Cisco ASR9000 Architecture and Network Virtualisation

BRKARC-2003

Session Goals

- **Understand the Cisco ASR 9000 system architecture**

  Key components RSP, switch fabric, Line card

  Packet flows, control plane, data plane

  nV (network virtualisation)

  IOS XR software high level overview

- **Not a dedicated IOS-XR troubleshooting session**

  Please refer to BRKSPG-3612 (Troubleshooting Cisco IOS XR) for IOS-XR related technical information

- **Not a dedicated platform configuration / troubleshooting session**

Agenda

- **Hardware Overview**

- **Carrier Class, Scalable System Architecture**

  – Fabric architecture

  – Line card architecture

  – Packet flow, control plane and data plane

  – Multicast packet flow and replication

  – Internal system Qos

- **nV Technology Overview**

 Cisco Public

ASR 9000 System Introduction and Chassis Overview

Cisco ASR 9000 – Next-Gen Edge Routing Platform
High level system design goals

- **System longevity : 10-15 years in the network**
- **Full HW portfolio / form factors**
- **Software modularity, portability, flexibility**
- **High power control plane / data plane scale**
- **10/100/??? Gigabit ethernet aggregation**
- **Complex feature paths : Video, Mobile, BNG**
- **Data centre internet and cloud gateway**

Cisco ASR 9000 – Next-Gen Edge Routing Platform
Achieving design goals

- **Significant hardware and software commonality**

- **Built on a repeatable/extensible "slice" arch.**

- **Edge-enhanced IOS XR software**

- **Shift to x86 CPU for improved power and virtualisation capabilites**

- **Highly integrated network processors for high data plane scale and feature flexibility**

- **Complete set of L2/L3 protocols and forwarding microcode**
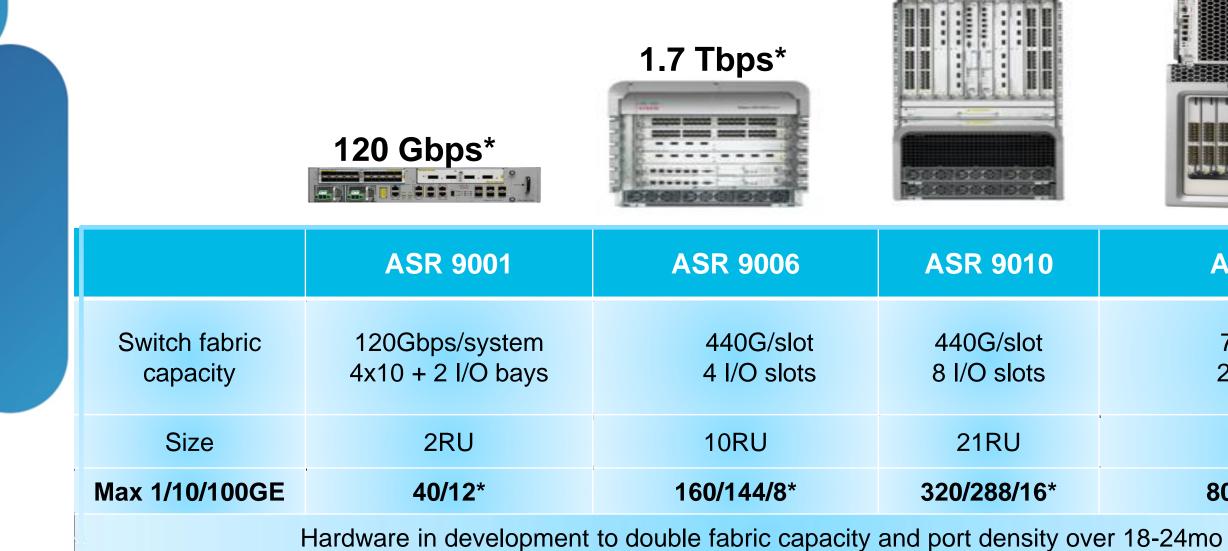
System Capacity : key components

- **Modular systems have multiple components with independent capabilities and interactions**

- **Important to understand how these fit together and relative costs**

- **Backplane**

- **System interconnect (generally a switch fabric)**

- **Linecards (or choose whatever name you like)**

- **ASR 9000 series has 5 current form factors, each with multiple current and future interconnect/fabric options, and multiple linecard options**

ASR 9000 Chassis Overview

Identical hardware, software, and forwarding
Across multiple physical form factors
Flexibility in data- to control-plane relationship

