-trace conditions all
```

Start / Stop Tracing

- Packet tracer master process on RP sends the specified conditions to all LCs
- LCs that own target interfaces program the NP and start marking packets
- Packet trace flag is preserved in the NP/fabric/punt/inject packet headers

```
packet-trace start
```

- When packet tracing is stopped, marking modules clear the condition and stop marking packets

```
packet-trace stop
```

Show Packet Trace Status – Simple Output

- Displays conditions buffered by packet trace master and status
  - Status is derived from the packet-trace start/stop command

```
RP/0/RSP0/CPU0:CORE-TOP#show packet-trace status
Packet Trace Master Process:

  Buffered Conditions:
    Interface HundredGigE0_5_0_1
    Interface HundredGigE0_5_0_3
    1 offset 14 value 0x7e4 mask 0xfff
    2 offset 30 value 0xc0a80003 mask 0xffffffff
    3 offset 34 value 0xc0a80002 mask 0xffffffff

  Status: Inactive

RP/0/RSP0/CPU0:CORE-TOP#
```

Conditions buffered by the packet trace master process on active RP

Status derived from the start/stop command

View Packet Trace Results

- Counter type:
  - *Marking* – packet has matched the condition and was marked for tracing
  - *Pass* – this counter describes the action performed on the packet. Packet is passed on for further processing.
  - *Drop* – this counter signals that the NP µcode made a decision to drop the packet. Drop reason is not provided. Try correlating the NP drop counters with the increment of this counter

```
show packet-trace results
```

View Packet Trace Results

```
RP/0/RSP0/CPU0:CORE-TOP#show packet-trace results
Thu Jul 11 17:03:57.477 UTC
T: D - Drop counter; P - Pass counter
Location | Source | Counter                    | T | Last-Attribute                          |           Count
--------   ------   -------------------------   -   ----------------------------------------   ---------------
0/5/CPU0   NP0      PACKET_MARKED                P   HundredGigE0_5_0_1                                  1000
0/5/CPU0   NP0      PACKET_TO_FABRIC             P                                                       1000
0/5/CPU0   NP0      PACKET_FROM_FABRIC           P                                                       1000
0/5/CPU0   NP0      PACKET_TO_INTERFACE          P   HundredGigE0_5_0_0                                  1000
```

Location

Source identifies the NP number on the location

Counter name

Type (explained on previous slide)

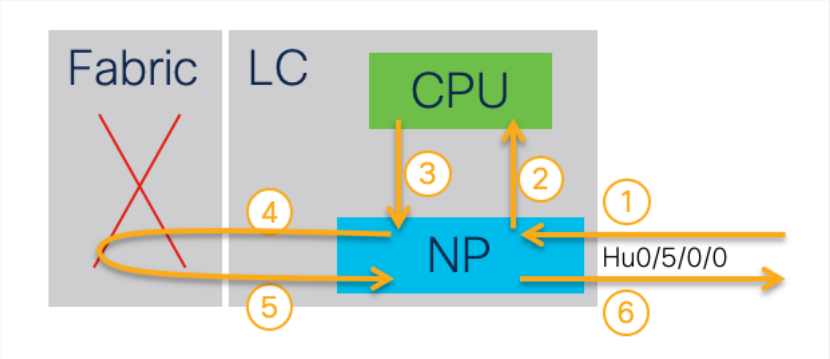
Counter value

Order of counters tries to illustrate the order of events in packet processing, but this may not always be the case.

- With every counter update, packet trace framework also receives the timestamp and allows for a freeform “attribute” that describes more closely the action.
- The use of attribute and its meaning differs between counters
- Current CLI only exposes the last attribute to the user.
- In future releases, user will be able to see the last 1023 timestamps and attributes per counter.

Example: ICMP Echo Request Processing Path

```
RP/0/RSP0/CPU0:CORE-TOP#sh packet-trace results
Wed Aug 26 13:36:52.679 UTC
T: D - Drop counter; P - Pass counter
Location     | Source       | Counter                    | T | Last-Attribute                          | Count
-------------  -------------  --------------------------   -   ----------------------------------------  ---------------
0/5/CPU0       NP0            PACKET_MARKED (1)            P   HundredGigE0_5_0_0                      100
0/5/CPU0       NP0            PACKET_FROM_INJECT (3)       P                                           100
0/5/CPU0       NP0            PACKET_FROM_FAB_INJECT (5)   P                                           100
0/5/CPU0       NP0            PACKET_TO_FABRIC (4)         P                                           100
0/5/CPU0       NP0            PACKET_TO_PUNT (2)           P                                           100
0/5/CPU0       NP0            PACKET_FROM_FABRIC (5)       P                                           100
0/5/CPU0       NP0            PACKET_TO_INTERFACE (6)      P   HundredGigE0_5_0_0                      100
RP/0/RSP0/CPU0:CORE-TOP#
```



References

- https://xrdocs.io/asr9k//tutorials/xr-embedded-packet-tracer/
- https://www.youtube.com/watch?v=5aCAwdQF8SE

Troubleshooting: Packet Capture

Packet Capture: Problem Packets

Example: incrementing drops


```
RP/0/RSP0/CPU0:rasr9000-2w-b#show controllers NP counters np0 location 0/0/CPU0 | include DROP
Sat Jan 18 18:46:52.618 EST
.
 370   RSV_DROP_XID_NO_MATCH                      209680463           0
 404   RSV_ING_VPWS_ERR_DROP                   3719838164404    11160601
 411   RSV_L2_SHG_DROP                            27390624           0
1171   MDF_PUNT_POLICE_DROP                    7924962278163    23809032
1178   MODIFY_PUNT_REASON_MISS_DROP                   1           0
1246   VIRTUAL_IF_GENERIC_INPUT_DROP                  1           0


RP/0/RSP0/CPU0:rasr9000-2w-b#show controllers NP counters np0 location 0/0/CPU0 | include DROP
Sat Jan 18 18:46:56.297 EST
.
 370   RSV_DROP_XID_NO_MATCH                      209680463           0
 404   RSV_ING_VPWS_ERR_DROP                   3719879236984    11161027
 411   RSV_L2_SHG_DROP                            27390624           0
1171   MDF_PUNT_POLICE_DROP                    7925049898728    23809936
1178   MODIFY_PUNT_REASON_MISS_DROP                   1           0
1246   VIRTUAL_IF_GENERIC_INPUT_DROP                  1           0
```

incrementing

Rate [PPS] or increments from last command run

Packet Capture: Problem Packets

Example: incrementing drops


```
RP/0/RSP0/CPU0:rasr9000-2w-b#monitor np counter RSV_ING_VPWS_ERR_DROP np0 count 3 location
0/0/CPU0
Sat Jan 18 19:02:36.386 EST
Warning: Every packet captured will be dropped! If you use the 'count'
         option to capture multiple protocol packets, this could disrupt
         protocol sessions (eg, OSPF session flap). So if capturing protocol
         packets, capture only 1 at a time.

Warning: A mandatory NP reset will be done after monitor to clean up.
         This will cause ~50ms traffic outage. Links will stay Up.
 Proceed y/n [y] >
```

Packet Capture: Problem Packets

Example: incrementing drops


```
RP/0/RSP0/CPU0:rasr9000-2w-b#monitor np counter RSV_ING_VPWS_ERR_DROP np0 count 3 location
0/0/CPU0
Sat Jan 18 19:02:36.386 EST
Warning: Every packet captured will be dropped! If you use the 'count'
         option to capture multiple protocol packets, this could disrupt
         protocol sessions (eg, OSPF session flap). So if capturing protocol
         packets, capture only 1 at a time.

