Cisco live!

January 29 – February 2, 2018 · Barcelona

BRKARC-2003

Cisco ASR 9000 System Architecture

Yongzhong Peng

Manager, Technical Marketing

SP Routing Infrastructure

Cisco live!

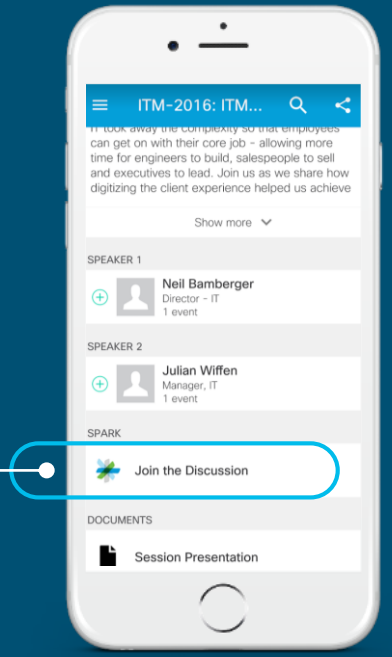
Cisco Spark

Questions?

Use Cisco Spark to communicate with the speaker after the session

How

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



cs.co/ciscolivebot#BRKARC-2003

Agenda

- ASR9000 Products Introduction
- ASR9000 System Hardware Architecture
- ASR9000 Distributed Control Plane
- ASR9000 Data Packet Processing
- ASR9000 QoS Architecture
- IOS-XR & IOS-XR 64 Bit
- Conclusion

ASR 9000 Products Introduction

Cisco ASR 9000 System Comprehensive Portfolio

Compact & Powerful Access/Aggregation	Flexible Service Edge	High Density Service Edge and Core
• Small footprint with full IOS-XR for distributed environments	• Optimized for ESE and MSE with high M-D scale for medium to large sites	• Scalable, ultra high density service routers for large, high-growth sites

One Platform, One OS, One Family

nV Satellites
ASR 9000v, NCS5000

ASR 9901

ASR 9001

Fixed 2RU
240 Gbps

ASR 9904

2 LC/6RU
16 Tbps

ASR 9006

4 LC/10RU
7 Tbps

ASR 9906

4 LC/14RU
32 Tbps

ASR 9010

8 LC/21RU
14 Tbps

ASR 9910

8 LC/21RU
64 Tbps

ASR 9912

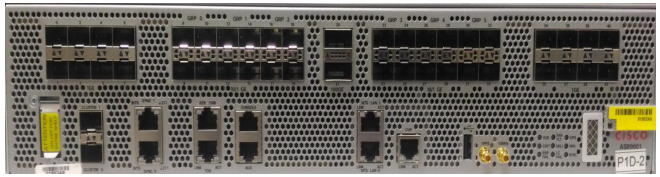
10 LC/30RU
80 Tbps

ASR 9922

20 LC/44RU
160 Tbps

MSE	E-MSE	Peering	P/PE	CE	Mobility	Broadband

ASR 9901 Highlights



Form Factor & BW

- ***2 RU*** *box with 2 Tomahawk NPU (ASR 9001 is 2 RU with 2 Typhoons)*
- *Depth of ~23 inches,(9001 is 18")*
- *456G Duplex BW (9001 is 120G Duplex)*

Ports / Port Density

- *Fixed ports available;* ***no MPAs***
- *42 ports on the faceplate : 16X1G, 24X1/10G, 2X100G(QSFP28)*
- *1G ports : LAN & MACSEC*
  10G ports: LAN & MACSEC
  100G port : LAN & MACSEC

SW & Licensing

- ***64 bit XR only***
- *PAYG mode for 120G,240G,360G and 456G*
- *Full feature parity with Tomahawk feature-set*

Mechanicals & Commons

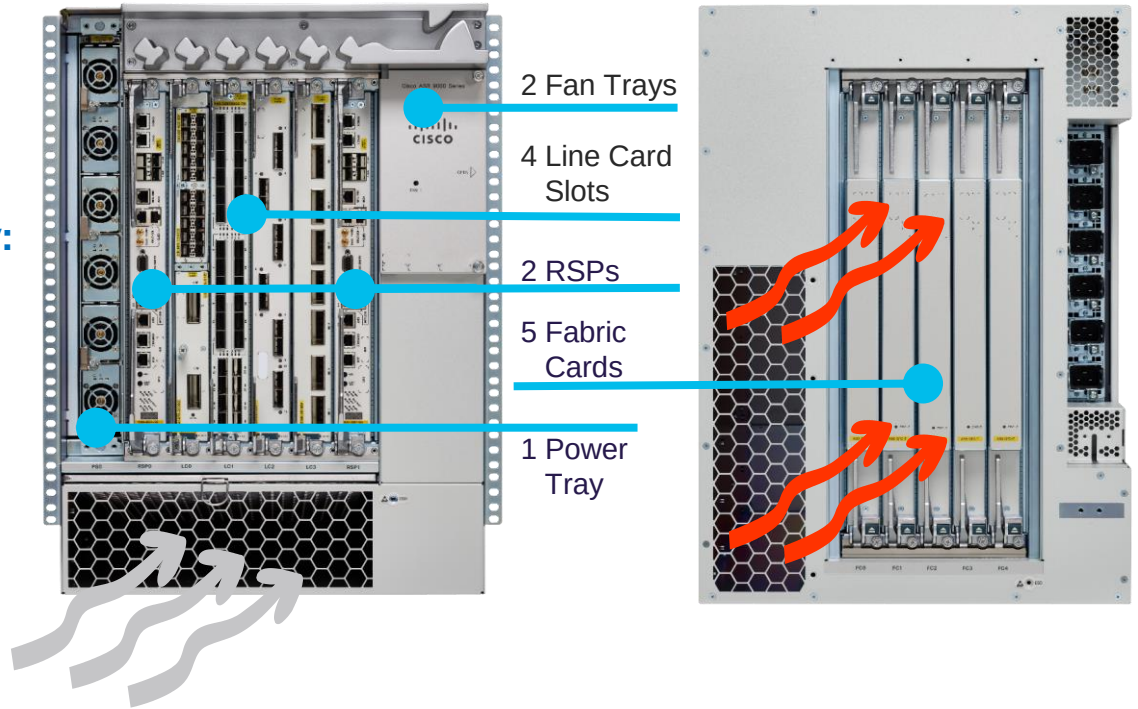
- *Redundant Power & Fan-trays*
- *Front to back Airflow*
- *NEBS, EMC Compliant*
- *All ports/power cabling on front plate; fan trays on backside*
- *Typical Power Consumed : 1200W*

ASR 9901 Port mapping

				GRP 0		GRP 1		GRP 2				GRP 3		GRP 4		GRP 5							
0	2	4	6	8	10	12	14	16	18	20		22	24	26	28	30	32	34	36	38	40		
1	3	5	7	9	11	13	15	17	19	21		23	25	27	29	31	33	35	37	39	41		
1GE				10/1 GE						100GE		10/1 GE						1GE					

Cisco ASR 9906 Overview

- ✓ **Scalable for Dense 100GE Applications** for future investment protection
- ✓ **Greater Capacity, Greater Flexibility:** Start with 2 RSPs, then add fabric cards to scale beyond 460G and enable N+1 fabric redundancy
- ✓ **Mid-Plane design** allows for compact chassis footprint



RSP880-LT

- RSP880-LT is the 3rd generation route processor card
  - 880Gbps/slot throughput (redundant configuration)
  - 4 Core Intel Broadwell1.8GHz processor
  - Available in both TR (16GB)/ SE (32GB) variants
  - Line Cards Supported
    - Tomahawk
    - Typhoon
    - VSM
    - SIP-700
  - Chassis Supported
    - 9006/9010/9910/9906/9904
  - **Feature parity cXR with RSP880 at FCS (eXR support planned for 6.4.x)**
  - Similar Control Plane Scale to RSP880
  - **No support for nV-Cluster and 3rd party applications**

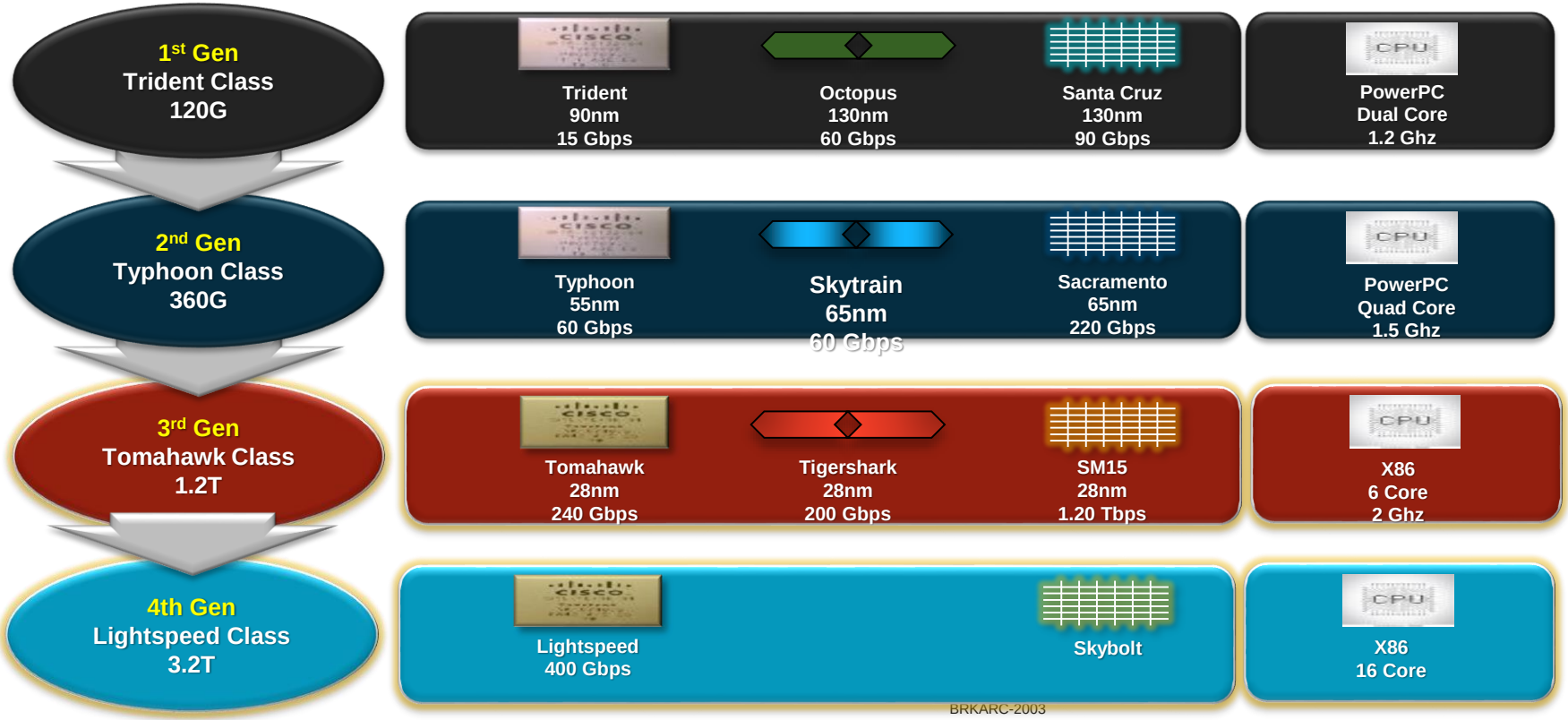
RSP880-LT provides is a cost effective migration path for RSP440 customers.

Hardware Differences	
RSP880	RSP880-LT
8 Cores, 1.9Ghz	4 Cores, 2.4Ghz
4 SFP+ external ports	No external SFP+ ports
2x32GB SSD	2x128GB SSD

Comparison of RSP880-LT to RSP880

	RSP880-LT	RSP880/A99-RSP
Bandwidth/Slot	880Gbps (Dual RSP)***	880Gbps (Dual RSP)
RIB Scale**	26M/27M (IPv4/IPv6)	26M/27M (IPv4/IPv6)
RIB Learning Rate	95K IPv4 Routes/sec	78K IPv4 Routes/sec
Platforms Supported	ASR 9904, 9006, 9010, 9910, 9906	**RSP880:** ASR 9904, 9006, 9010 **A99-RSP:** ASR 9910, 9906
3rd Party Application Support	Not Supported	Supported
eXR Support	Planned for support in 6.4.x	Supported

ASR 9000 Linecards Evolution



BRKARC-2003

ASR 9K Ethernet Line Card Overview

-TR, -SE

2nd LC Typhoon
NPU: 60Gbps, ~45Mpps

A9K-36x10GE A9K-2x100GE A9K-24x10GE

A9K-40GE
A9K-4T16GE

A9K-MOD80 A9K-MOD160

MPAs
20x1GE
2x10GE
4x10GE
1x40GE
2x40GE

3rd LC Tomahawk
NPU: 240Gbps, ~150Mpps

A9K-4x100GE A9K-8x100GE A9K-12x100GE

MOD400/MOD200 24/48x10/1G LC A9K-400G-DWDM

MPAs
1x100GE
2x100GE
20x10GE
+
Typhoon MPAs

Cost Optimized 4x100GE LAN Line card

HW Specs	• Tomahawk based, 5-fabric linecard with: 2x NPU and 5Mbit internal TCAM • Supported on all ASR 90xx and 99xx systems
Features	• Parity with Tomahawk TR linecards at FCS except for Timings features (post FCS) • Features not supported: VidMon, Cluster, LISP, NSH, BNG, Satellite, MPLS-TP
Scale	• Limited scale for TCAM dependent features; rest of the scale on-par with other Tomahawk –TR cards
SW support	• cXR: 6.2.3, 6.3.2, 6.4.1 and onwards. SMU option possible for 6.1.4, 6.3.1, subject to business case • eXR: 6.4.1

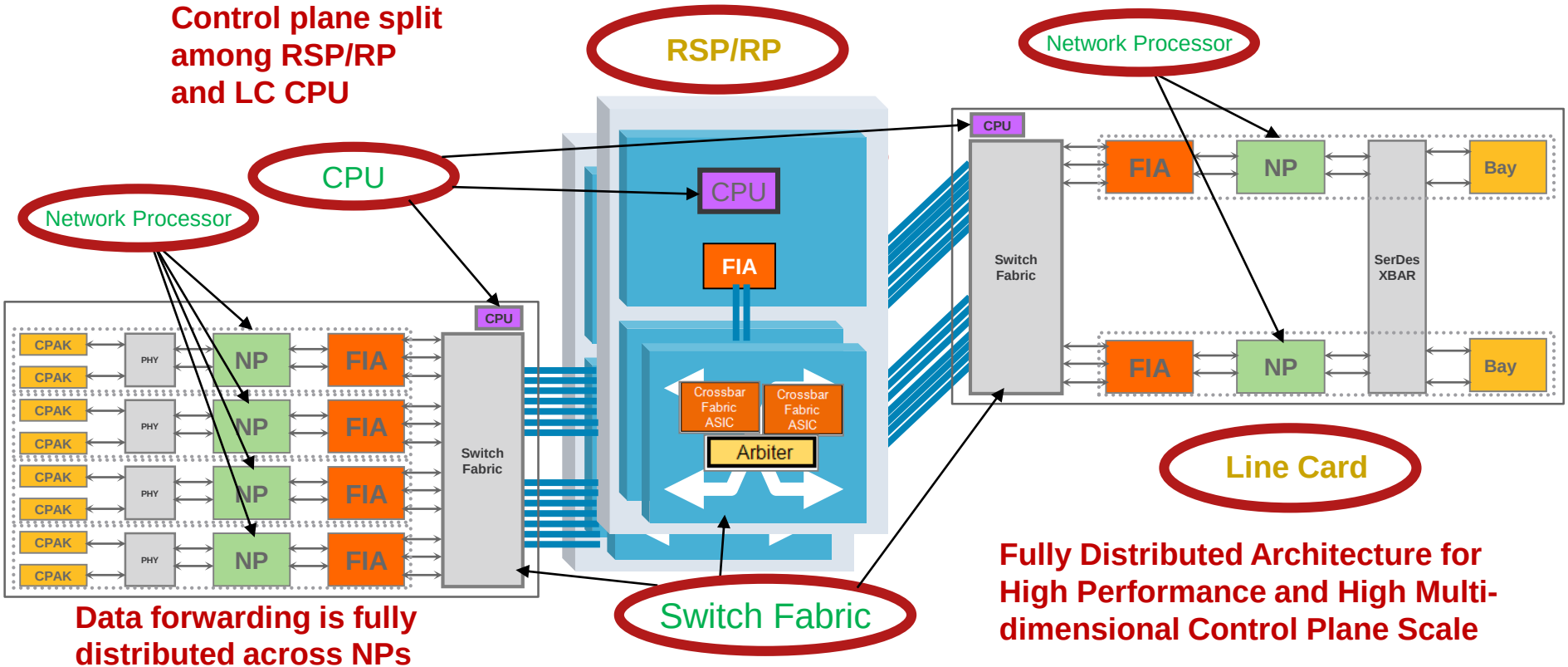
Pop Quiz ????