15Tbps*

3.5 Tbps*

1.7 Tbps*

120 Gbps*

	ASR 9001	ASR 9006	ASR 9010	ASR 9922
Switch fabric capacity	120Gbps/system 4x10 + 2 I/O bays	440G/slot 4 I/O slots	440G/slot 8 I/O slots	770G/slot 20 I/O slot
Size	2RU	10RU	21RU	44RU
Max 1/10/100GE	**40/12***	**160/144/8***	**320/288/16***	**800/720/40***
Hardware in development to double fabric capacity and port density over 18-24mo				

actual, shipping, year 2012, engineering-math, not doubled, usable densities

Cisco Public

Cisco live!
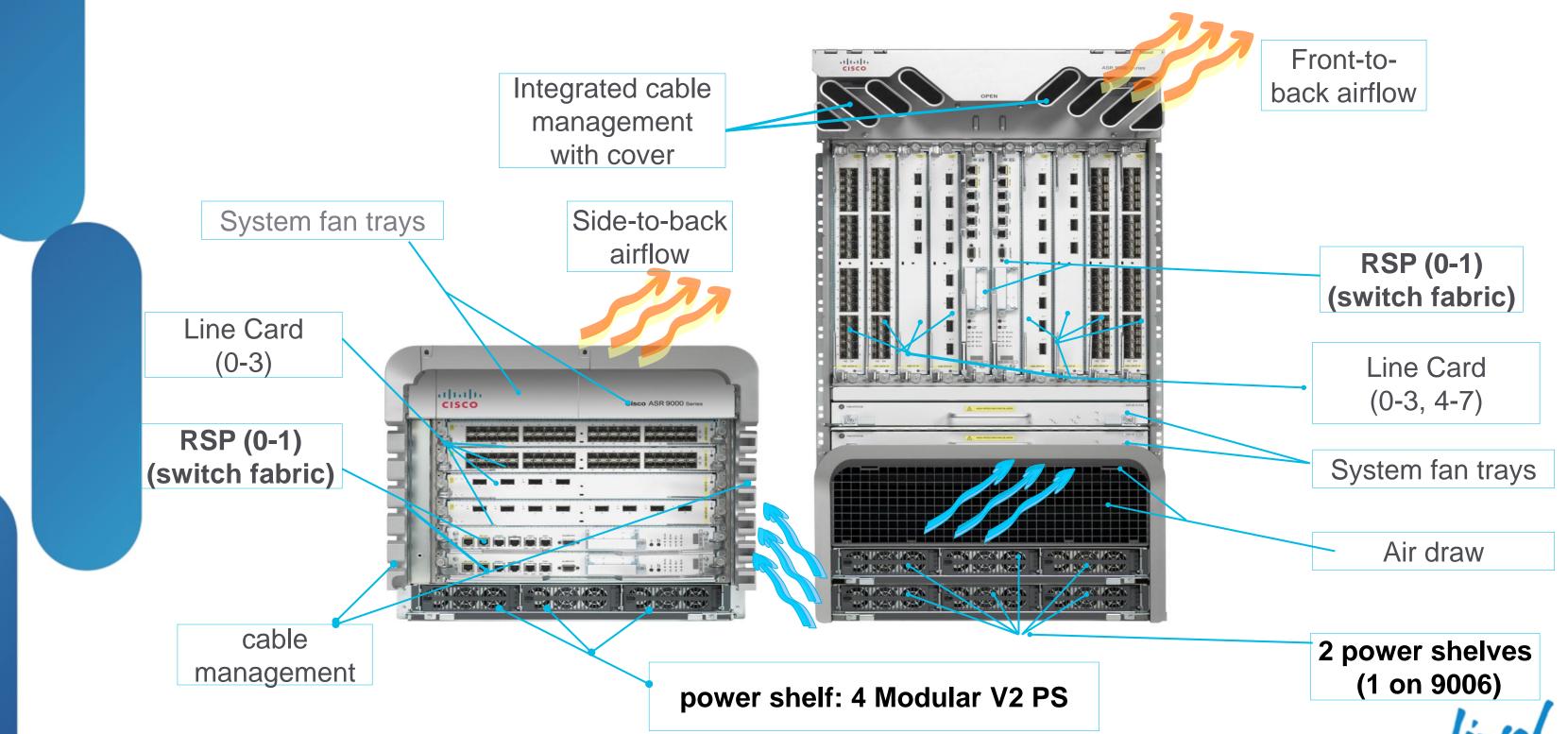
ASR 9010 and ASR 9006 Chassis

Integrated cable management with cover

Front-to-back airflow

System fan trays

Side-to-back airflow

RSP (0-1) (switch fabric)

Line Card (0-3)

Line Card (0-3, 4-7)

RSP (0-1) (switch fabric)