Warning: A mandatory NP reset will be done after monitor to clean up.
         This will cause ~50ms traffic outage. Links will st
 Proceed y/n [y] >
 Monitor RSV_ING_VPWS_ERR_DROP on NP0 ... (Ctrl-C to quit)
Sat Jan 18 19:02:44 2014 -- NP0 packet
 From TenGigE0/0/0/0: 157 byte packet, bytes[0-3] invalid!
0000: 00 00 02 01 61 90 00 00 c0 02 01 02 81 00 00 0a   ....a...@.......
0010: 08 00 45 00 00 8b 00 00 00 00 40 3d f8 30 c0 01   ..E.......@=x0@.
0020: 01 01 c0 01 01 02 00 00 00 00 00 00 00 00 00 00   ..@.............
0030: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00   ................
0040: 10 f3 11 05 00 00 00 00 00 00 00 00 00 00 00 00   ................
0050: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00   ................
0060: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00   ................
0070: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00   ................
0080: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00   ................
0090: 00 00 00 00 00 00 00 00 00 00 00 00 00            .............
```

Ignore [internal]

UP to 300 B
No CRC

Packet Capture: Problem Packets

Example: incrementing drops


```
 From TenGigE0/0/0/0: 234 byte packet, bytes[0-3] invalid!
0000: 00 00 02 01 61 90 00 00 c0 02 01 02 81 00 00 0a    ....a...@.......
0010: 08 00 45 00 00 d8 00 00 00 00 40 3d f7 e3 c0 01    ..E..X....@=wc@.
0020: 01 01 c0 01 01 02 00 00 00 00 00 00 00 00 00 00    ..@.............
0030: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00    ................
0040: 10 f3 11 05 00 00 00 00 00 00 00 00 00 00 00 00    .s..............
0050: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00    ................
0060: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00    ................
0070: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00    ................
0080: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00    ................
0090: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00    ................
00a0: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00    ................
00b0: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00    ................
00c0: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00    ................
00d0: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00    ................
00e0: 00 00 00 00 00 00 00 00 00 00                      ..........

          (count 3 of 3)


Cleanup: Confirm NP reset now (~50ms traffic outage).
 Ready? [y] >
RP/0/RSP0/CPU0:rasr9000-2w-b#
```

Packet Capture: Problem Packets

Decoding


```
00 00 02 01 61 90 00 00 c0 02 01 02 81 00 00 0a
08 00 45 00 00 8b 00 00 00 00 40 3d f8 30 c0 01
01 01 c0 01 01 02 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
10 f3 11 05 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00
```

Decode using Wireshark 'Import From Hex Dump'

```
Ethernet II, Src: WesternD_02:01:02 (00:00:c0:02:01:02), Dst: Xerox_01:61:90 (00:00:02:01:61:90)
    Destination: Xerox_01:61:90 (00:00:02:01:61:90)
    Source: WesternD_02:01:02 (00:00:c0:02:01:02)
    Type: 802.1Q Virtual LAN (0x8100)
802.1Q Virtual LAN, PRI: 0, CFI: 0, ID: 10
    000. .... .... .... = Priority: 0
    ...0 .... .... .... = CFI: 0
    .... 0000 0000 1010 = ID: 10
    Type: IP (0x0800)
Internet Protocol, Src: 192.1.1.1 (192.1.1.1), Dst: 192.1.1.2 (192.1.1.2)
    Version: 4
    Header length: 20 bytes
    Differentiated Services Field: 0x00 (DSCP 0x00: Default; ECN: 0x00)
    Total Length: 139
```

Packet Capture: HW Counters On Lightspeed

Example: HW counters

```
RP/0/RSP0/CPU0:xrg-402-asr9906#monitor np ?
  counter         next packet to increment a specific NP counter(cisco-support)
  crc-frame-err   Ingress packets with CRC or Framing errors(cisco-support)
  fabric-egress   Egress packets from fabric(cisco-support)
  interface       interface to display counters(cisco-support)
  net-ingress     Ingress packets from network links(cisco-support)
RP/0/RSP0/CPU0:xrg-402-asr9906#monitor np
```

Packet Capture: Transit Packets

Example: IPv4 L3VPN ingress

```
RP/0/RSP0/CPU0:rasr9000-2w-b#show running-config ipv4 access-list CAPTURE
Sat Jan 18 20:13:35.941 EST
ipv4 access-list CAPTURE
 10 permit ipv4 192.4.1.0/24 10.10.6.0/24 capture
 20 permit ipv4 any any
!
RP/0/RSP0/CPU0:rasr9000-2w-b#show running-config interface TenGigE 0/0/0/2
Sat Jan 18 20:13:50.654 EST
interface TenGigE0/0/0/2
 vrf TRAFFIC
 ipv4 address 192.4.1.1 255.255.255.0
 ipv4 access-group CAPTURE ingress
!
RP/0/RSP0/CPU0:rasr9000-2w-b#show controllers NP counters np0 location 0/0/CPU0 | include
ACL_CAPTURE_NO_SPAN
Sat Jan 18 20:14:26.109 EST
 477   ACL_CAPTURE_NO_SPAN                              6802507        38003
RP/0/RSP0/CPU0:rasr9000-2w-b#show controllers NP counters np0 location 0/0/CPU0 | include
ACL_CAPTURE_NO_SPAN
Sat Jan 18 20:14:28.819 EST
 477   ACL_CAPTURE_NO_SPAN                              6905417        380
RP/0/RSP0/CPU0:rasr9000-2w-b#show controllers NP counters np0 location 0/0/C
ACL_CAPTURE_NO_SPAN
Sat Jan 18 20:14:34.597 EST
 477   ACL_CAPTURE_NO_SPAN                              7124969        37991
```

Count in NP

Let all else go!

Apply to transit

NP ACL “capture” counter incrementing

Packet Capture: Transit Packets

Example: IPv4 L3VPN ingress

```
RP/0/RSP0/CPU0:rasr9000-2w-b#monitor np counter ACL_CAPTURE_NO_SPAN np0 count 3 location 0/0/CPU0
Sat Jan 18 20:31:53.311 EST
Warning: Every packet captured will be dropped! If you use the 'count'
         option to capture multiple protocol packets, this could disrupt
         protocol sessions (eg, OSPF session flap). So if capturing protocol
         packets, capture only 1 at a time.

Warning: A mandatory NP reset will be done after monitor to clean up.
         This will cause ~50ms traffic outage. Links will stay Up.
 Proceed y/n [y] >
```

Packet Capture: Transit Packets

Example: IPv4 L3VPN ingress

```
RP/0/RSP0/CPU0:rasr9000-2w-b#monitor np counter ACL_CAPTURE_NO_SPAN np0 count 3 location 0/0/CPU0
Sat Jan 18 20:31:53.311 EST
Warning: Every packet captured will be dropped! If you use the 'count'
         option to capture multiple protocol packets, this could disrupt
         protocol sessions (eg, OSPF session flap). So if capturing protocol
         packets, capture only 1 at a time.