Which of the following RSP is supported in all of ASR9006, ASR9010, ASR9904, ASR9906 and ASR9910?

- **RSP880-LT**
- **RSP880**

ASR 9000 Hardware Architecture

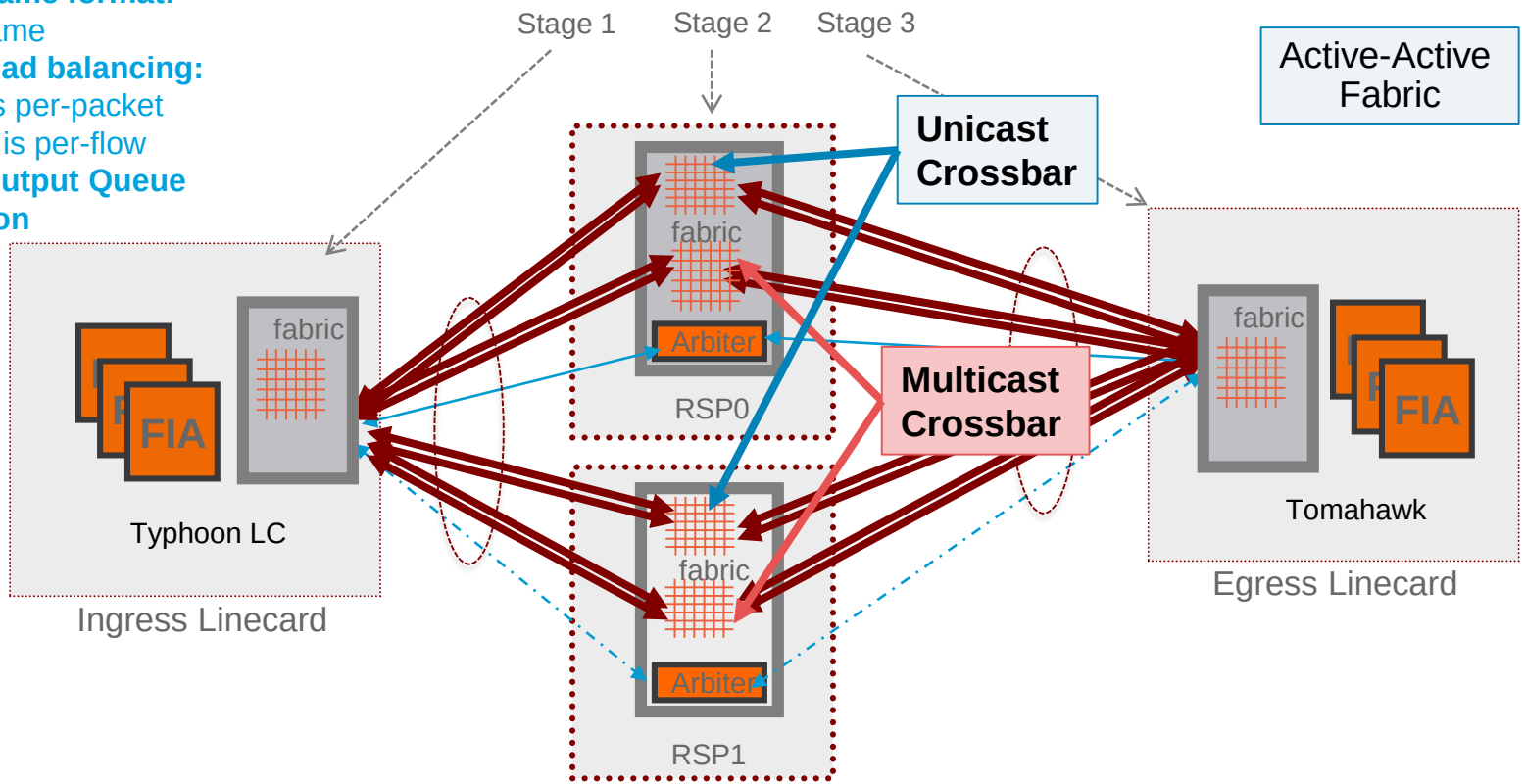
ASR 9000 System Architecture "At-a-Glance"



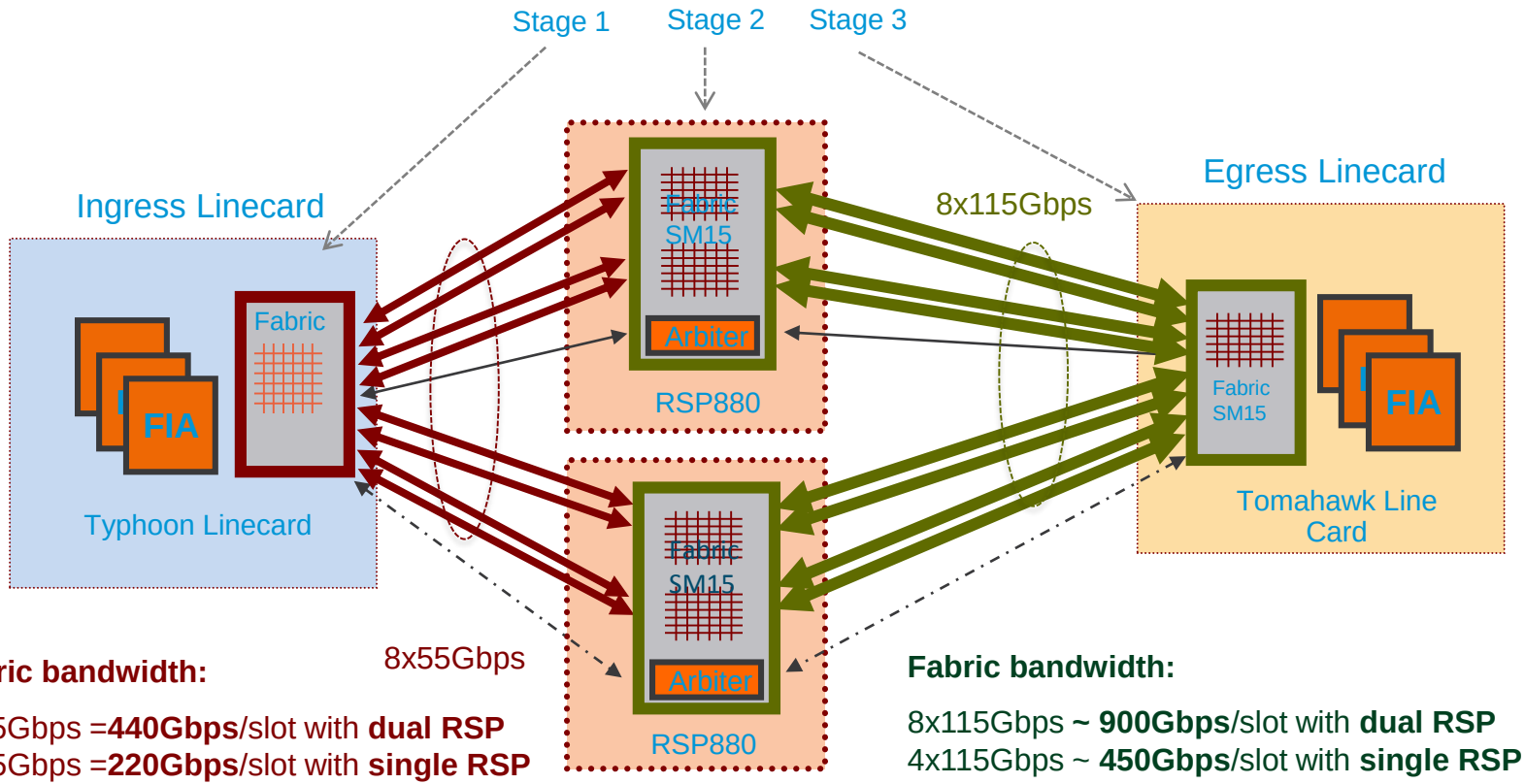
ASR 9000 Switch Fabric High-Level Architecture

3-Stage Non-Blocking Fabric (Separate Unicast and Multicast Crossbars)

Fabric frame format:
Super-frame
Fabric load balancing:
Unicast is per-packet
Multicast is per-flow
Virtual Output Queue
Arbitration



ASR90xx – RSP880 and Mixed LC



Fabric bandwidth:

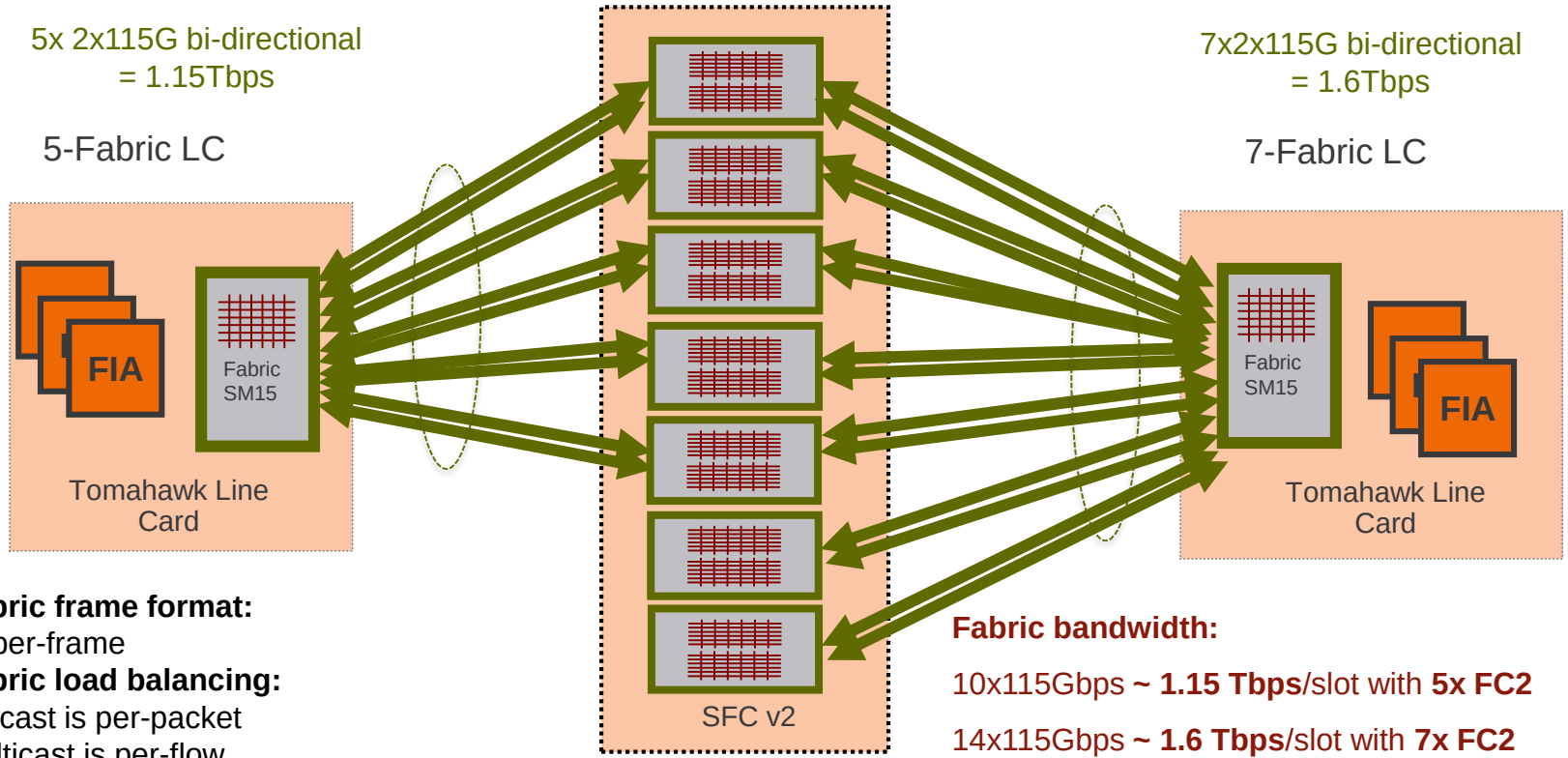
8x55Gbps =**440Gbps**/slot with **dual RSP**
4x55Gbps =**220Gbps**/slot with **single RSP**

Fabric bandwidth:

8x115Gbps **~ 900Gbps**/slot with **dual RSP**
4x115Gbps ~ **450Gbps**/slot with **single RSP**

ASR99xx Switch Fabric Card (FC2)

6+1 All Active 3-Stage Fabric Planes, Scale to 1.6Tbps/LC, Separated Ucast/Mcast Crossbars



Fabric frame format:
Super-frame
Fabric load balancing:
Unicast is per-packet
Multicast is per-flow

Fabric bandwidth:

10x115Gbps **~ 1.15 Tbps**/slot with **5x FC2**

14x115Gbps **~ 1.6 Tbps**/slot with **7x FC2**

Tomahawk LC: 8x100GE Architecture

Slice Based Architecture

CPAK: 100G, 40G, 10G

Macsec Suite B+, G.709, OTN, Clocking

L2/L3/L4 lookups, all VPN types, all feature processing, mcast replication, QoS/Queuing, ACL, etc …

VoQ buffering, Fabric credits, mcast hashing, scheduler for fabric and egress port