Warning: A mandatory NP reset will be done after monitor to clean up.
         This will cause ~50ms traffic outage. Links will stay Up.
 Proceed y/n [y] >
 Monitor ACL_CAPTURE_NO_SPAN on NP0 ... (Ctrl-C to quit)

Sat Jan 18 20:32:34 2014 -- NP0 packet
 From TenGigE0/0/0/2: 250 byte packet, bytes[0-3] invalid!
0000: 00 11 0b 00 61 92 00 00 c0 04 01 02 08 00 45 60   ....a...@.....E`
0010: 00 ec 00 00 00 00 40 3d a8 08 c0 04 01 02 0a 0a   .l....@=(.@.....
0020: 06 5d 00 00 00 00 00 00 00 00 00 00 00 00 00 00   .]..............
0030: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00   ................
0040: 10 f3 11 05 00 00 00 00 00 00 00 00 00 00 00 00   .s..............
0050: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00   ................
0060: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00   ................
0070: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00   ................
0080: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00   ................
0090: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00   ................
.
```

Those 3 packets are dropped!

Packet Capture: Transit Packets

Example: IPv4 L3VPN ingress

```
.
            (count 2 of 3)
Sat Jan 18 20:32:36 2014 -- NP0 packet
 From TenGigE0/0/0/2: 220 byte packet, bytes[0-3] invalid!
0000: 00 11 0b 00 61 92 00 00 c0 04 01 02 08 00 45 00   ....a...@.....E.
0010: 00 ce 00 00 00 00 40 3d a8 bc c0 04 01 02 0a 0a   .N....@=(<@.....
0020: 06 27 00 00 00 00 00 00 00 00 00 00 00 00 00 00   .'..............
0030: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00   ................
0040: 10 f3 11 05 00 00 00 00 00 00 00 00 00 00 00 00   .s..............
0050: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00   ................
0060: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00   ................
0070: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00   ................
0080: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00   ................
0090: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00   ................
00a0: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00   ................
00b0: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00   ................
00c0: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00   ................
00d0: 00 00 00 00 00 00 00 00 00 00 00 00               ............
            (count 3 of 3)

Cleanup: Confirm NP reset now (~50ms traffic outage).
 Ready? [y] >
RP/0/RSP0/CPU0:rasr9000-2w-b#
```

Those 3 packets were dropped!

Packet Capture: Transit Packets

Decoding the packet

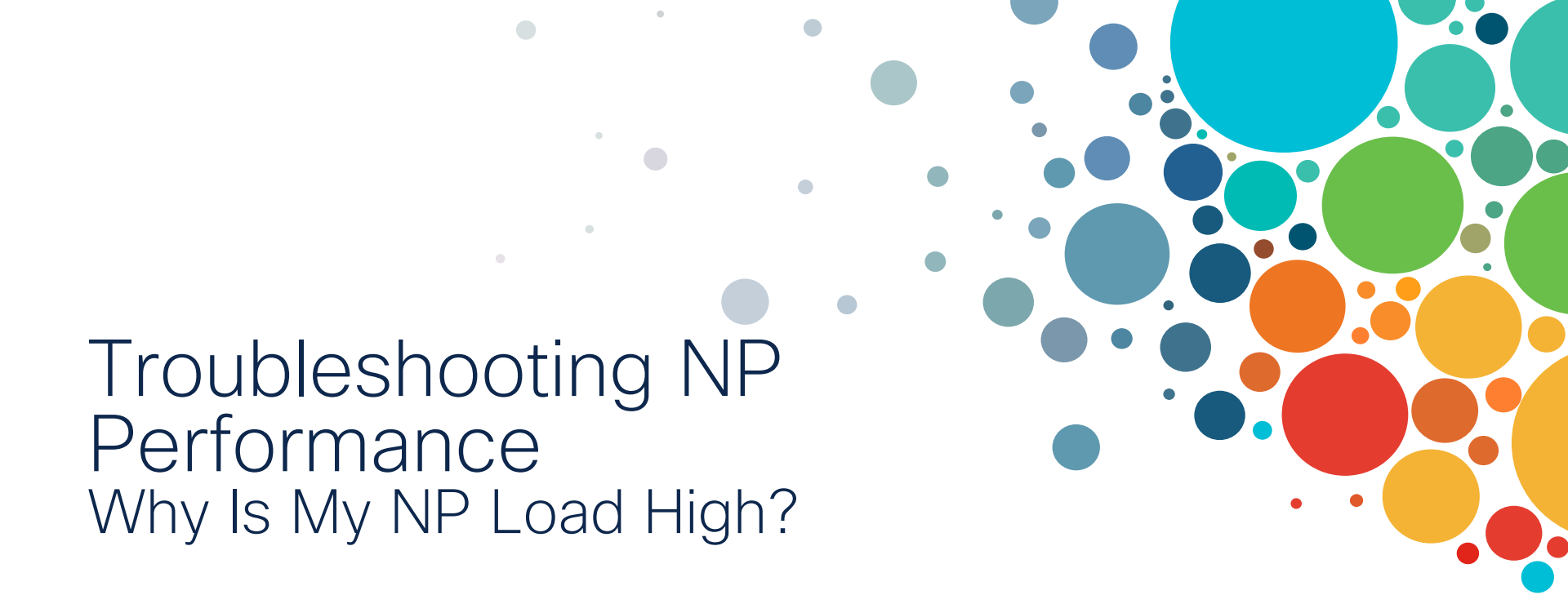
```
00 11 0b 00 61 92 00 00 c0 04 01 02 08 00 45 60
00 ec 00 00 00 00 40 3d a8 08 c0 04 01 02 0a 0a
06 5d 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
10 f3 11 05 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00 00 00 00 00 00 00 00 00 00
```

Decode using Wireshark 'Import From Hex Dump'

```
Ethernet II, Src: WesternD_04:01:02 (00:00:c0:04:01:02), Dst: Franklin_00:61:92 (00:11:0b:00:61:92)
Internet Protocol, Src: 192.4.1.2 (192.4.1.2), Dst: 10.10.6.93 (10.10.6.93)
    Version: 4
    Header length: 20 bytes
    Differentiated Services Field: 0x60 (DSCP 0x18: Class Selector 3; ECN: 0x00)
    Total Length: 236
    Identification: 0x0000 (0)
    Flags: 0x00
    Fragment offset: 0
    Time to live: 64
    Protocol: SHIM6 header (0x3d)
```

Troubleshooting NP Performance

Why Is My NP Load High?



cisco Live!

High NP Load

- Typical: complex feature processing keeps packets in NP buffers for extended time
- Atypical: excessive egress replications keep NP pipeline busy
- Enhanced version of np_perf shell facility on TH/LS/LS+ line cards:
  - Better monitoring of NP buffer pool(s) utilisation
  - Provides insight into average and peak buffer utilisation over a longer period of time, while not overloading the CPU
  - Available since XR release 7.1.2

Enhanced np_perf On Tomahawk

```
[xr-vm_node0_2_CPU0:~]$np_perf -h
     Usage: np_perf -e<channel> [-n<samples>] [-t<ms sample time>] -[opts]
     -e                  NP channel (0,1,2,3)
     -n <sample>, --num number of samples to take, min:100, def:2000,
     -R, --rfd           show average flow-control on time per source
         <optional> -T, thread mode, def = FALSE (console mode)
         <optional> -D <sec>, run duration in second
                     between 10 and 200000, def = 0 (one shot)
                     repeat -n samples(>= 2000) for at least <sec>
                     -T will be on automatically
         <optional> -k <rfd_threshold>, between 100 and 10000, def = 3000
         <optional> -o <selection>, def = run all selections
                    1: 10G ~ 100G interface
                    2: 40G ~ 100G interface
                    3: QSGMII interface
                    4: special interface
                    5: Group 1~8
                    6: all bank 0~31
                    7: ICFD/OCFD queue
                    8: TM global queue
                    9: TM output queue
         <optional> -x <xfi port_id>, between 0 and 47.
                    100G: port_id % 12 == 0
                     40G: port_id % 4 == 0
                     10G: all other
```

Mandatory arguments

irrelevant output omitted

Enhanced np_perf On Lightspeed/Lightspeed+

```
[xr-vm_node0_0_CPU0:~]$np_perf
     Usage: np_perf -e<channel> [-n<samples>] [-t<ms sample time>] -[opts]
     -e, NP channel (0..3)
     -r, Global Packet Buffer (GPB) info/monitor
         [-w], show current thresholds
         [-T], thread mode, def = FALSE (console mode)
         [-D <sec>], run duration in second
             between 3 and 200000, def = 0 (one shot)
             repeat -n samples(>= 2000) for at least <sec>
             -T will be on automatically with -D
         [-k <soft threshold>], between 1 and 40000, def = 30000
         [-o <selection>], (info operation)
             0: all ingress and egress info (default)
             1: ingress 100G ports info
             2: all ingress info
             3: ingress port info
             4: ingress PSA info
             5: ingress CMN/Shared space info
             6: all egress info
             7: egress Channel info
             8: egress PSA info
             9: egress CMN space info
         [-x <port_id>], (monitor operation)
             0 ~ 42: select one of ingress interface port
             43: ingress all 100G ports
             44: ingress all ports
             45: ingress SecChn1 service lpbk
             46: ingress SecChn2 HPI
             47: egress  Primary channel
             48: egress  SecChn1 recycle lpbk
             49: egress  SecChn2 local lpbk
             50: ingress CMN space
             51: ingress shared space
             52: egress  CMN space
```

Mandatory arguments

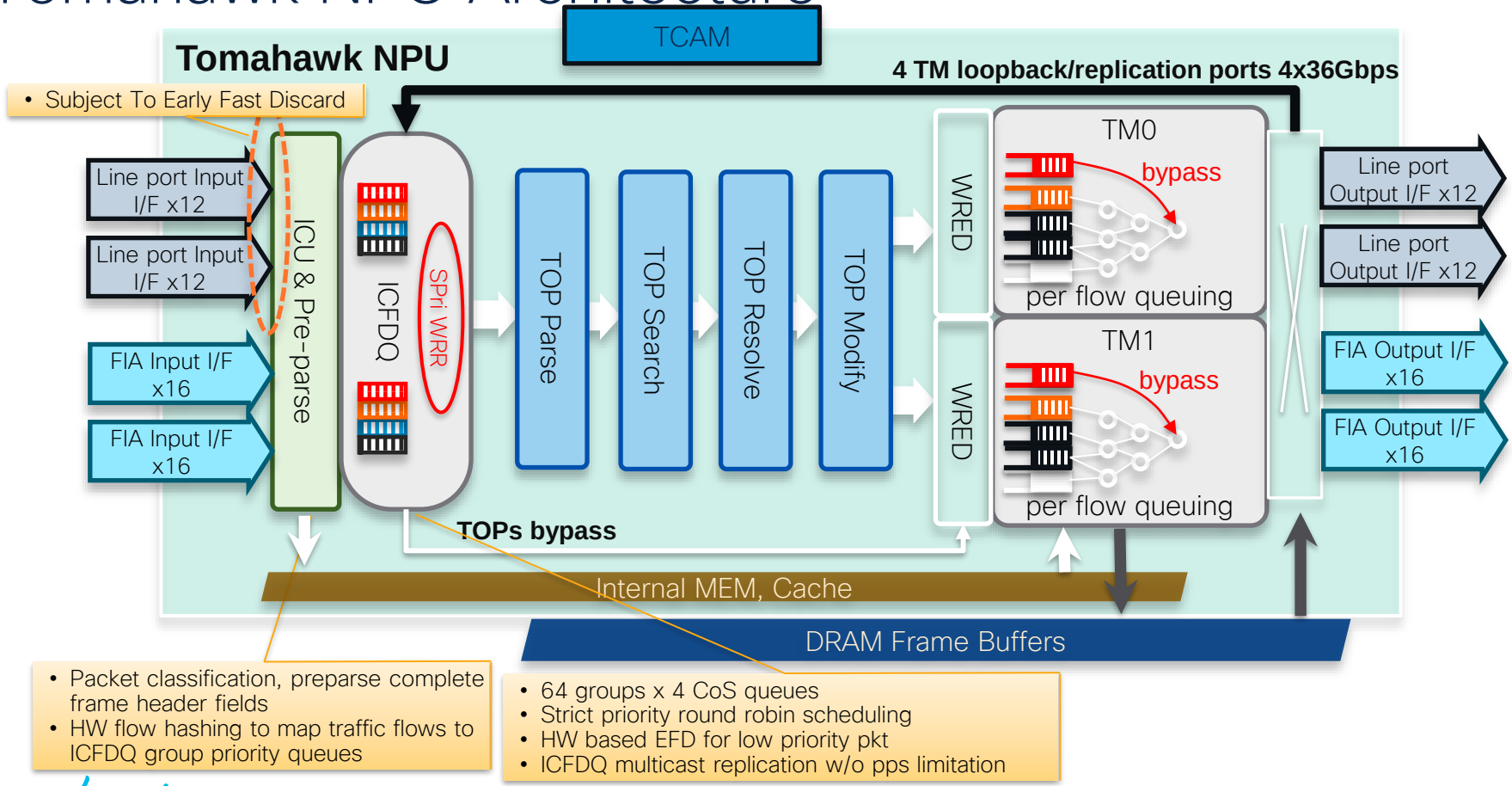
irrelevant output omitted

Enhanced np_perf Operation

- Continuously read current utilisation of selected buffer pool
- Print one line of output after every ‘-n’ reads
- Release LC CPU:
  - Tomahawk: release CPU for 1us on every 500 reads
  - Lightspeed: release CPU for 0.5us on every 1000 reads
- If current read exceeds the ‘-k’ threshold, print current usage
- If ‘-D’ or ‘-T’ is specified, run in the background as a thread of:
  - Tomahawk: prm_server_to process
  - Lightspeed: npu_server process (npu_server_main thread)

- Logfile name example when using ‘-D’ option:
- /misc/scratch/np/npu_gpb_dump_0_0_CPU0_np0_x43_20200721-163940.442912.txt

Tomahawk NPU Architecture

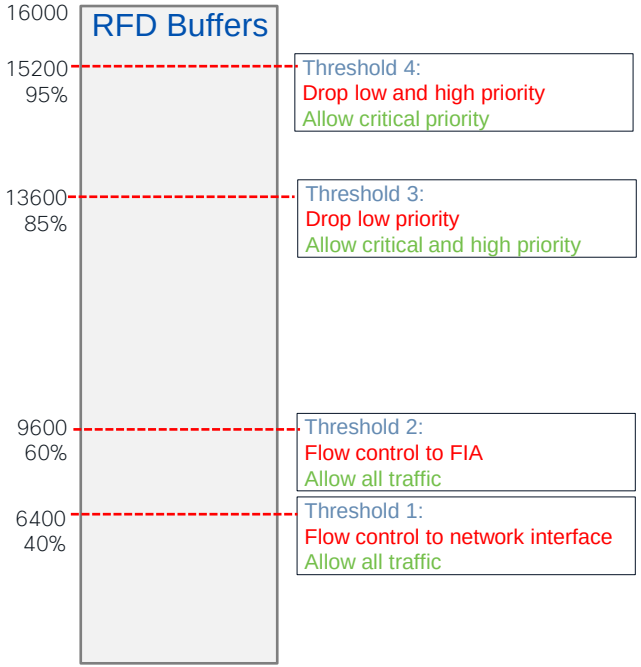


Ingress Packet Pre-Processing On Tomahawk

- Tomahawk NP has a single RFD buffer pool shared between ingress/egress path processing
- At times of congestion important packet are protected by Early Frame Discard (EFD) mechanism:
  - EFD thresholds are based on RFD buffer utilisation:
    - Low/High threshold per network interface
    - Global threshold: based on % of total RFD buffers in use
- EFD is performed in HW, before passing packet to TOP engine feature processing
- EFD is only performed on packets received on network interfaces
- Packets that pass EFD are stored in 256-byte RFD buffers.

Tomahawk EFD Thresholds

Global EFD Thresholds



Per Network Interface EFD Thresholds

HundredGigE Interface

Threshold	Value
3600 / 4200(*)	Drop low and high priority Allow critical priority
3400 / 3800(*)	Drop low priority Allow critical and high priority

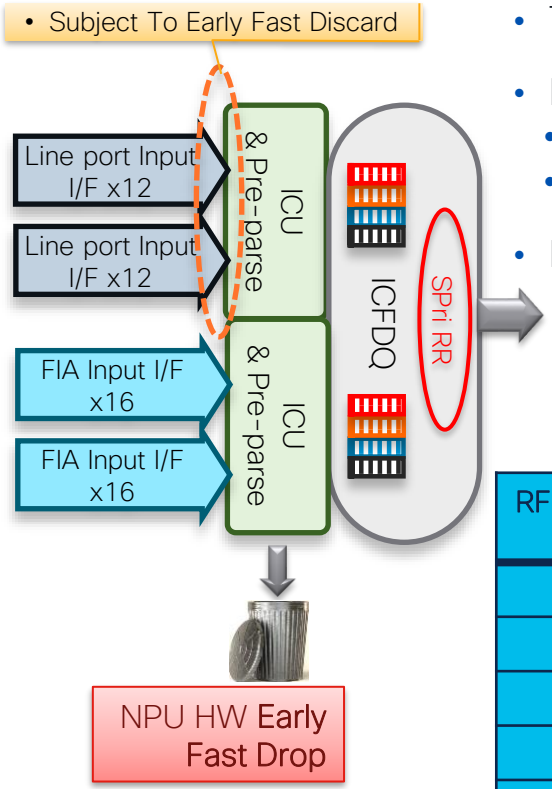
(*) new CLI (7.1.x and later, via CSCvu03480):