FPOE, Auto-Spread, DWRR, RBH, replication

LC CPU

PHY

Tomahawk NP

240G

FIA

240G

XBAR

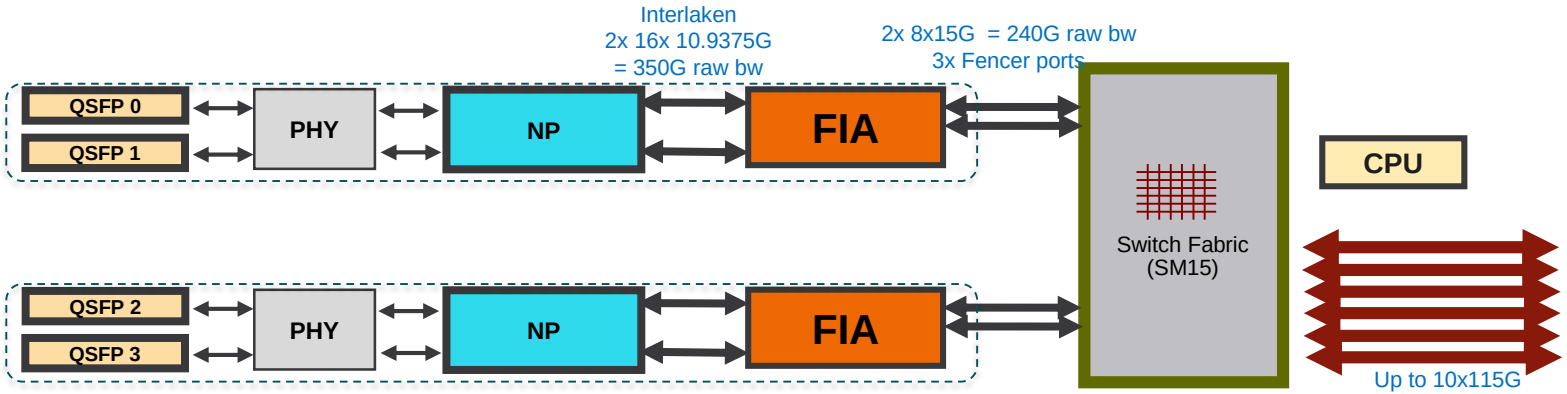
Up to 14x115G

Separated Switch Fabric

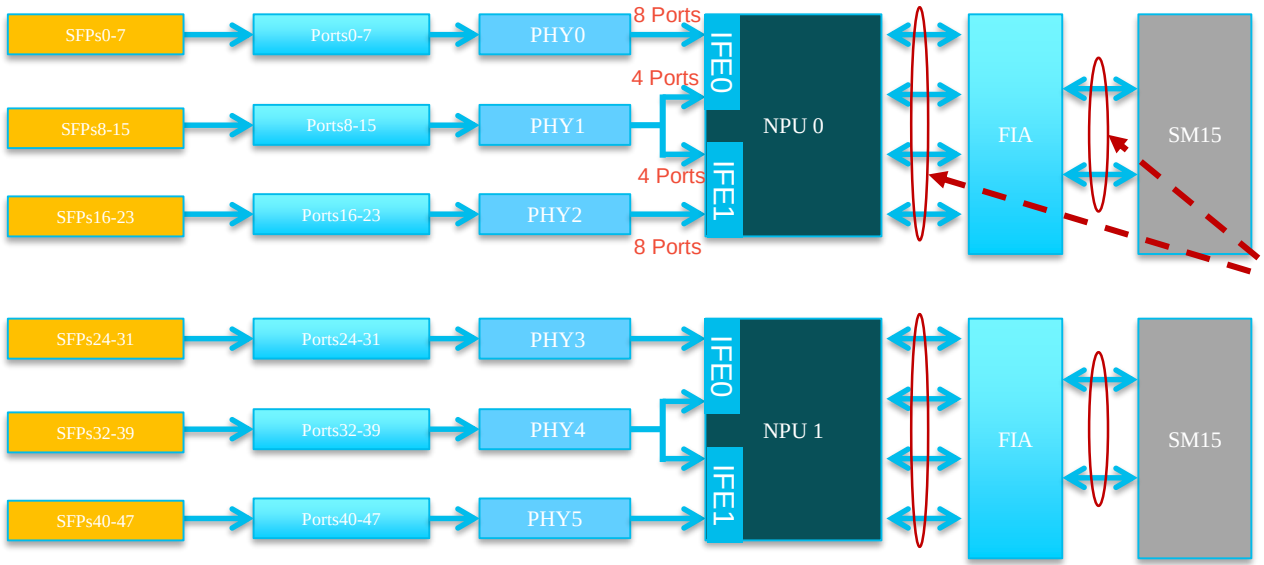
Per Slice Power Management: (100-200W Power Savings)

PE1(admin-config)# hw-module power saving location ? slice [0-3]

4x100GE LAN-based linecard architecture



A9K-48x10G-1G Architecture

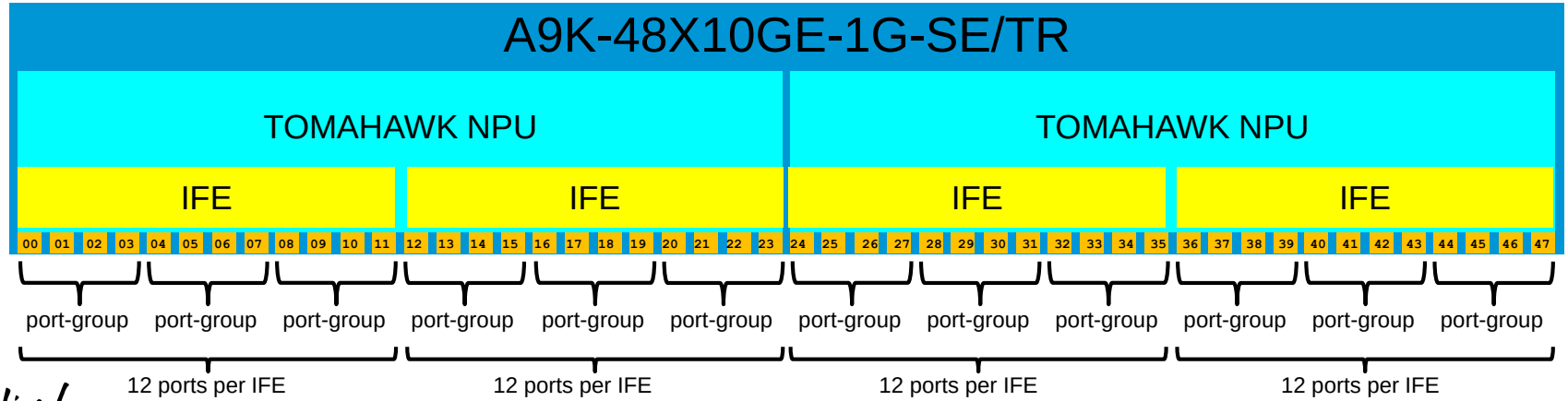


- 240G Raw Bandwidth/ 200G Data Traffic Bandwidth
- Potentially Oversubscribed

- Supports all 1G and 10G. 1G and 10G are configurable follow certain rule
- 48-port has a two NPU slices with shared 80Mb TCAM
- Phy does not support OTN Framing or MACSec

A9K-48x10G-1G Port Organization

- A9K-48X10GE-1G-SE/TR has two NPU slices, A9K-24X10GE-1G-SE/TR has one
- Each NPU slice consists of two Interface Engines (IFE)
- Interface Ports are organized in port-groups
- Each IFE consists of three port-groups
- Each port-group consists of 4 interface ports



A9K-48x10-1G Port Speed Rules

- Interface Port Speed is configurable: 1G / 10G
- Per default, all ports are 10G
- The 4 Rules:
  1. All ports in a port-group (group of four ports) have to have the same speed.
  2. If the first port-group on an IFE (group of three port-groups) is configured for 1G, then all ports in that same IFE must be 1G.
  3. If the first port-group of an IFE is 10G, then the other two port-groups in that same IFE can be any combination of 1G or 10G.
  4. When configuring port speeds, all ports of a given linecard need to be configured together in one single CLI command.
- Deviations from these rules will results in a CLI rejection
- Changing port speed on an interface does not impact traffic on other interfaces

Possible Port Speed Combination per IFE

- 5 possible port speed combinations for each IFE
- IFE are independent of each other for port speed configuration

	Interface Engine (IFE)											
	P0	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
Configuration 1	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G	1G
Configuration 2	10G	10G	10G	10G	1G	1G	1G	1G	1G	1G	1G	1G
Configuration 3	10G	10G	10G	10G	10G	10G	10G	10G	1G	1G	1G	1G
Configuration 4	10G	10G	10G	10G	1G	1G	1G	1G	10G	10G	10G	10G
Configuration 5	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G	10G

Line Card Components

Runs distributed control plane protocols for increased scale

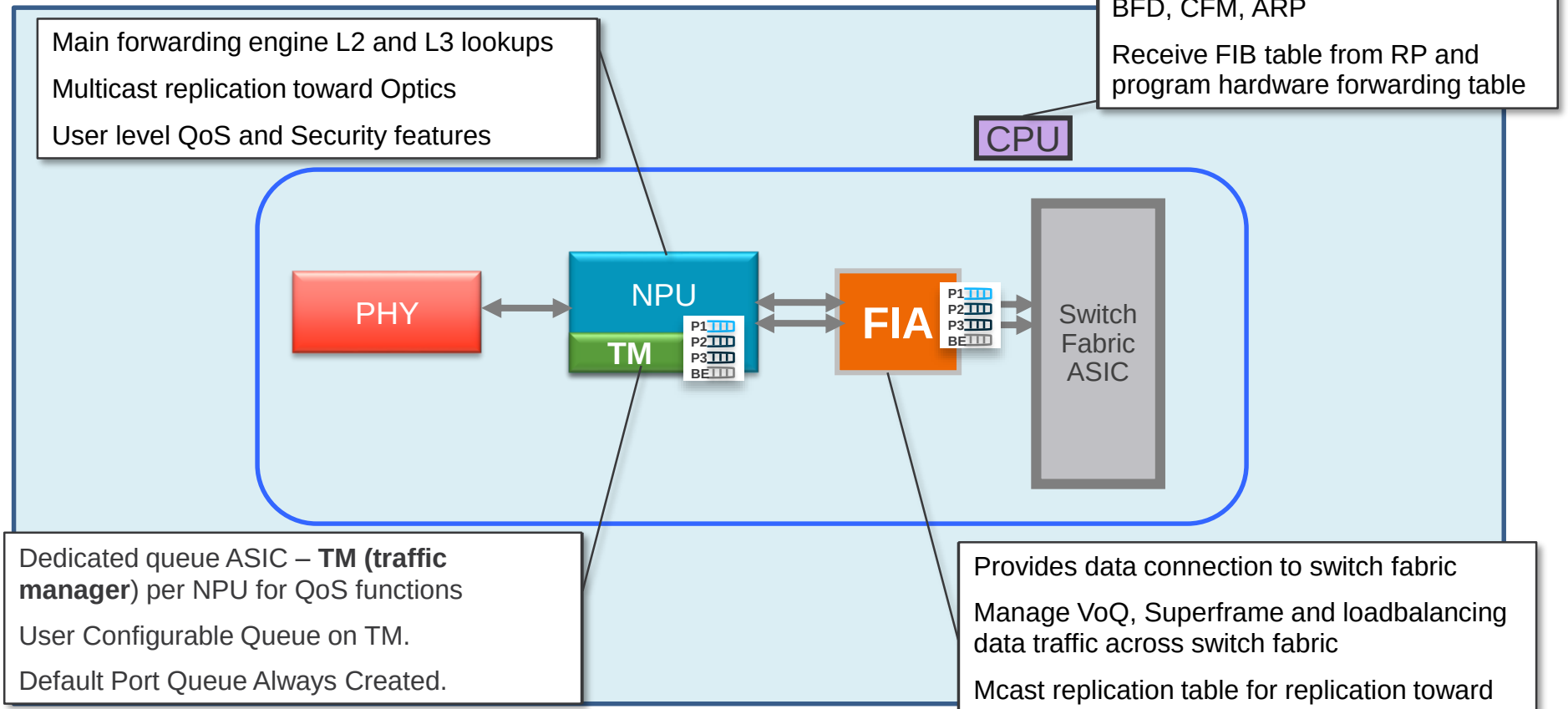
BFD, CFM, ARP

Receive FIB table from RP and program hardware forwarding table

Main forwarding engine L2 and L3 lookups

Multicast replication toward Optics

User level QoS and Security features



Dedicated queue ASIC – **TM (traffic manager**) per NPU for QoS functions

User Configurable Queue on TM.

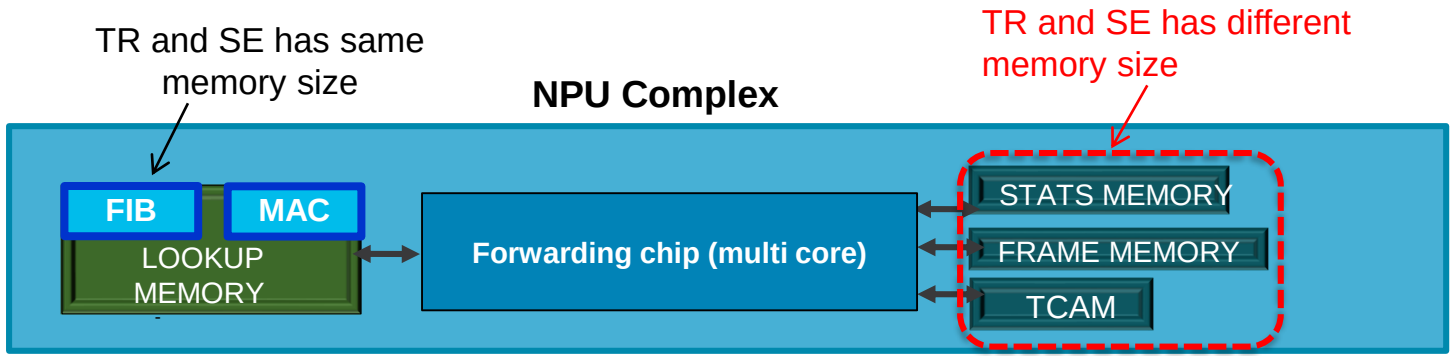
Default Port Queue Always Created.