```
hw-module location <location> early-fast-discard burst-absorption
```

TenGigE Interface

Threshold	Value
690	Drop low and high priority Allow critical priority
360	Drop low priority Allow critical and high priority

Tomahawk HW Based EFD



- Tomahawk implements HW based priority EFD
- Existing Typhoon SW based EFD for priority classification and discard logic
  - *hw location 0/x/cpu0 early-fast-discard <ip, mpls, vlan cos, i/o encp> val op*
  - **Supports IPv4 and IPv6**
- HW priority discard criteria
  1. Input Frame Resource Congestion State:
     - RFD consumption (per network interface and global)
  2. Packet priority as classified by ICU: 4 classes, 3 actually used
     - Network control, High priority (ToS/Exp/Vlan Cos/DSCP >= 6) and Low priority

RFD Usage	Line Side Forward Priority (No Drop)	Line Side Early Fast Drop Priority	Line Side Flow Control	Fabric Side Flow Control
>95%	Control	High, Low Priority	On*	On
>85%	Control, High priority	Low priority	On*	On
>60%	All	None	On*	On
>40%	All	None	On*	Off
<40%	All	None	Off	Off

* If not CLI disabled

What Causes NP Fast Drops?

- EFD drops a packet when RFD threshold is hit (i.e. RFD utilisation is high)
- What may cause high RFD utilisation:
  - RFD leak
  - NP is overloaded
- How to confirm RFD leak:

RFD utilisation while traffic flows	RFD utilisation when traffic stops	Likely Cause
Average ~= Peak ~= RFD threshold	Average ~= Peak ~= RFD threshold	RFD leak ➔ bug in ucode
Average ~= Peak ~= RFD threshold	Average ~= Peak ~= 0	NP overload ➔ heavy features (e.g. BVI, uRPF, etc.)
Average < Peak (by order of magnitude)	Average ~= Peak ~= 0	Bursty traffic, possibly heavy features

Example - Input

- run np_perf in background during 3600 seconds and save output into a file (-D)
- Read only RFD buffer utilisation (-R)
  - Limit to specific port (-x): 100G ports and both FIA ports
  - execute 50000 reads before printing a summary line (-n)
  - report syslog when instantaneous RFD utilisation is above 4000 (-k)

```
run ssh lc0_xr /pkg/bin/np_perf -e0 -R -n50000 -k4000 -x0 -D3600
run ssh lc0_xr /pkg/bin/np_perf -e0 -R -n50000 -k4000 -x12 -D3600
run ssh lc0_xr /pkg/bin/np_perf -e0 -R -n50000 -k4000 -x24 -D3600
run ssh lc0_xr /pkg/bin/np_perf -e0 -R -n50000 -k4000 -x32 -D3600
```

Example – Output Snippet

Output of –k option: whenever the threshold specified by '-k' argument is exceeded, output is produced into syslog and standard output

```
Apr  6 17:33:21.734422, NP_1 xfi_0 , N: 1427 , rfd: 3811   (> 3700)
Apr  6 17:33:21.734422, NP_1 xfi_12, N: 140  , rfd: 3818   (> 3700)
Apr  6 17:33:21.734422, NP_1 xfi_24, N: 4969 , rfd: 3831   (> 3700)
Apr  6 17:33:21.734422, NP_1 xfi_32, N: 889  , rfd: 3819   (> 3700)
```

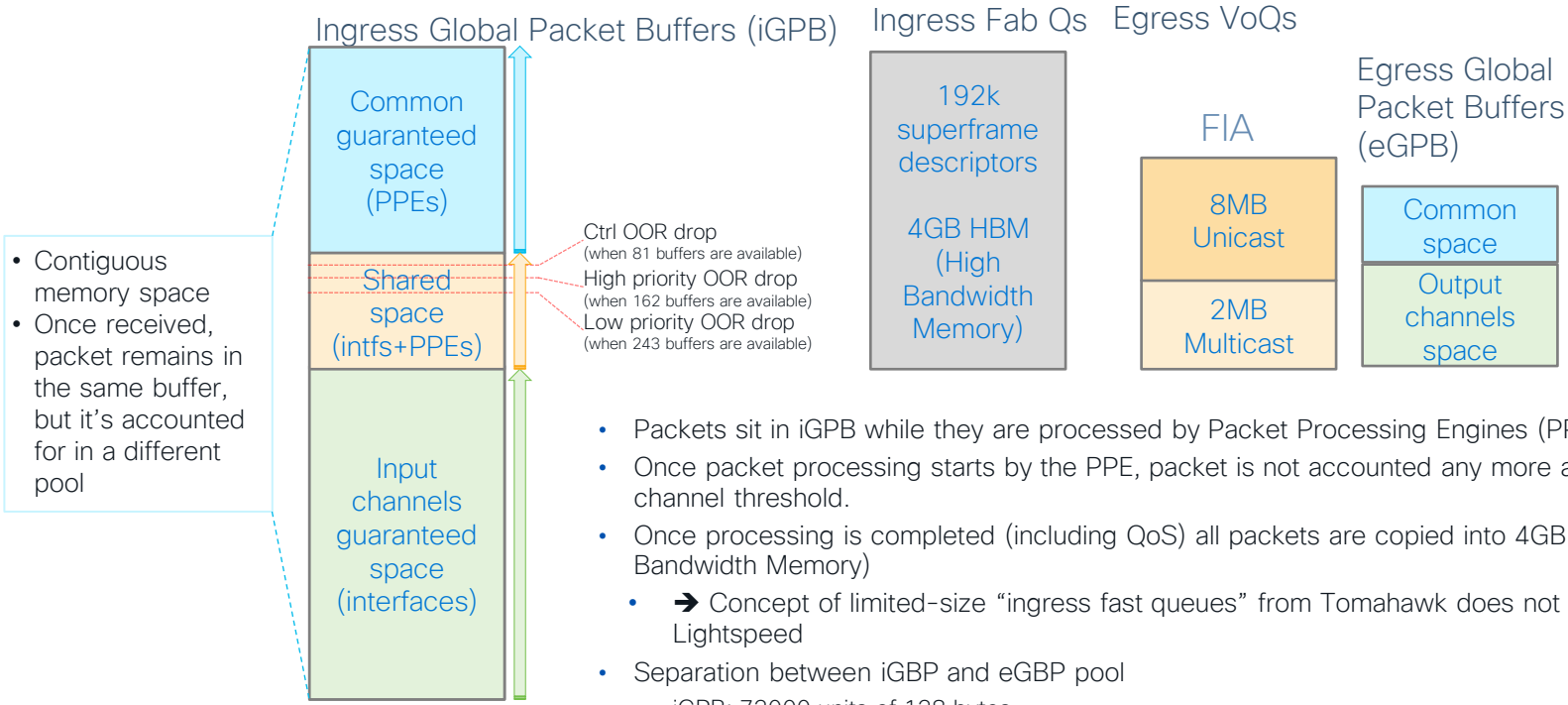
← 3811 RFD buffers in use on Hu0/4/0/3
← 3818 RFD buffers in use on Hu0/4/0/2
← 3831 RFD buffers in use on fabric interface 0
← 3719 RFD buffers in use on fabric interface 1