Provides data connection to switch fabric

Manage VoQ, Superframe and loadbalancing data traffic across switch fabric

Mcast replication table for replication toward NPs

Network Processor Architecture Details



- TCAM: VLAN tag, QoS and ACL classification
- Stats memory: interface statistics, forwarding statistics etc
- Frame memory: buffer, Queues
- Lookup Memory: forwarding tables, FIB, MAC, ADJ
- TR/SE
  - Different TCAM/frame/stats memory size for different per-LC QoS, ACL, logical interface scale
  - Same lookup memory for same system wide scale mixing different variation of LCs doesn't impact system wide scale

-TR: transport optimized, -SE: Service edge optimized

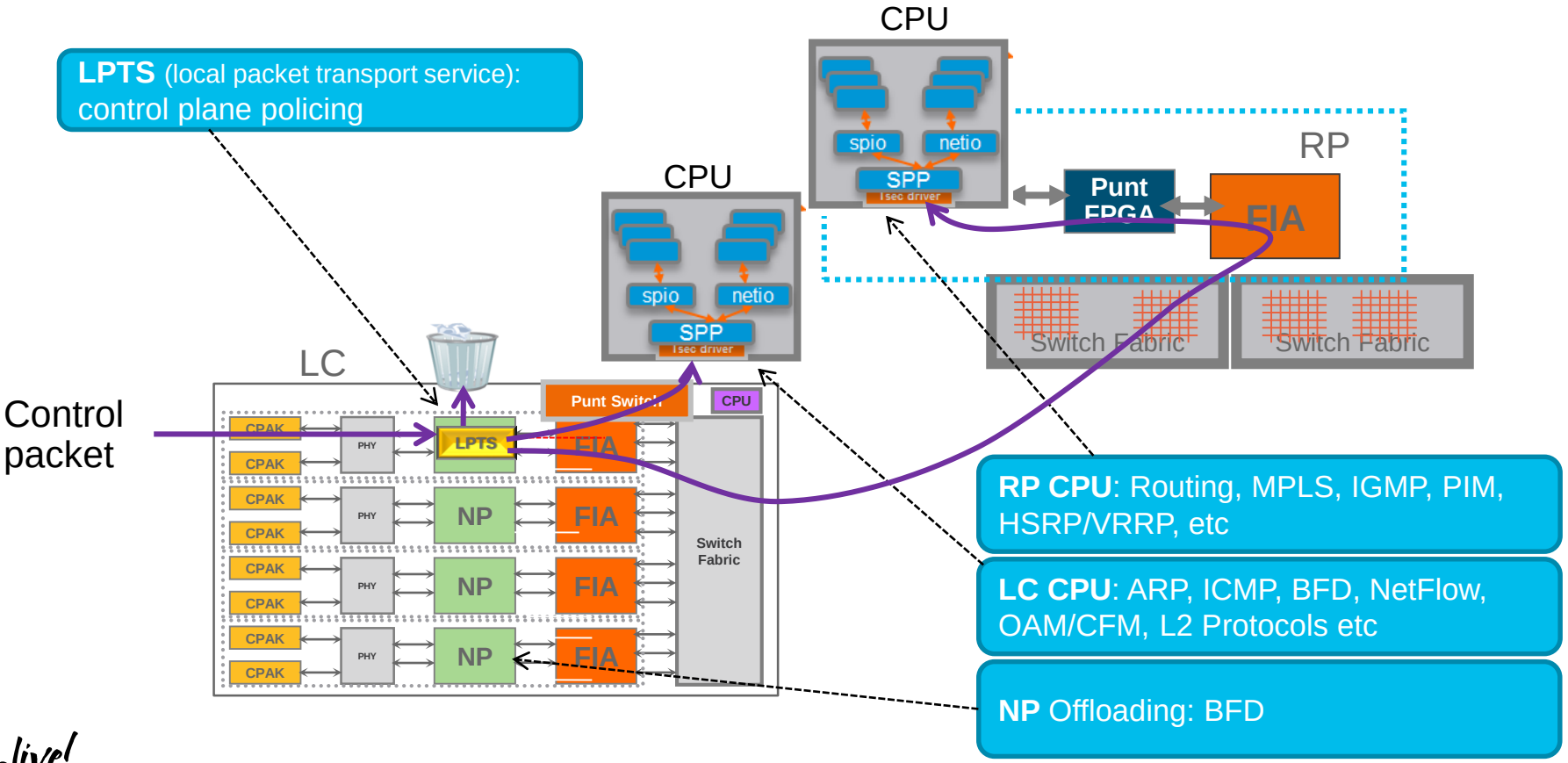
Pop Quiz ????

For the A9K-48x10G-1G linecard, if the first port-group of an IFE is configured as 1G, what is the possible port speed for the other two port-groups of the same IFE?

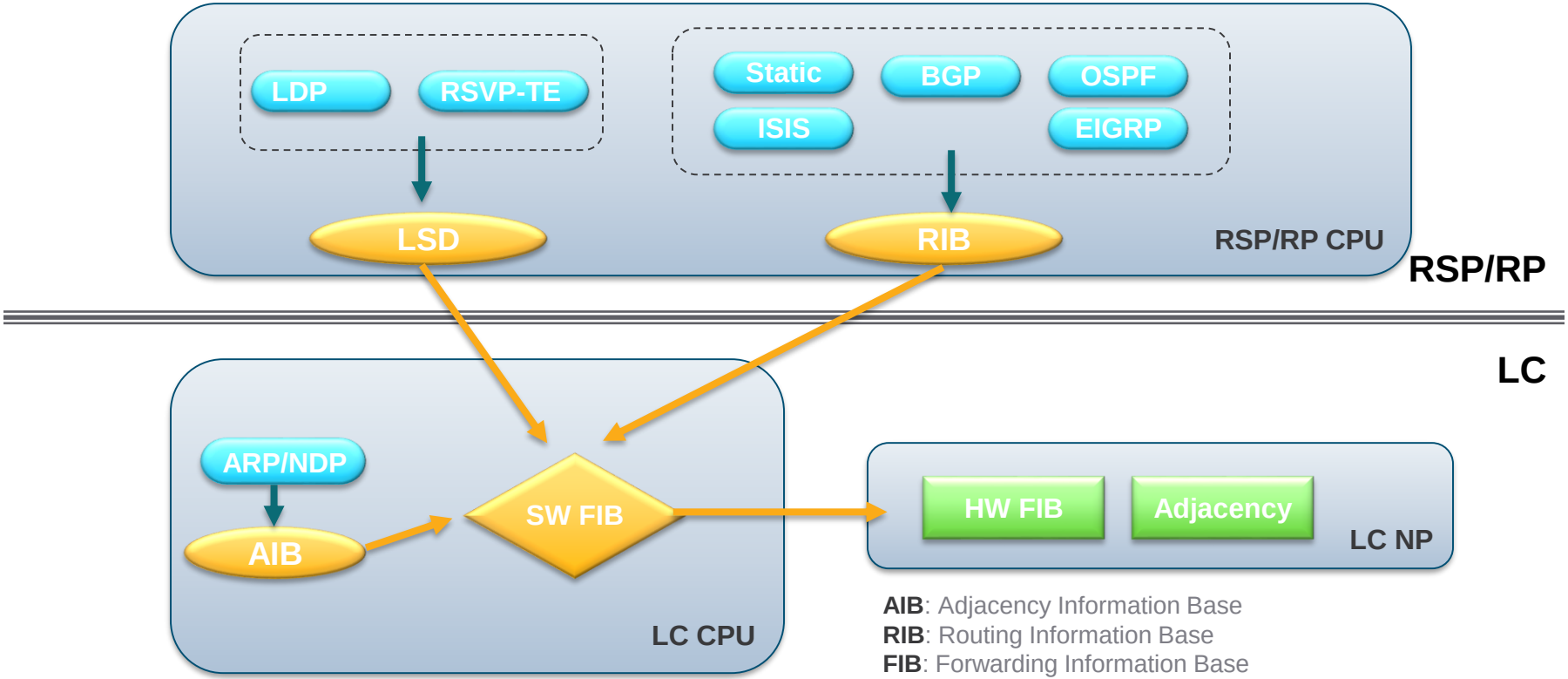
- **10G**
- **10G or 1G**
- **1G**

ASR 9000 Distributed Control Plane

ASR9000 Fully Distributed Control Plane

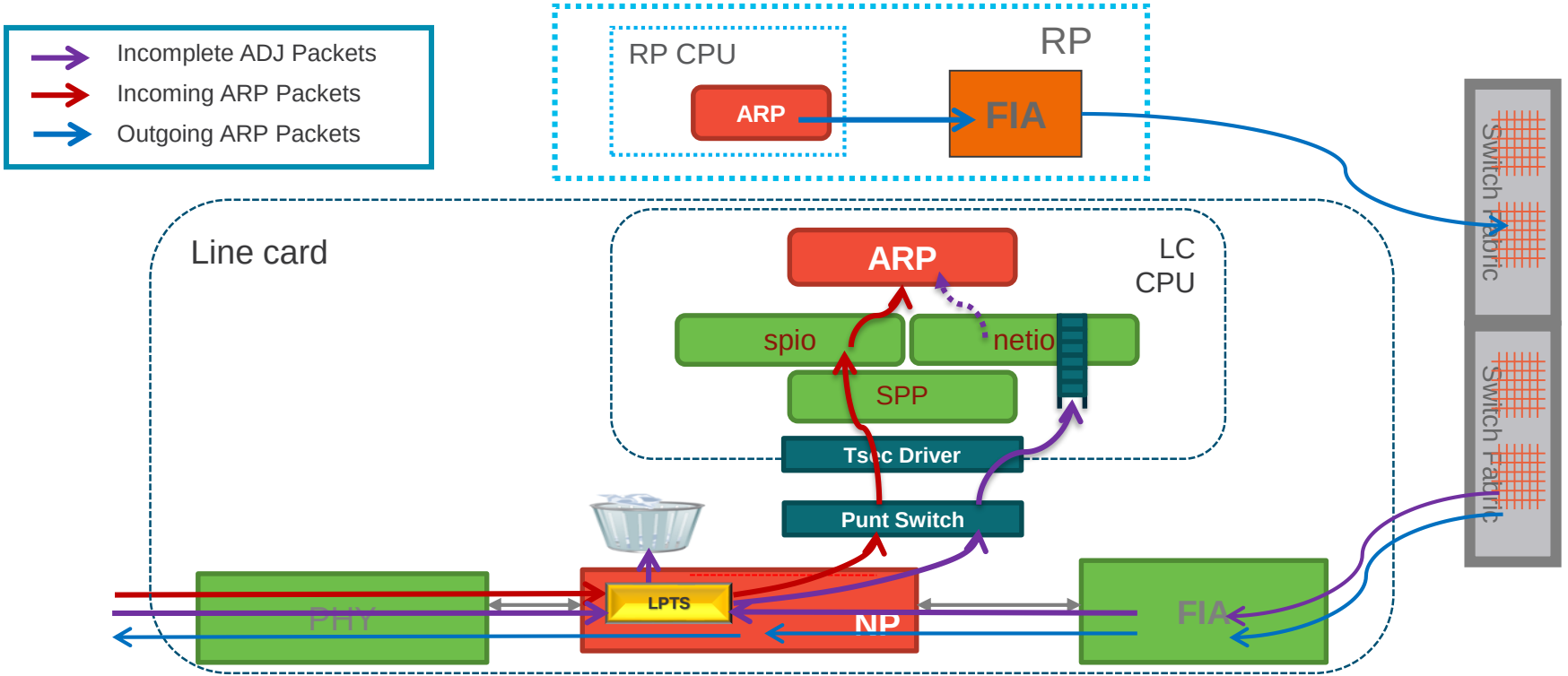


L3 Control Plane Architecture



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

Distributed ARP Processing

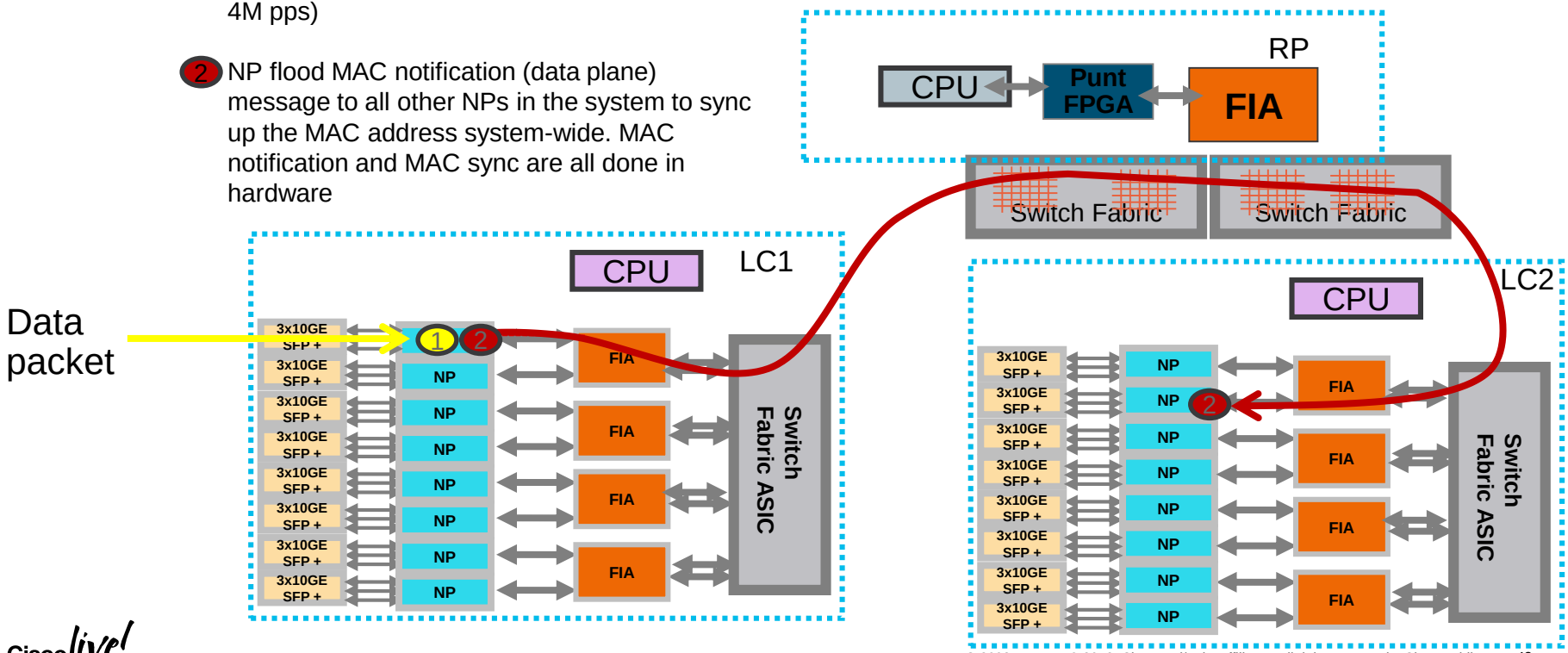


MAC Learning and Sync

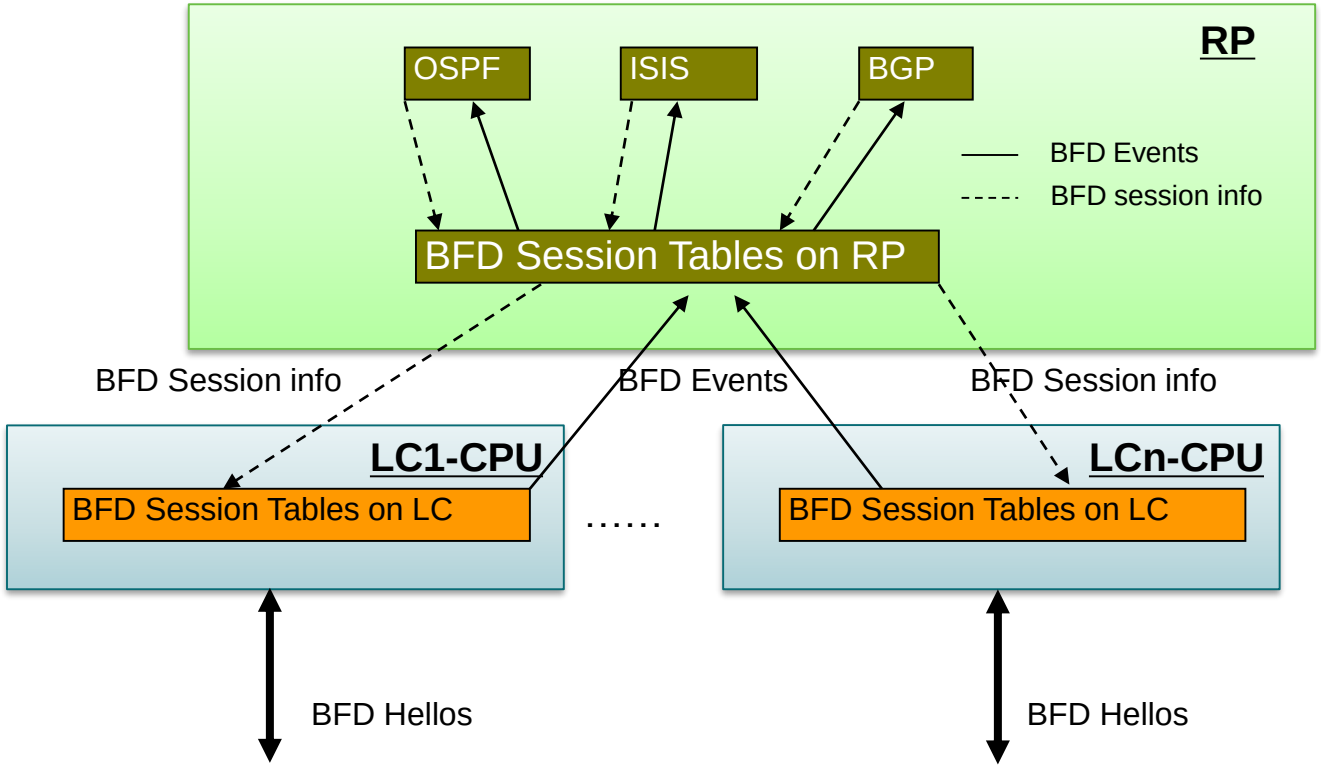
Hardware based MAC learning: ~4Mpps/NP

1. NP learn MAC address in hardware (around 4M pps)

2. NP flood MAC notification (data plane) message to all other NPs in the system to sync up the MAC address system-wide. MAC notification and MAC sync are all done in hardware

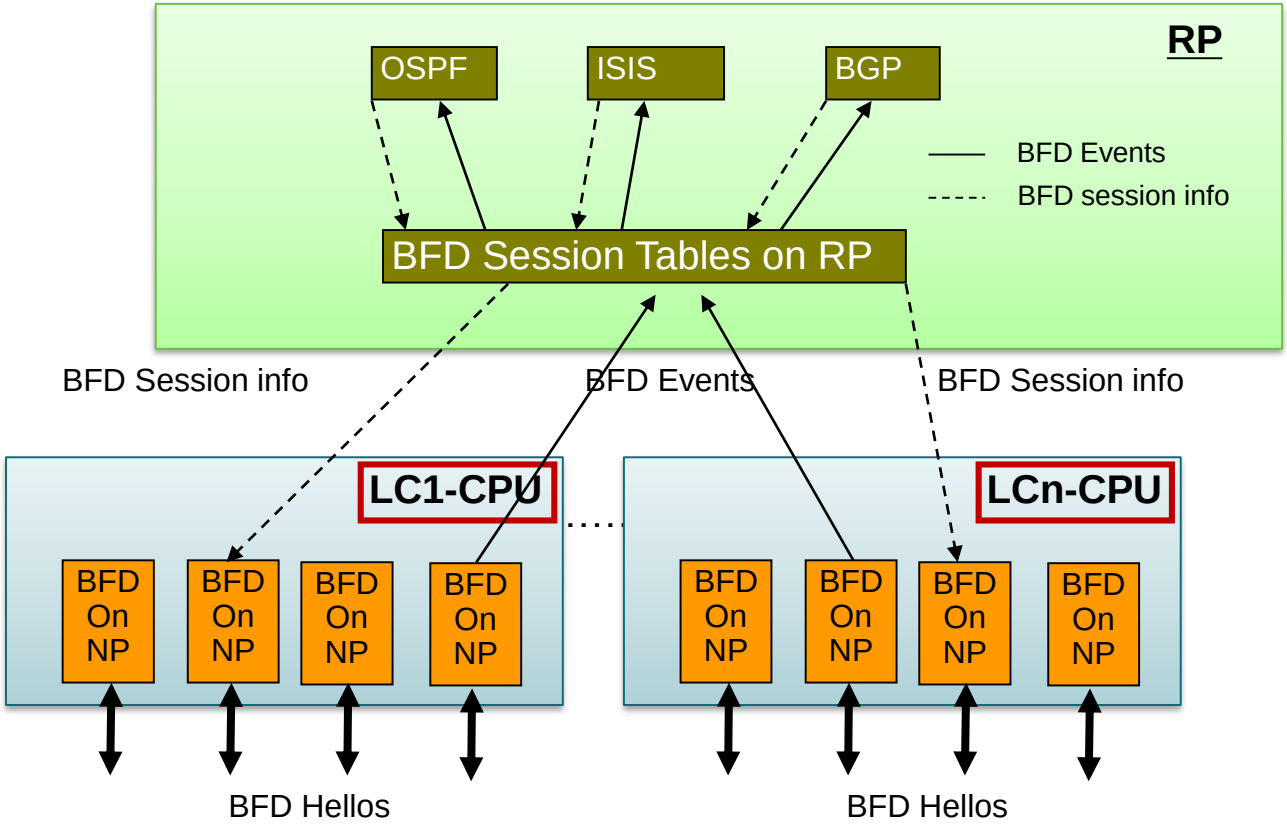


Distributed BFD Architecture

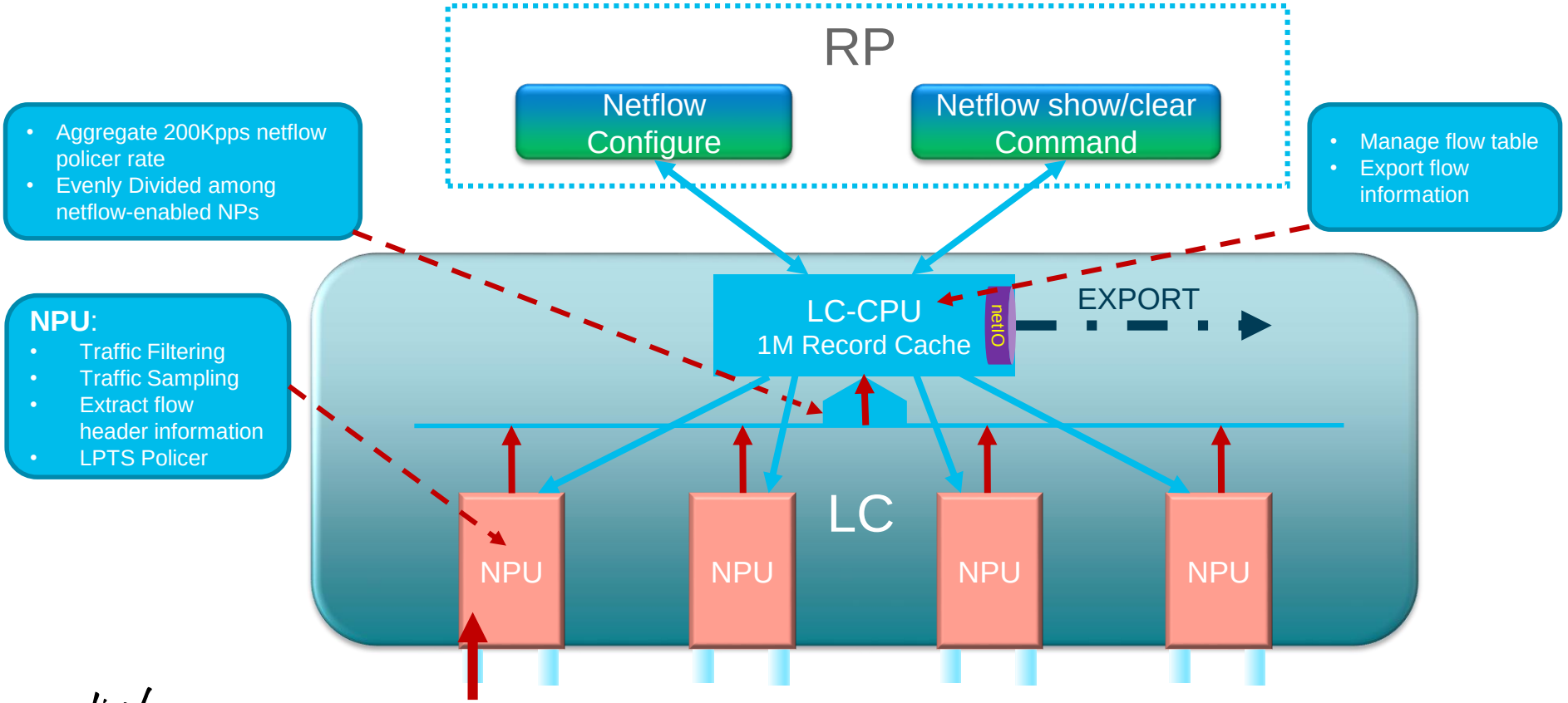


HW-offloaded BFD

```
hw-module bfd-hw-offload enable location 0/0/CPU0
```



Distributed Netflow Architecture



Pop Quiz ????

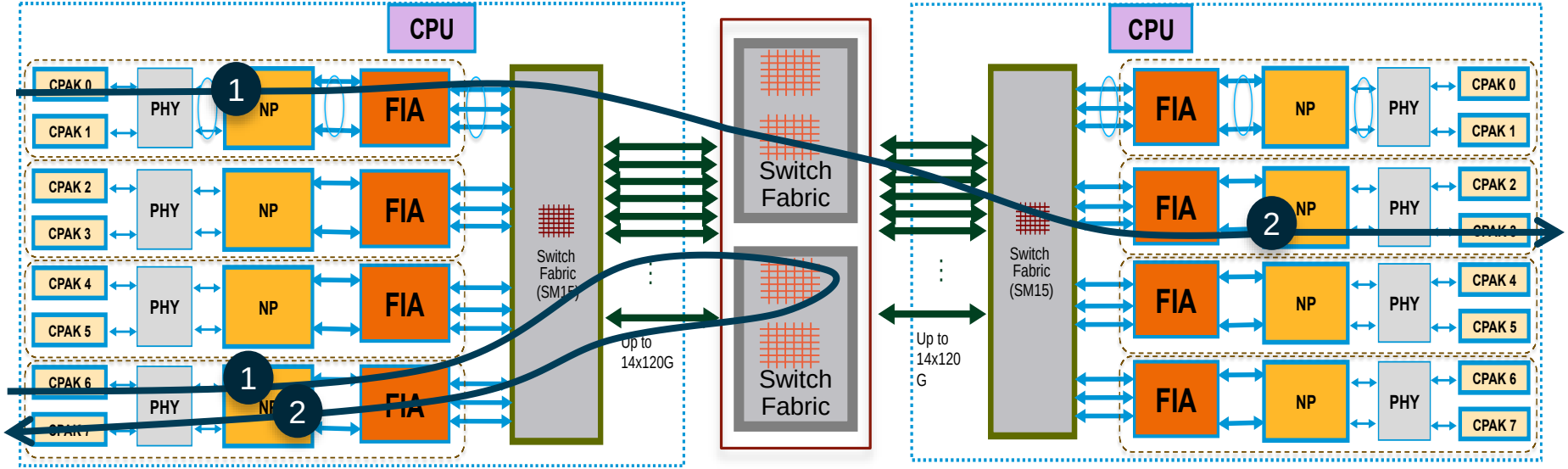
In an ASR9906 system which has 2 RSP, 4 linecards and 5 switch fabric cards, how many total CPUs are available to host ASR9000 control-plane functionalities?

- **2**
- **4**
- **6**
- **11**

ASR 9000 Data Packet Processing

Distributed Two-Stage Packet Processing

- Ingress lookup yields packet egress port and applies ingress features
- Egress lookup performs packet-rewrite and applies egress features



Uniform packet flow for simplicity and predictable performance

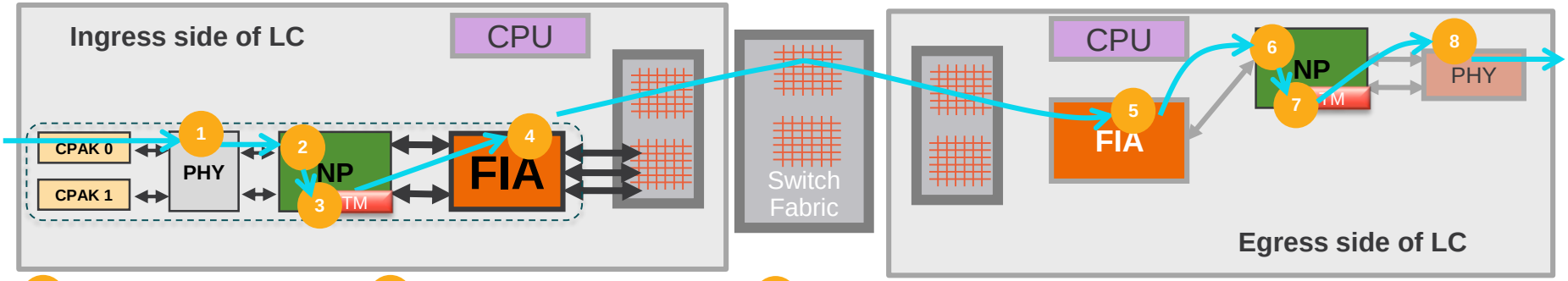
ASR9000 Life of a Packet – Tomahawk LC

2
- Ingress L2/L3 FIB lookup, ACL/QoS lookup
- Ingress PBR/ABR, ACL, uRPF
- Ingress QoS: classification, marking, policing
- Packet Punting
- Ingress ECMP/LAG hashing

4
- Buffering packet from NP
- Requesting fabric credit
- Manage superframe and load-balancing packet across fabric
- Manage system VoQ

6
- Egress L2/L3 FIB lookup, ACL/QoS lookup
- Egress PBR/ABR, ACL, uRPF
- Egress QoS: classification, marking, policing, shaping
- Incomplete Adj Packet Punting
- Egress ECMP/LAG hashing

8
- MACSEC Emcryption
- G.709/OTN/WAN-PHY/LAN-PHY
- Line Clocking



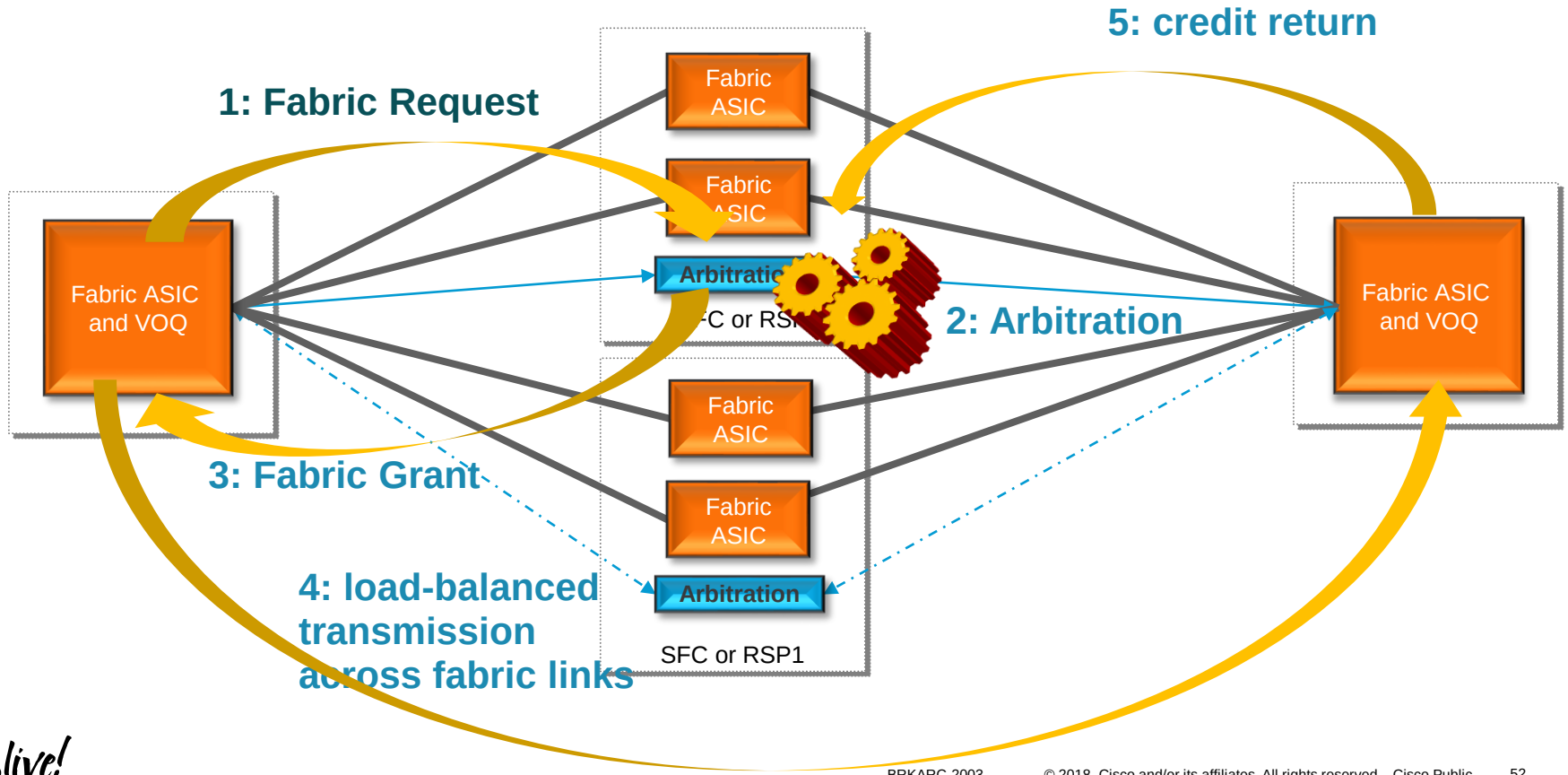
1
- MACSEC Decryption
- G.709/OTN/WAN-PHY/LAN-PHY
- Line Clocking