```
Jun  7 17:32:31.120 HundredGigE0_4_0_3     180  3811
Jun  7 17:32:32.323 HundredGigE0_4_0_3     109  3109
Jun  7 17:32:33.540 HundredGigE0_4_0_3     101  2807
Jun  7 17:32:34.712 HundredGigE0_4_0_3     301  3801
```

Average and peak RFD utilisation over the last '-n' reads

XFI port	Interface
0	Hu0/4/0/3
12	Hu0/4/0/2
24	FIA port 0
32	FIA port 1

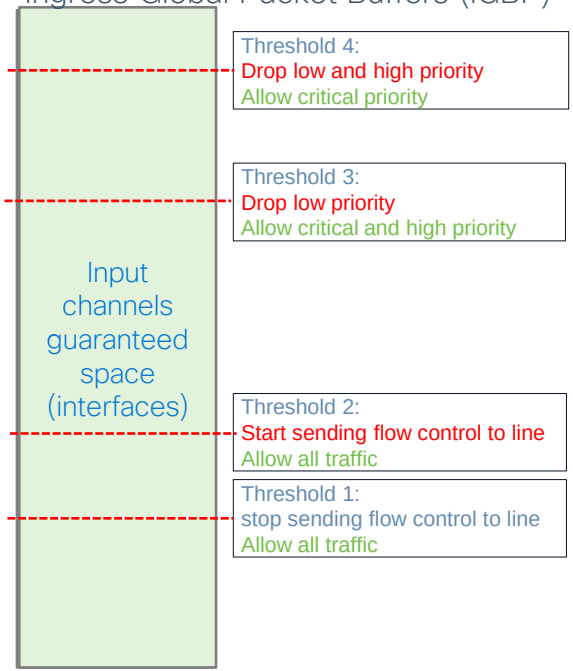
Lightspeed iGPB and eGPB



- Packets sit in iGPB while they are processed by Packet Processing Engines (PPE)
- Once packet processing starts by the PPE, packet is not accounted any more against input channel threshold.
- Once processing is completed (including QoS) all packets are copied into 4GB HBM (High Bandwidth Memory)
  - ➔ Concept of limited-size "ingress fast queues" from Tomahawk does not apply on Lightspeed
- Separation between iGBP and eGBP pool
  - iGPB: 72000 units of 128 bytes
  - eGPB: 26000 units of 128 bytes

Lightspeed EFD Thresholds – Input channel

Ingress Global Packet Buffers (iGBP)



Threshold values per single HundredGigE interface

Threshold	Description	Value
4	Drop low and high priority	10766
3	Drop low priority	8944
2	Start sending flow control to line	6509
1	Stop sending flow control to line	4112

View NP Buffer Thresholds

np_perf -e<n> -r -w

```
          Ingress GPB
                                        Taildrop Thresh              L2
Port                 MinGuar CTL    HP     LP2    LP1      Xoff   Xon
--------------------+-------+-----------------------------+--------------
HundredGigE0_0_0_8   8944    11899  10766  8944   8944     6509   4112
HundredGigE0_0_0_9   8944    11899  10766  8944   8944     6509   4112
HundredGigE0_0_0_10  8944    11899  10766  8944   8944     6509   4112
HundredGigE0_0_0_11  8944    11899  10766  8944   8944     6509   4112

                                              DropThreshRemaining
Type                 MinGuar  Max     CTL     HP      LP2     LP1
--------------------+---------------+----------------------------
Common Space         20000    26000
Shared Space         0        11860   81      162     243     324

          Egress GPB
                            Lp                  Hp
Egress Channel       MinGuar  BpThresh  MinGuar  BpThresh
--------------------+------------------+-----------------
Primary              2540     1268      2540     1268

Type                 MinGuar
--------------------+--------
Common Space         20216
```

Where's LS+ Perf Limit? Test Results

Interface	In(bps)	Out(bps)
Hu0/0/0/5	82.6G/ 82%	94.5G/ 94%
Hu0/0/0/0	97.0G/ 97%	94.6G/ 94%
Hu0/0/0/13	82.5G/ 82%	90.0G/ 89%
Hu0/0/0/17	96.4G/ 96%	89.8G/ 90%
Hu0/0/0/8	90.4G/ 90%	87.1G/ 87%
Hu0/0/0/9	90.1G/ 89%	86.0G/ 85%
Hu0/0/0/10	91.5G/ 91%	88.2G/ 88%
Hu0/0/0/11	91.2G/ 91%	88.0G/ 87%

NP2

Average values
(based on multiple snapshots during the test)

Features:
- ingress/egress QoS
- MPLS imposition (L3VPN
- Ingress netflow

```
          Node: 0/0/CPU0:
----------------------------------------------
            Load                  Packet Rate
  NP2: 56% utilization        232878440 pps
  NP2: 26% utilization        233491860 pps
  NP2: 56% utilization        241413480 pps
  NP2: 53% utilization        298298940 pps
  NP2: 42% utilization        233614360 pps
  NP2: 64% utilization        236480540 pps
  NP2: 57% utilization        230305960 pps
```

varying readout due to bursty traffic

np_perf Threads Running In Parallel

```
# +++++++++++ ingress all 100G ports
run ssh lc5_xr /pkg/bin/np_perf -e0 -r -n1000000 -x43 -D3600