3
- Ingress Queuing Processing
- Bypassed in case no ingress queuing support

5
- Re-assembling packets from superframe
- Send packet to corresponding NP
- Release buffer and fabric credit

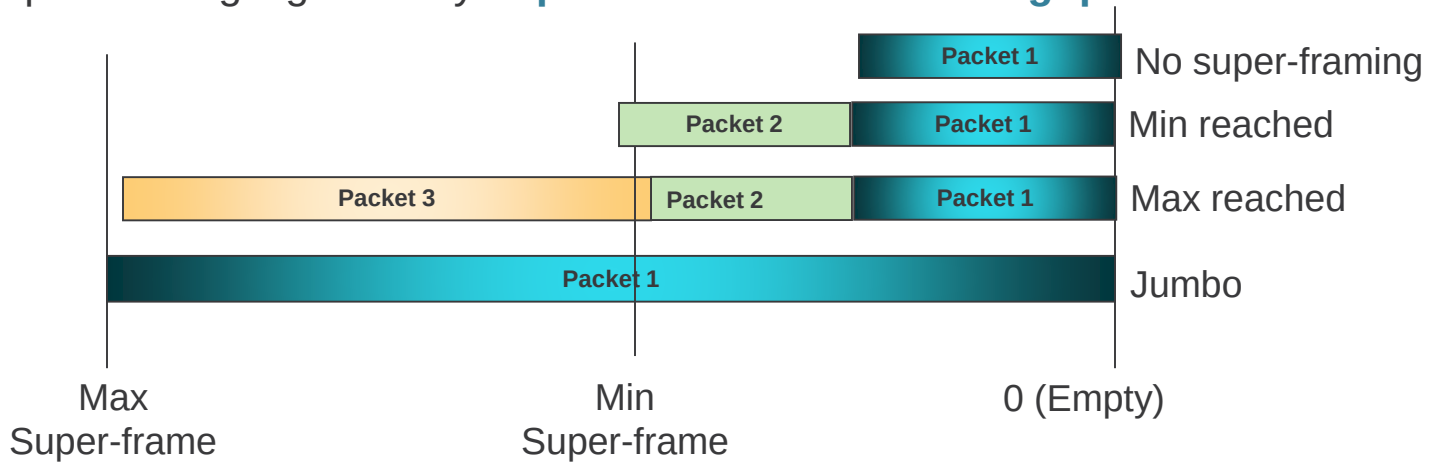
7
- Egress Queuing Processing

Switch Fabric Arbitration

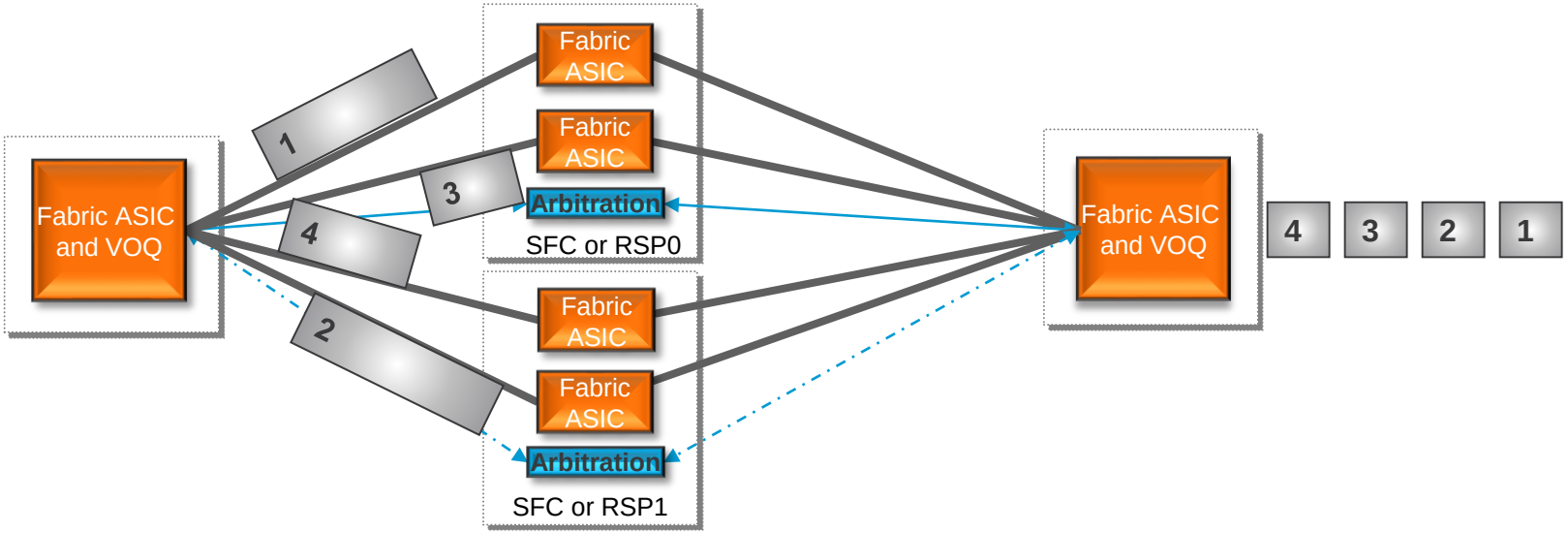


Fabric Super-framing Mechanism

- **Multiple unicast frames** from/to same destinations aggregated into **one super frame**
- Super frame is created if there are frames waiting in the queue, up to 32 frames or when min threshold met, can be aggregated into one super frame
- Super frame only apply to unicast, **not multicast**
- Super-framing significantly **improves total fabric throughput**

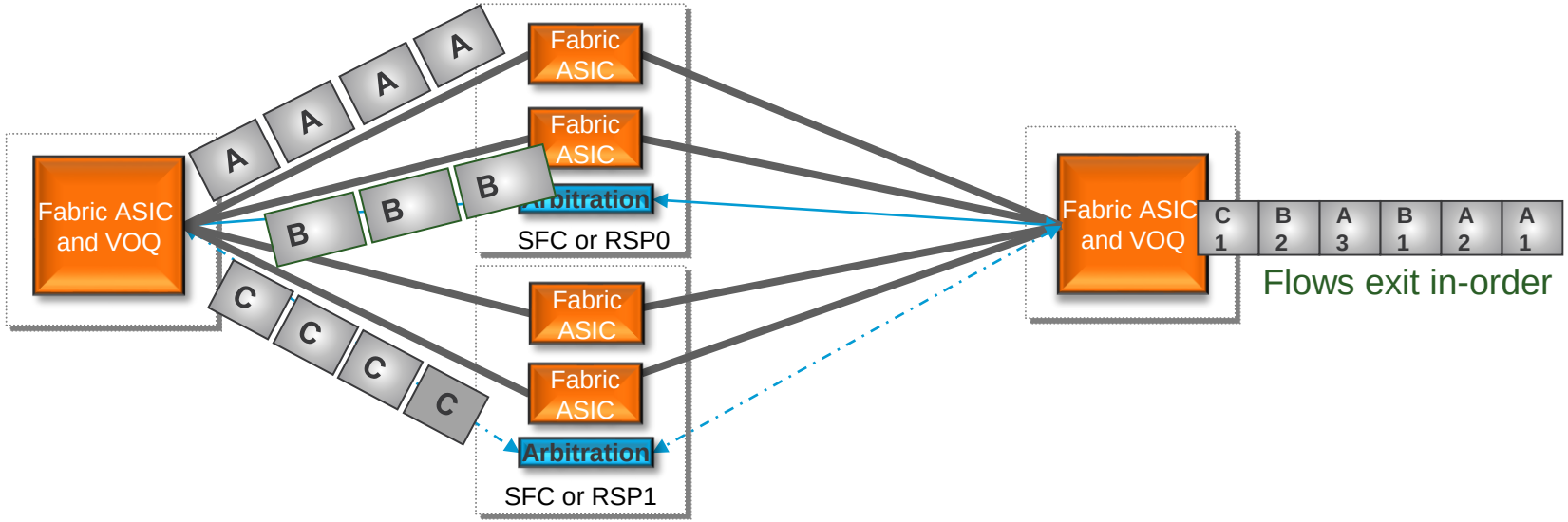


Fabric Load Balancing – Unicast



- Unicast traffic sent across first available fabric link to destination (maximizes efficiency)
- Each frame (or super frame) contains sequencing information
- All destination fabric ASIC have re-sequencing logic
- Additional re-sequencing latency is measured in nanoseconds

Fabric Load Balancing – Multicast

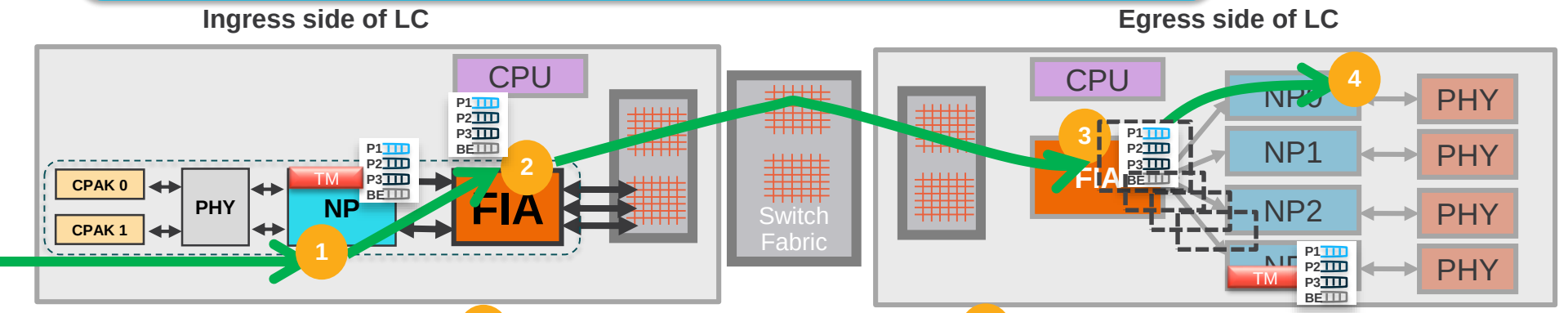


- Multicast traffic hashed based on (S,G) info to maintain flow integrity
  - Very large set of multicast destinations preclude re-sequencing
- Multicast traffic is non arbitrated – sent across a different fabric plane

ASR 9000 QoS Architecture

ASR9000 Priority-Based QoS Architecture

- Dedicated Traffic Manager(TM) for Traffic Queuing
- User Configurable QoS Policy on Ingress/Egress NP
- End-to-End priority propagation → Guarantee bandwidth, low latency for high priority traffic
- Unicast VOQ and back pressure



1
- Ingress (sub-)interface QoS Queues
- User Configurable Ingress QoS Policy

User-configuration with Ingress MQC

2
- 4xVOQ per VQI
- Up to 8K VOQs per TSK FIA (vs 4k per SKT FIA)

3
4x Egress Destination Qs per VQI, aggregated at egress port rate

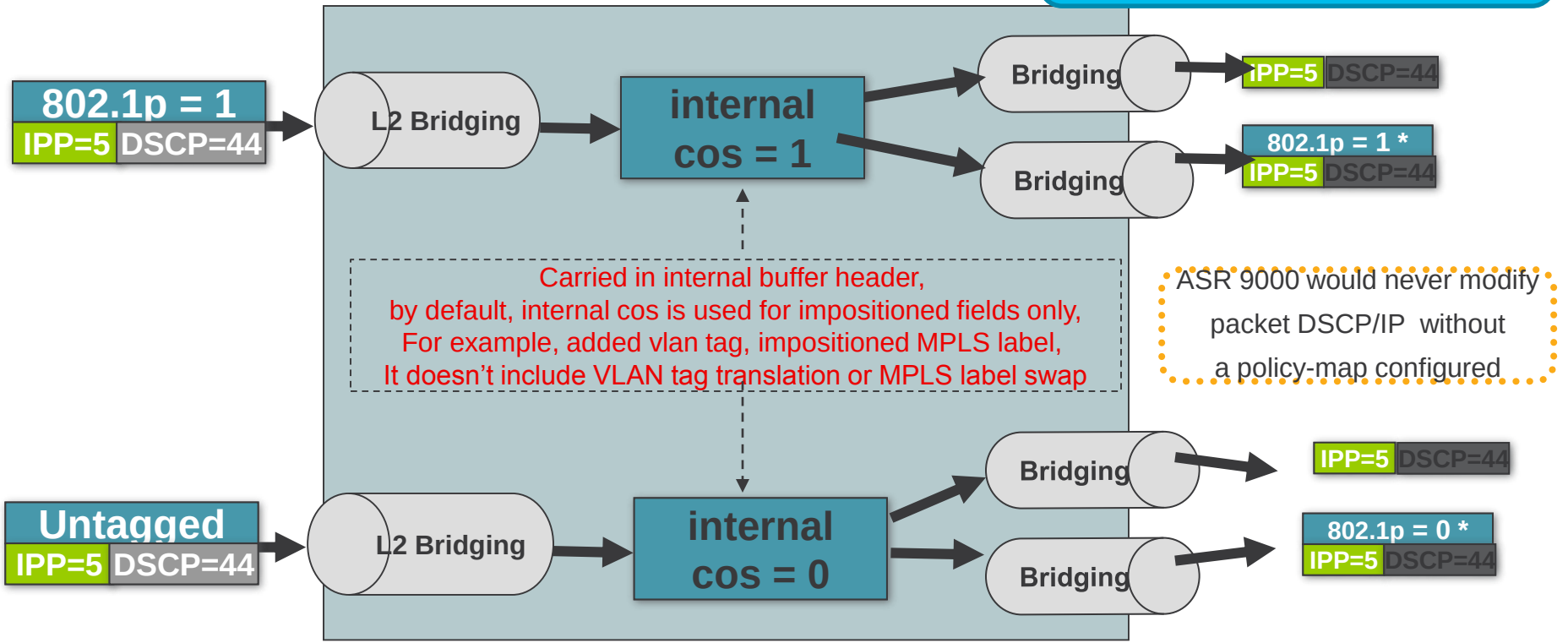
4
Egress (sub-)interface QoS Queues

User-configuration with Egress MQC

Implicit Configuration
Not User-controllable

Default Implicit Trust Model

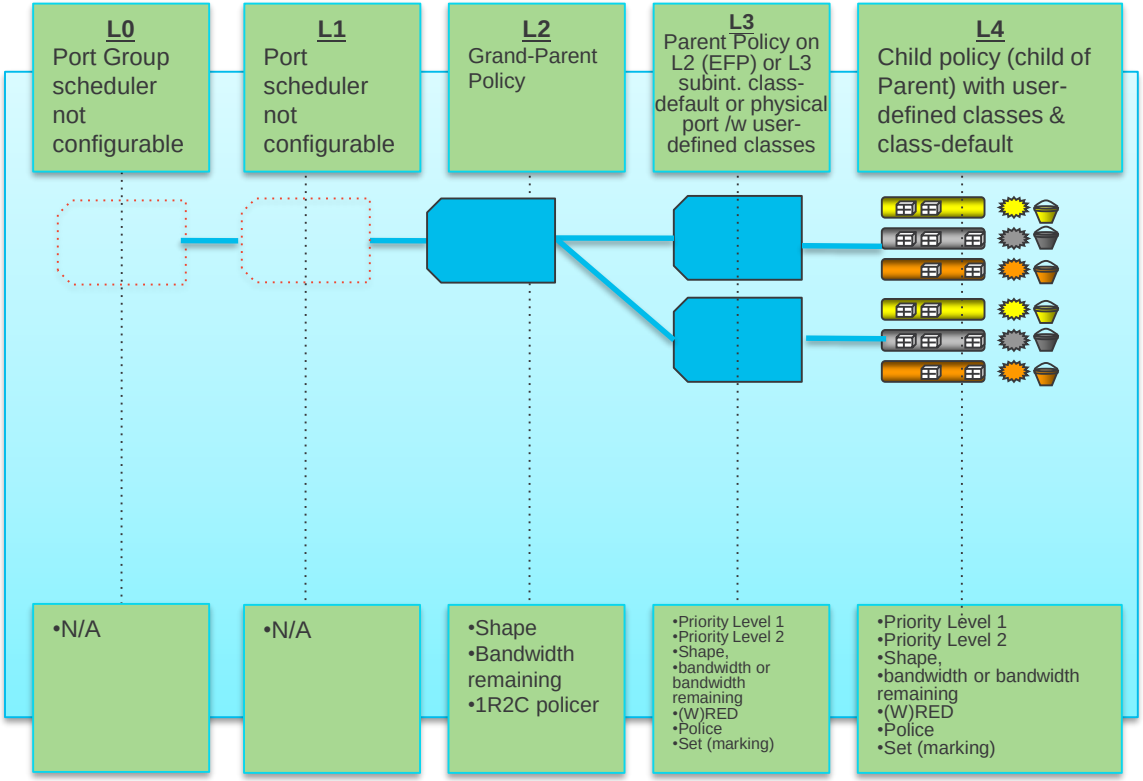
L2 IF: trust outer Cos
L3 IF: trust DSCP
L3 MPLS: trust outer EXP



Ingress line card

Egress line card

3-Layer Hierarchical QoS (H-QoS)

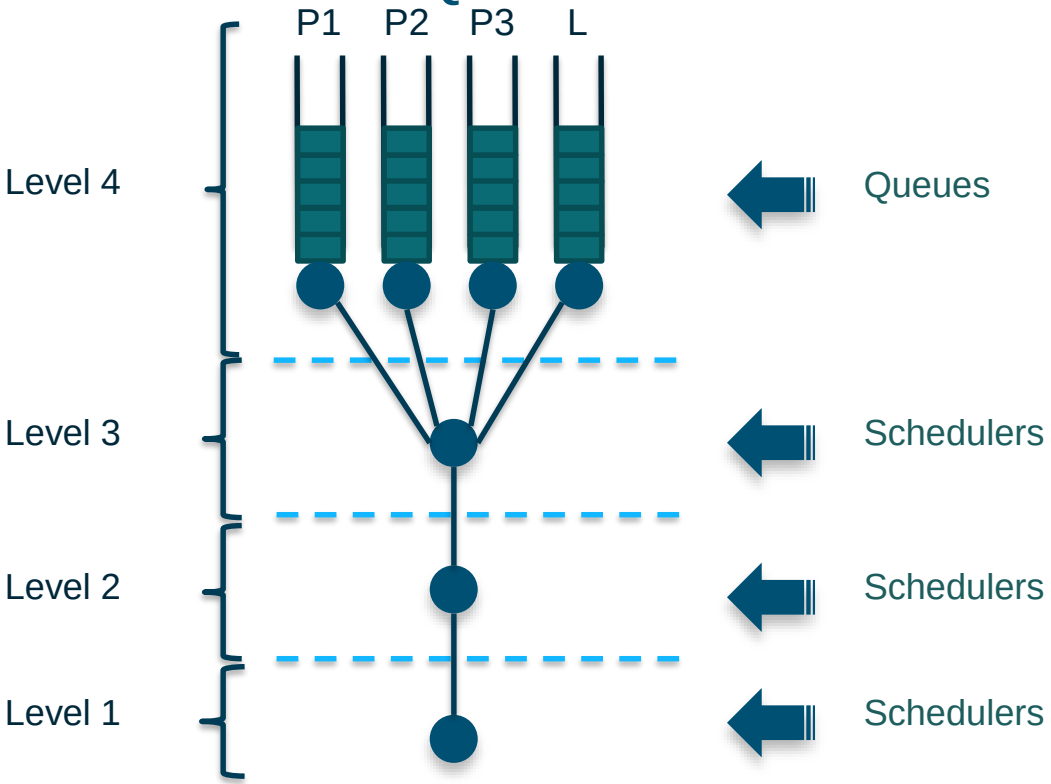



```
policy-map child
 class Pr1
  police rate 64 kbps
  priority level 1
 class Pr2
  police rate 10 mbps
  priority level 2
 class Cl3
  bandwidth 3 mbps
class Cl4
  bandwidth 1 mbps
!
policy-map parent
 class parent1
  shape average 100 mbps
  service-policy child
 class parent2
  shape average 25 mbps
  service-policy child
 class class-default
!
policy-map grand-parent
 class class-default
   shape average 500 mbps
   service-policy parent
```

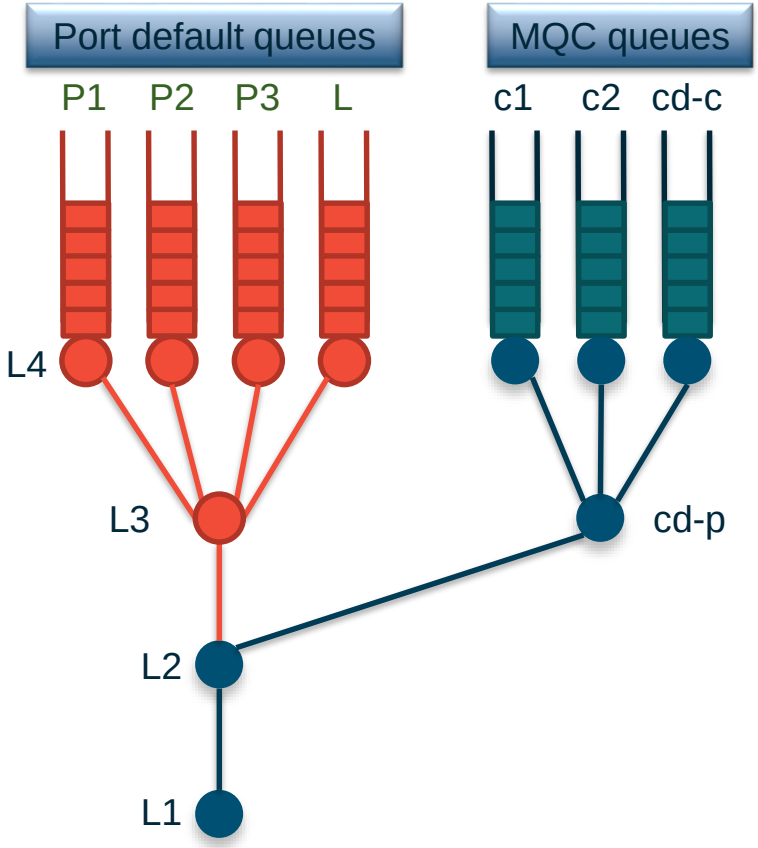
H-QoS – Supported Classification/Policy

Policy-map hierarchy level	Classification support	Policy Support
Grand-parent	Only class-default	• Shape Average • Bandwidth remaining • 1R2C policer with only drop/transimit action(no set/mark, Tomahawk card only)
Parent	User defined fields with restrictions based on format/interface types.	• Priority/WRED Queue and Queue-limit on Leaf only • Policer/Shaper/Marking/non-Priority Queue/Bandwidth/Bandwidth Remaining
Child	User defined fields with restrictions based on format/interface types.	• Priority/WRED Queue and Queue-limit on Leaf only • Policer/Shaper/Marking/non-Priority Queue/Bandwidth/Bandwidth Remaining

Default Interface Queues



MQC Hierarchy in Queuing ASIC




```
policy-map child
 class c1
  priority level 1
  police rate 640 kbps
 class c2
  bandwidth 20 mbps
 class class-default
  bandwidth 1 mbps
!
policy-map parent
 class class-default
  shape average 35 mbps
  service-policy child
!
interface GigabitEthernet0/0/0/0
 service-policy output parent
```

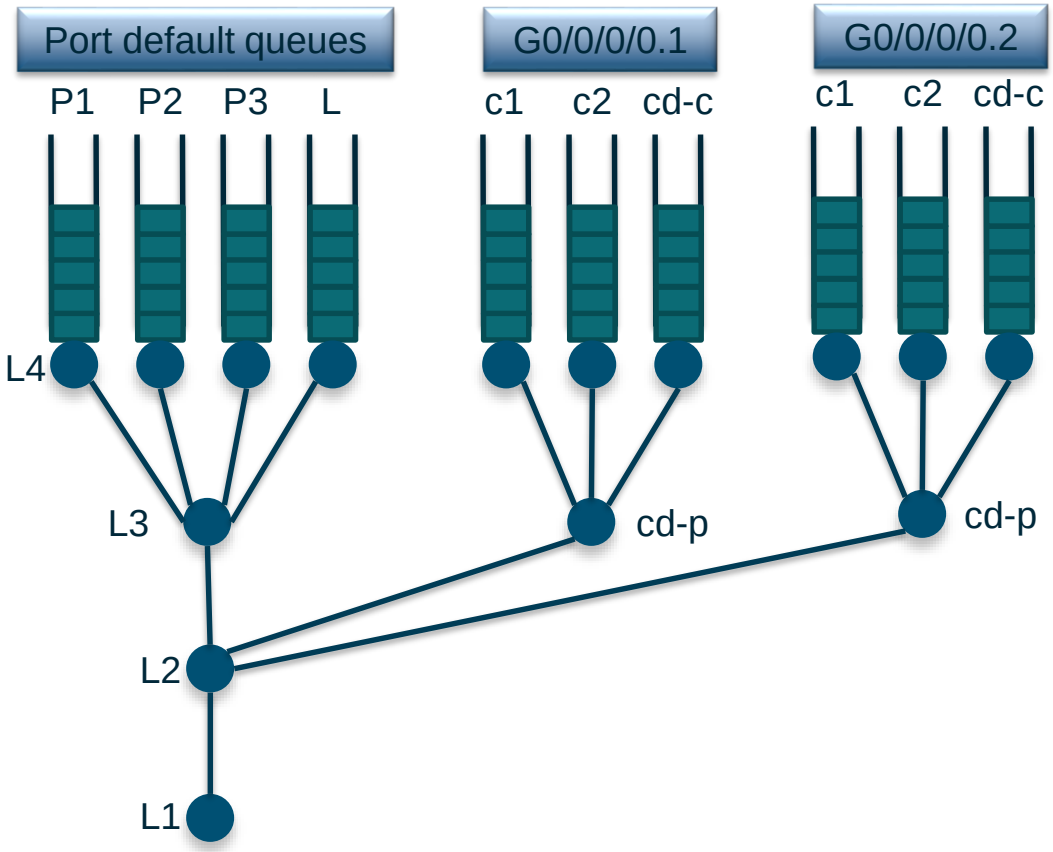
cd-c

cd-p

Inactive entity

Active entity

MQC Hierarchy in Queuing ASIC




```
policy-map child
 class c1
  priority level 1
  police rate 640 kbps
 class c2
  bandwidth 20 mbps
 class class-default
  bandwidth 1 mbps
!
policy-map parent
 class class-default
  shape average 35 mbps
  service-policy child
!
interface GigabitEthernet0/0/0/0.1
 service-policy output parent
!
interface GigabitEthernet0/0/0/0.2
 service-policy output parent
```

cd-c

cd-p

Inactive entity

Active entity

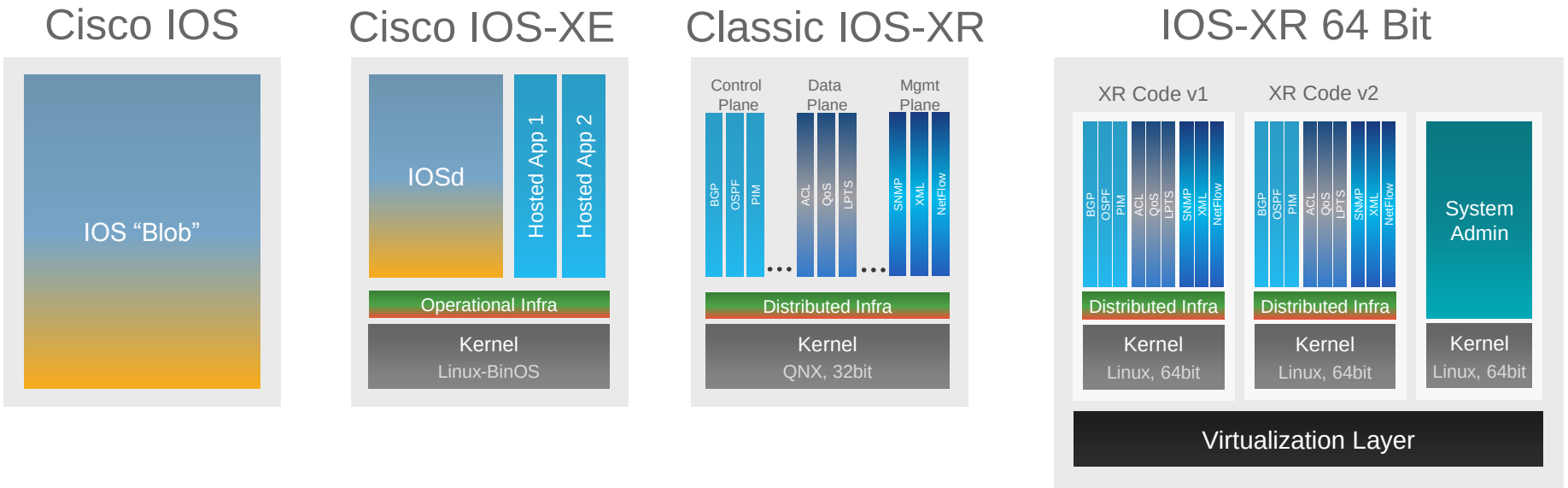
Pop Quiz ????

For an incoming untagged layer 2 frame, what is the priority value used when the frame is processed inside ASR9000?