# +++++++++++ ingress CMN space
run ssh lc5_xr /pkg/bin/np_perf -e0 -r -n1000000 -x50 -D3600

# +++++++++++ ingress shared space
run ssh lc5_xr /pkg/bin/np_perf -e0 -r -n1000000 -x51 -D3600

# +++++++++++ egress Primary channel
run ssh lc5_xr /pkg/bin/np_perf -e0 -r -n1000000 -x47 -D3600

# +++++++++++ egress CMN space
run ssh lc5_xr /pkg/bin/np_perf -e0 -r -n1000000 -x52 -D3600
```

NP iGPB utilisation (Top 5)

LP threshold:
8944

HP threshold:
10766 – 8944 = 1822

Observations:

- LP not approaching 50%
- HP not approaching 1.5%

Time	Port	Avg	Peak	Wm	HpAvg	HpPeak	HpWm
Jun 20 01:03:44.723901	**HundredGigE0_0_0_8**	13	4164	10	0	8	10
Jun 20 01:05:59.736445	HundredGigE0_0_0_8	13	3988	10	0	8	8
Jun 20 00:56:42.636269	HundredGigE0_0_0_8	14	3602	10	0	8	0
Jun 20 01:00:57.933415	HundredGigE0_0_0_8	13	3395	12	0	8	0
Jun 20 01:02:50.073372	HundredGigE0_0_0_8	13	3356	10	0	8	8

sorted

sorted

Time	Port	Avg	Peak	Wm	HpAvg	HpPeak	HpWm
Jun 20 01:05:00.272599	HundredGigE0_0_0_8	13	2004	10	0	10	0
Jun 20 01:04:45.076353	HundredGigE0_0_0_8	9	1417	10	0	10	0
Jun 20 01:03:20.181699	HundredGigE0_0_0_8	10	1544	10	0	10	8
Jun 20 01:01:43.596496	HundredGigE0_0_0_8	15	2652	12	0	10	0
Jun 20 01:00:32.929748	HundredGigE0_0_0_8	13	2094	12	0	10	0

Time	Port	Avg	Peak	Wm	HpAvg	HpPeak	HpWm
Jun 20 00:57:09.381446	**HundredGigE0_0_0_9**	10	2974	12	0	8	0
Jun 20 01:05:06.236427	HundredGigE0_0_0_9	10	2922	12	0	8	0
Jun 20 00:59:53.835802	HundredGigE0_0_0_9	11	2912	10	0	8	8
Jun 20 01:03:36.333409	HundredGigE0_0_0_9	10	2886	10	0	8	8
Jun 20 01:01:44.785470	HundredGigE0_0_0_9	9	2858	10	0	8	8

sorted

sorted

Time	Port	Avg	Peak	Wm	HpAvg	HpPeak	HpWm
Jun 20 01:05:46.171598	HundredGigE0_0_0_9	10	1684	12	0	22	0
Jun 20 01:02:46.365554	HundredGigE0_0_0_9	8	1556	12	0	11	0
Jun 20 01:06:05.979454	HundredGigE0_0_0_9	10	1536	10	0	10	8
Jun 20 01:03:01.581849	HundredGigE0_0_0_9	11	2103	148	0	10	0
Jun 20 01:02:30.778633	HundredGigE0_0_0_9	9	1716	10	0	10	8

NP iGPB utilisation (Top 5)

- Common space at ~60% of max
- Empty shared space
- ➔ ingress pipeline PPEs are not overloaded

Thresholds

```
          Ingress GPB
Type                  MinGuar   Max      CTL      HP       LP2      LP1
--------------------+---------+--------+-------------------------------
Common Space          20000     26000
Shared Space          0         11860    81       162      243      324
```

sorted

Measurements

```
Time                      Space                Avg     Peak    Wm
------------------------+--------------------+-------------------
Jun 20 01:05:59.891605    Ing Cmn Space        4240    15085   15239
Jun 20 01:03:45.665564    Ing Cmn Space        4180    14724   15097
Jun 20 00:58:10.271313    Ing Cmn Space        4231    14099   14256
Jun 20 01:04:43.159634    Ing Cmn Space        4239    13632   13869
Jun 20 00:59:06.216773    Ing Cmn Space        4263    13429   13638
Jun 20 01:01:53.595518    Ing Cmn Space        4209    13350   13628
Jun 20 00:58:41.276052    Ing Cmn Space        4216    12999   13264
Jun 20 00:59:16.923882    Ing Cmn Space        4192    12997   13318
Jun 20 01:03:51.374061    Ing Cmn Space        4194    12955   13071
Jun 20 00:59:37.209121    Ing Cmn Space        4226    12737   12933
```

Measurements

```
Time                      Space                Avg     Peak    Wm
------------------------+--------------------+-------------------
Jun 15 16:27:51.575630    Ing Shared Space     0       0       0
```

NP eGPB Egress Primary Pool Utilisation (Top 5)

Thresholds

Egress Channel	Lp MinGuar	Lp BpThresh	Hp MinGuar	Hp BpThresh
Primary	2540	1268	2540	1268

Measurements

Time	Port	Avg	Peak	Wm	HpAvg	HpPeak	HpWm
Jun 20 00:59:07.161242	Egress Primary	43	1742	1742	4	1420	1510
Jun 20 01:00:41.680441	Egress Primary	45	1732	1742	4	1386	1524
Jun 20 01:02:38.946098	Egress Primary	44	1724	1724	4	1456	1620
Jun 20 01:01:40.263936	Egress Primary	44	1724	1724	4	1370	1460
Jun 20 00:57:35.265513	Egress Primary	42	1724	1724	3	1274	1594

sorted

Observations:

- LP not approaching 70%
- HP not approaching 65%

Measurements

sorted

Time	Port	Avg	Peak	Wm	HpAvg	HpPeak	HpWm
Jun 20 00:57:21.642064	Egress Primary	44	1680	1680	4	1620	1626
Jun 20 01:06:28.024603	Egress Primary	46	1680	1718	4	1596	1600
Jun 20 01:05:11.635047	Egress Primary	43	1682	1728	4	1588	1612
Jun 20 01:05:52.364817	Egress Primary	45	1680	1686	4	1578	1622
Jun 20 00:57:30.865742	Egress Primary	46	1688	1688	4	1558	1592

NP eGPB Egress CommonPool Utilisation
(Top 5)

Thresholds

```
Type                    MinGuar
--------------------+--------
Common Space            20216
```

Measurements

sorted

```
Time                      Space                 Avg    Peak    Wm
------------------------+---------------------+-------------------
Jun 20 01:03:10.035183  Egr Cmn Space        3343   10691   10771
Jun 20 01:02:08.450449  Egr Cmn Space        3358   10547   10987
Jun 20 01:03:02.794878  Egr Cmn Space        3367   10478   10948
Jun 20 01:03:53.247814  Egr Cmn Space        3357   10471   10668
Jun 20 01:05:35.467574  Egr Cmn Space        3366   10275   10535
Jun 20 01:02:39.539830  Egr Cmn Space        3354   10271   10493
Jun 20 00:56:49.441531  Egr Cmn Space        3327   10270   10487
Jun 20 01:02:46.863921  Egr Cmn Space        3342   10209   10350
Jun 20 01:01:07.863248  Egr Cmn Space        3344   10177   10558
Jun 20 00:59:42.809658  Egr Cmn Space        3375   10132   10571
```

Observations:

- Not approaching 55%

Additional Fast Drop Troubleshooting Commands

```
show interface <interface>
show controllers <interface> stats
show policy-map interface <interface>
show qoshal default-queue interface <interface>
show qoshal loopback-queue interface <interface>
show controllers np fast-drop np<np_number> location <location>
show controller np counters np<np_number> location <location>
show interfaces <interface> accounting rates
```