- **2**
- **COS bit**
- **802.1p Value**
- **0**

IOS-XR & IOS-XR 64 Bit

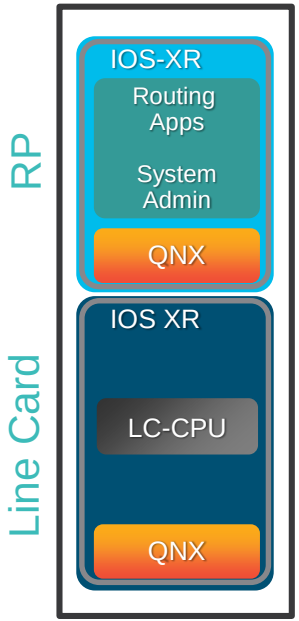
Cisco IOS – A Recap



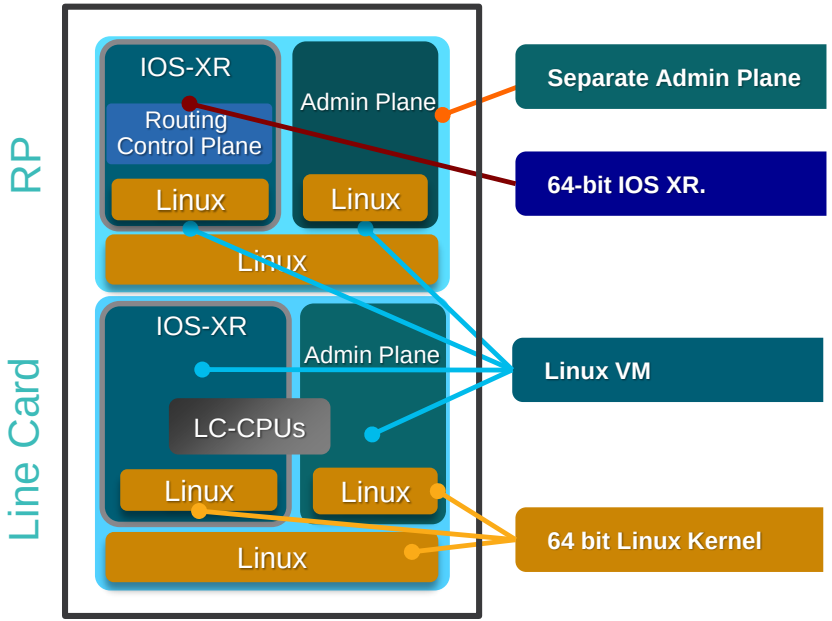
1990s → 2000s → 2003-04 → Present Day

Incremental Development, with Industry leading investment protection

IOS XR Evolution: XR 64 bit Architecture



Classic IOS XR



Linux-based Virtualized
IOS XR 64 Bit

IOS XR 64 Bit Packaging

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	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

IOS XR 64 Bit Packaging - GISO – Golden ISO

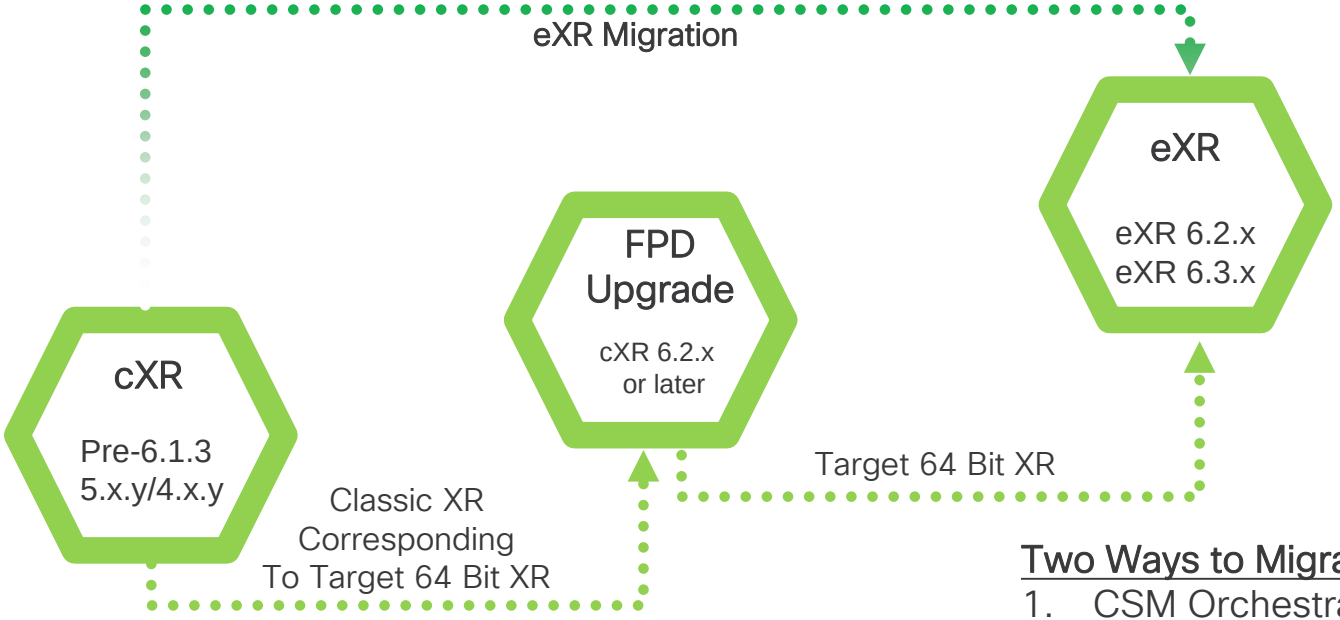
- GISO is a customized iso which is built as per individual customer needs.
- GISO contains Mini ISO + rpms + SMUs + config

How to build GISO: use tool provided on the router at /pkg/bin/gisobuild.py in XR domain.

Usage: gisobuild.py [-h] -i BUNDLE_ISO [-r RPMREPO] [-c XRCONFIG] [-l GISOLABEL] [–m] [-v]
Example: **gisobuild.py -i asr9k-mini-x.iso -r . -c config-file -l v1**

Script Parameter	Expansion	Explanation	Required/optional
-I	BUNDLE_ISO	Path to mini ISO	Required
-r	RPMREPO	Path to RPM repo	Optional
-c	XRCONFIG		Optional
-l	GISOLABEL	GISO Label	Optional
-m	migration	To build migration tar for ASR9K only.	Optional
-v	version	Print script version and exit	Optional
-h	help	Print help menu	N/A

Migrating Classic XR to IOS XR 64 Bit



Two Ways to Migrate:

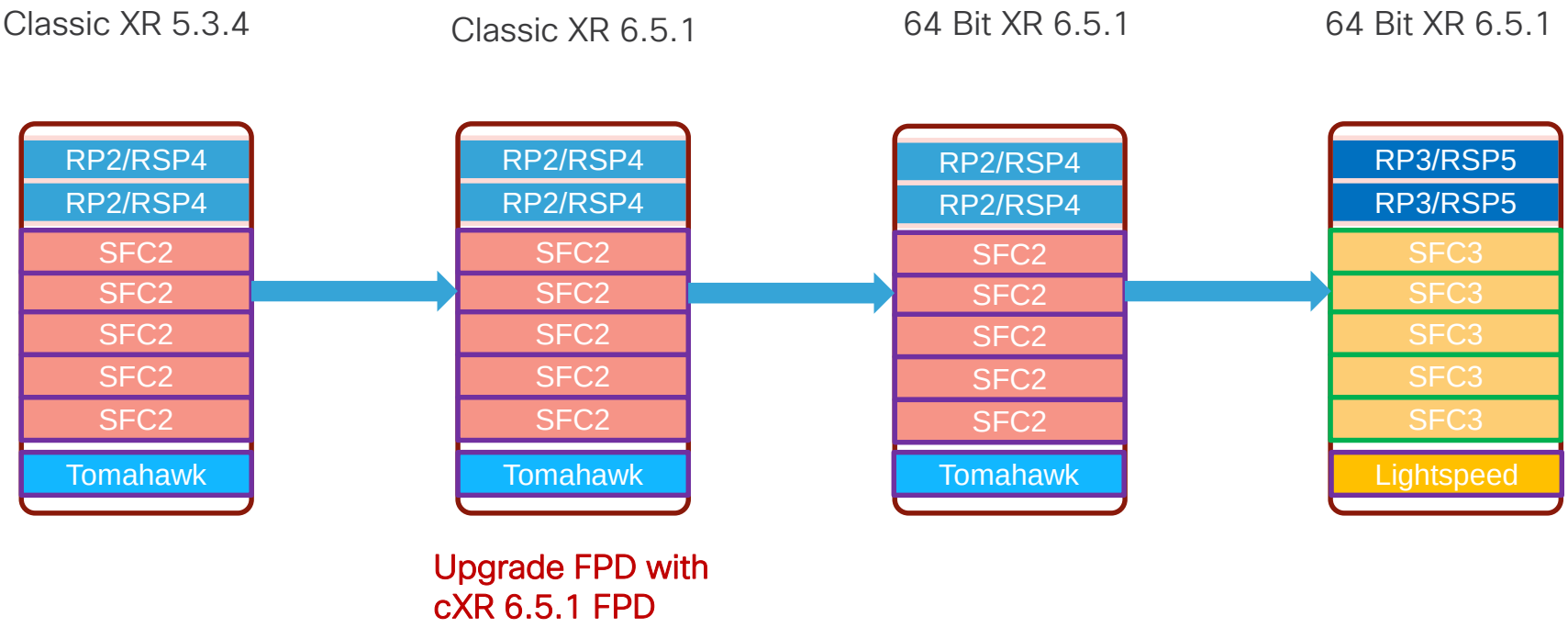
1. CSM Orchestrated Migration
2. MOP: Manual Migration

Follow below link for details:
https://www.cisco.com/c/en/us/td/docs/routers/asr9000/migration/guide/b-migration-to-ios-xr-64-bit/b-migration-to-ios-xr-64-bit_chapter_011.html

Migration Pre-requisites

HW Component Check	Upgrade cXR	Operational Status	Backup cfg to External Server (Calvados, XR)
• All hardware components, Chassis, RSP/RP, LC, FC, FAN and PEM, should be supported in ASR9K 64 bit. Any unsupported hardware may fail to boot after migration	• Any pre-6.1.3 release needs to be upgraded to classic XR corresponding to target IOS XR 64 Bit version	• All hardware components must be in operational state before migration	• Back up admin/XR configuration to external server

ASR 9000 System Migration Example to Lightspeed



Questions?

Conclusion

- **ASR9000 - Truly Carrier-Class Edge Router Provides:**
  - **Rich Features, Flexible Service Capability**
  - **Variety of Hardware to Meet Different Capacity Requirements**
- **Fully Distributed Architecture for High Performance and System Scalability**
- **Uniform, Open and Modularized Software Architecture**

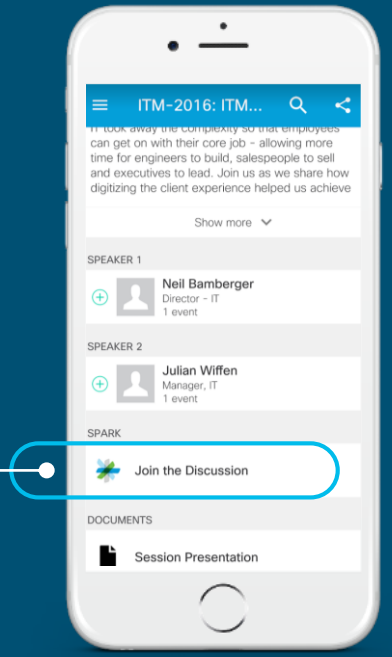
Cisco Spark

Questions?

Use Cisco Spark to communicate with the speaker after the session

How

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

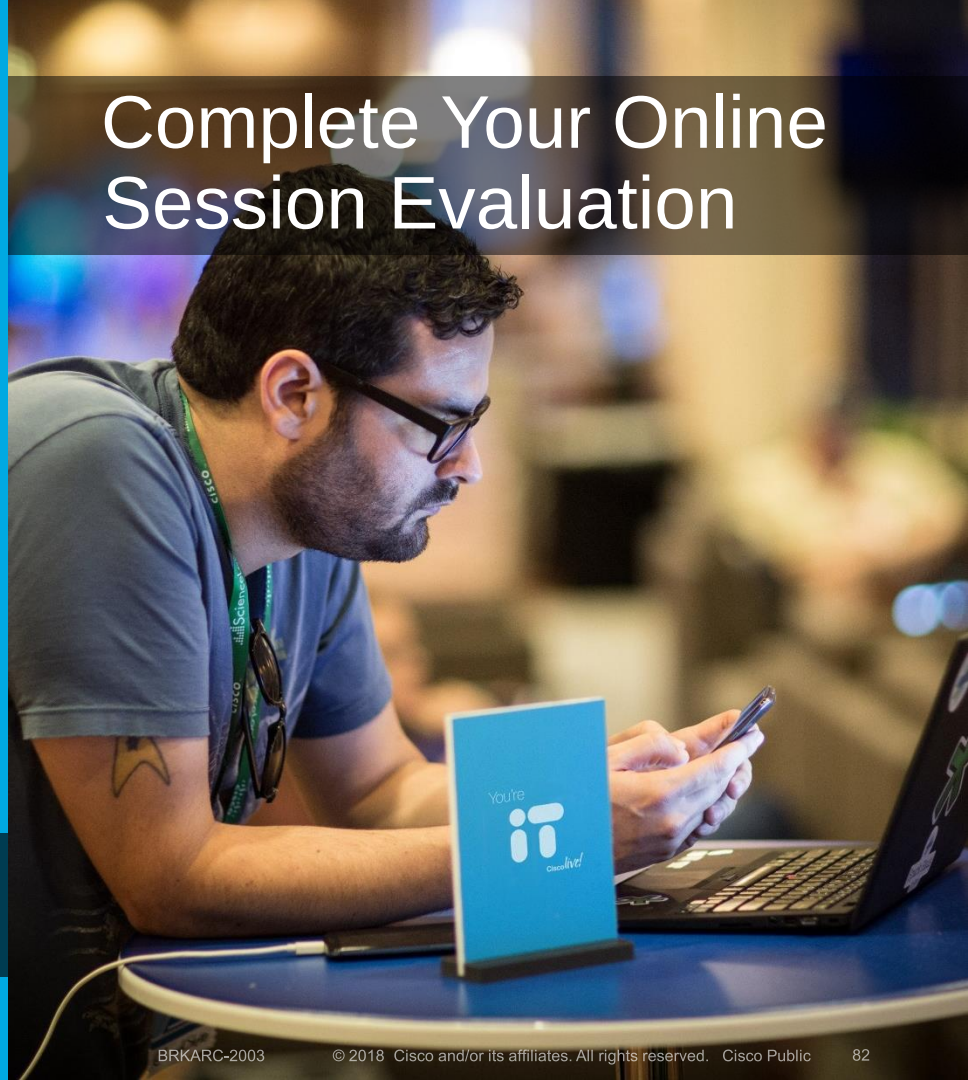


cs.co/ciscolivebot#BRKARC-2003

Complete Your Online Session Evaluation

- Please complete your Online Session Evaluations after each session
- Complete 4 Session Evaluations & the Overall Conference Evaluation (available from Thursday) to receive your Cisco Live T-shirt
- All surveys can be completed via the Cisco Live Mobile App or the Communication Stations

Don’t forget: Cisco Live sessions will be available for viewing on-demand after the event at CiscoLive.com/Online.



Continue Your Education

- Demos in the Cisco campus
- Walk-in Self-Paced Labs
- Tech Circle
- Meet the Engineer 1:1 meetings
- Related sessions

Thank you


You're
Cisco live!