```
telemetry model-driven
 sensor-group NP
  sensor-path Cisco-IOS-XR-asr9k-np-oper:hardware-module-np/nodes/node/nps/np/efd
  sensor-path Cisco-IOS-XR-asr9k-np-oper:hardware-module-np/nodes/node/nps/np/fast-drop
  sensor-path Cisco-IOS-XR-asr9k-np-oper:hardware-module-np/nodes/node/nps/np/counters
  sensor-path Cisco-IOS-XR-asr9k-np-oper:hardware-module-np/nodes/node/nps/np/load-utilization
```

Agenda

- ✓ System Architecture: System anatomy & health
- ✓ Operating System & Configuration: IOS-XR & configuration models
- ✓ Control, Management, Security: Processing of control & exceptions
- ✓ Transit Packet/Frame Journey: Life of L3/L2 unicast/multicast
- ✓ MPLS Operation: Processing, forwarding & L3/L2 service operation
- ✓ Troubleshooting: Diagnostics, counters, drops, and packet capture

Complete your Session Survey

- Please complete your session survey after each session. Your feedback is important.
- Complete a minimum of 4 session surveys and the Overall Conference survey (open from Thursday) to receive your Cisco Live t-shirt.

- All surveys can be taken in the Cisco Events Mobile App or by logging in to the Session Catalog and clicking the " Attendee Dashboard" at https://www.ciscolive.com/emea/learn/sessions/session-catalog.html

Continue Your Education

Visit the Cisco Showcase for related demos.

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

Attend any of the related sessions at the DevNet, Capture the Flag, and Walk-in Labs zones.

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

The bridge to possible

Thank you

cisco Live!


CISCO Live!
ALL IN

Glossary

1R2C	One rate two color
2R3C	Two rate 3 color
802.1Q	An IEEE [Institute of Electrical and Electronics Engineers] standard
AAA	Authentication, Authorization, and Accounting
AAL5	ATM Adaptation Layer 5
AC	Attachment Circuit
ACE	Access Control Entry
ACL	Access Control List
ADJ	Adjacency
ARP	Address Resolution Protocol
ASIC	Application-Specific Integrated Circuit
ATM	Asynchronous Transfer Mode
B	Byte
bc	Burst committed
Bcast	Broadcast
BD	Bridge Domain
be	Burst excess
BFD	Bidirectional Forwarding Detection
BGP	Border Gateway Protocol
BPDU	Bridge Protocol Data Unit
BVI	Bridge Virtual Interface
CAM	Content Addressable Memory
CBWFQ	Class-Based Weighted Fair Queuing
CDP	Cisco Discovery Protocol
CEF	Cisco Express Forwarding
cir	Committed information rate

CLNSE	Connectionless Network Service
CoS	Class of Service
CoPP	Control Plane Policing
CPU	Central Processing Unit
CRC	Cyclic Redundancy Check
CSC	Carrier Supporting Carrier
DBUS	Data bus
dCEF	Distributed Cisco Express Forwarding
DB	Database
DCI	Data Center Interconnect
DDR	Double Data Rate
DFC	Distributed Forwarding Card
DoS	Denial of Service
DRAM	Dynamic Random Access Memory
DSCP	Differentiated Services Code Point
DTP	Dynamic Trunking Protocol
DWDM	Dense Wavelength Division Multiplexing
EFD	Early Fast Discard
EFP	Ethernet Flow Point
EIGRP	enhanced Internal Gateway Routing Protocol
ELAM	Embedded Logic Analyzer Module
EOBC	Ethernet Out of Band Channel
EoMPLS	Ethernet over Multiprotocol Label Switching
eq	Equal
ES+	Ethernet Services Plus
ESI	Ethernet Segment Identity

Glossary

EVI	EVPN Instance
FIA	Fabric Interface ASIC
FPD	Field Programmable Device
FPGA	Field Programmable Gate Array
FW	Firmware
Gbits	Gigabits
Gbps	Gigabit per second
GByte	Gigabyte
GE	Gigabit Ethernet
GHz	Gigahertz
GPB	Google Protocol Buffer
GRE	Generic routing Encapsulation
HA	High Availability
HbH	Hop by Hop
HBM	High Bandwidth Memory
HSRP	Hot Standby Router Protocol
H/W	Hardware
H-QoS	Hierarchical Quality of Service
ICMP	Internet Control Message Protocol
ID	Identity
IDS	Intrusion Detection system
IFIB	Internal FIB [Forwarding Information Base]
IOS	Internet Operating system
IP	Internet Protocol
IPCP	IP [Internet Protocol] Control Protocol [Part of PPP]

IPSec	Internet Protocol Security
IPv4	Internet Protocol version 4
IPv6	Internet Protocol version 6
IPV6CP	IPv6 [Internet Protocol version 6] control Protocol [Part of PPP}
IRB	Integrated Routing & Bridging
ISIS	Intermediate System - Intermediate System [Internal gateway routing protocol]
L2	Layer 2 [OSI Open Systems Interconnection] OR Level 2
L2PT	Layer 2 Protcol Tunneling
L2VPN	Layer 2 VPN [Virtual Private Network]
L3	Layer 3 [OSI Open Systems Interconnection] or Level 3
L4	Layer 4 [OSI Open Systems Interconnection] or Level 4
LACP	Link Aggregation Control Protocol
LAN	Local Area Network
LC	Line Card
LCDBUS	Line Card Data Bus
LCP	Link Control Protocol [Part of PPP]
LCRBUS	Line Card Results Bus
LDP	Label Distribution Protocol
LER	Label Edge Router
LFI	Link Fragment Interleave
LFIB	Label Forwarding Information Base
LLQ	Low Latency Queue
LPTS	Local Packet Transport Services
LSP	Label Switched Path
LSR	Label Switching Router

Glossary

MAC	Media Access Control
Mbps	Megabits per second
MByte	Megabyte
Mcast	Multicast
MET	Multicast Expansion Table
MHz	Megahertz
MIB	Management Information Base
MIPS	Multiprocessor without Interlock Pipeline Stages
Mod	Modulo
Mpps	Megapackets per second
MPLS	Multiprotocol Label Switching
MPLS-TP	Multiprotocol Label Switching - Transport Profile
MPP	Management Plane Protection
MQC	Modular Quality of service Command line interface
MSDP	Multicast Source Discovery Protocol
MSFC	Multilayer Switch Feature Card
MSS	Maximum Segment Size (TCP)
MTU	Maximum Transmission Unit
MUX	Multiplexer
NAT	Network Address Translation
ND	Neighbor Discovery [protocol]
NP	Network Processor
NPU	Network Processor Unit
NSF	Non-Stop forwarding
NTP	Network Time Protocol

NVRAM	Non-Volatile Random Access Memory
OSM	Optical Services Module
OSPF	Open Shortest Path First [protocol]
PA	Port Adapter
PAgP	Port Aggregation Protocol
PDU	Protocol Data Unit
PFC	Policy Feature Card
PFM	Platform Fault Manager
PHP	Penultimate Hop Popping
PIFIB	Pre-IFIB [Internal Forwarding Information Base]
PLU	Packet Lookup Unit
PoP	Point of Presence
POS	Packet Over sonet
PPP	Point to Point Protocol
PPS	Packets Per Second
PSIRT	Product Security Incident Reponse Team [Cisco]
PW	Pseudo-Wire
QoS	Quality of Service
RADIUS	Remote Authentication Dial In Service [protocol]
RARP	Reverse ARP [Address Resolution Protocol]
RBUS	Results bus
RIP	Routing Information Protocol
RJ45	An 8 wire wiring standard
RP	Routing Processor OR Route Processor
RPF	Reverse Path Forwarding

Glossary

RSP	Routing and Switching Processor OR Route Switch Processor [Cisco]
RSVP	Resource reservation protocol
RTBH	Remote Triggered Black Holing
SCP	Secure Copy
SDRAM	Synchronous Dynamic Random Access Memory
SFP	Small Form-factor Pluggable
Sh	Shaper
SIP	Shared Port Adapter Interface Processor [Cisco]
SNMP	Simple Network Management Protocol
SP	Service Provider OR Switching Processor
SPA	Shared Port Adapter
SRAM	Static Random Access Memory
SSH	Secure Shell [protocol]
SSO	Stateful Switch Over
SSRAM	Synchronous Static Random Access Memory
SUP	Supervisor [Cisco]
SW	Switching
TAC	Technical Assistance Center [Cisco]
TACACS	Terminal access Control Access-Control System [protocol]
TCAM	Tertiary Content Addressable Memory
TCB	Transmission Control Block
TCL	Tool Command Language
TCP	Transmission Control Protocol
TDM	Time Division Multiplexing
TTL	Time To Live

tx	Transmit
uC	Microcontroller
UDLD	Unidirectional Link Detection
uRPF	Unicast Reverse Path Forwarding
VACL	VLAN [Virtual Local Access Network] Access control List
VLAN	Virtual Local Access Network
VOQ	Virtual Output Queueing
VPLS	Virtual Private LAN [Local Access Network] Service
VPN	Virtual Private Network
VQI	Virtual Queue Identifier
VRF	Virtual Routing and Forwarding
VRRP	Virtual Router Redundancy Protocol
VTP	Virtual Trunking Protocol
VTY	Virtual Terminal line
WAN	Wide Area Network
WFQ	Weighted Fair Queuing [Cisco]
WRR	Weighted Round Robbin
XML	Extensible Markup Language


CISCO Live!
ALL IN