System fan trays

cable management

Air draw

power shelf: 4 Modular V2 PS

2 power shelves (1 on 9006)
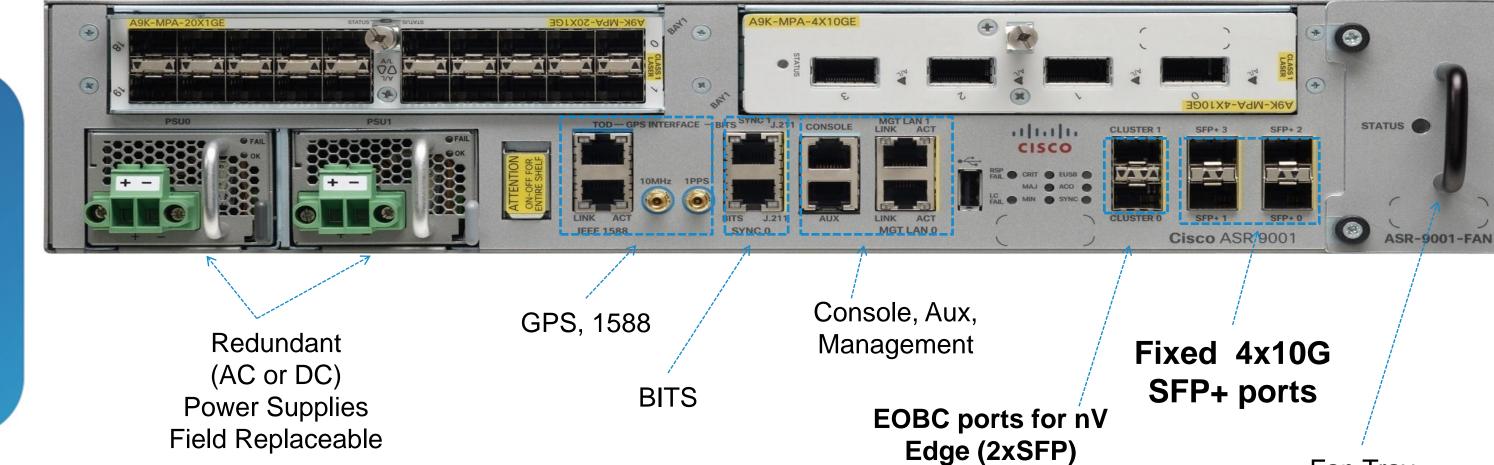
ASR 9001 "Iron Man" Chassis

Sub-slot 0 with 20xGEMPA   **Sub-slot 1 with 4x10GE MPA**

GPS, 1588

Console, Aux,
Management

Redundant
(AC or DC)
Power Supplies
Field Replaceable

BITS

Fixed 4x10G
SFP+ ports

**EOBC ports for nV
Edge (2xSFP)**

Fan Tray
Field Replaceable

Supported MPA (Media Port Adapter)**: 20xGE, 2/4x10GE, 1x40GE**

ASR 9922 "Megatron" Chassis bandwidth system

8+8 ACs or
11+1 DCs

Front-to-back airflow

20x Line Cards
(10 top, 10 bottom)

Height : 44 RU (AC & DC)
Depth : 28.65" (800mm)
Width : 17.75" (fits 19" rack)

Fan trays 4x
(ASR-9922-FAN)

2 x RPs

4+1 Fabric Cards initially
6+1 Fabric Cards in future

Cisco Public

ASR 9000 System Components
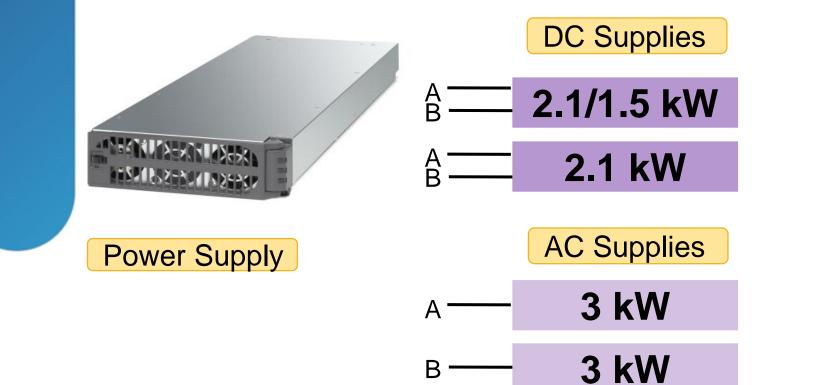
Power and Cooling

ASR-9010-FAN

ASR-9006-FAN

- Fans unique to chassis
- Variable speed for ambient temperature variation
- Redundant fan-tray
- Low noise, NEBS and OSHA compliant

Power Supply

DC Supplies

A
B
2.1/1.5 kW

A
B
2.1 kW

AC Supplies

A
3 kW

B
3 kW
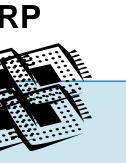
- Single power zone
- All power supplies run in active mode
- Power draw shared evenly
- 50 Amp DC Input or 16 Amp AC for Easy CO Install

ASR 9006/9010/9922 RSP (Route/Switch Processors)

	RSP	RSP440	9922-RP
Processors	PPC/Freescale 2 Core 1.5GHz	Intel x86 4 Core 2.27 GHz	Intel x86 4 Core 2.27 GHz
RAM	RSP-4G: 4GB RSP-8G: 8GB	**RSP440-TR: 6GB** **RSP440-SE: 12GB**	**-TR: 6GB** **-SE: 12GB**
nV EOBC ports	No	Yes, 2 x 1G/10G SFP+	Yes, 2 x 1G/10G SFP+
Switch fabric bandwidth	92G + 92G (with dual RSP)	220+220G (with dual RSP)	660+110 (7-fabric model)
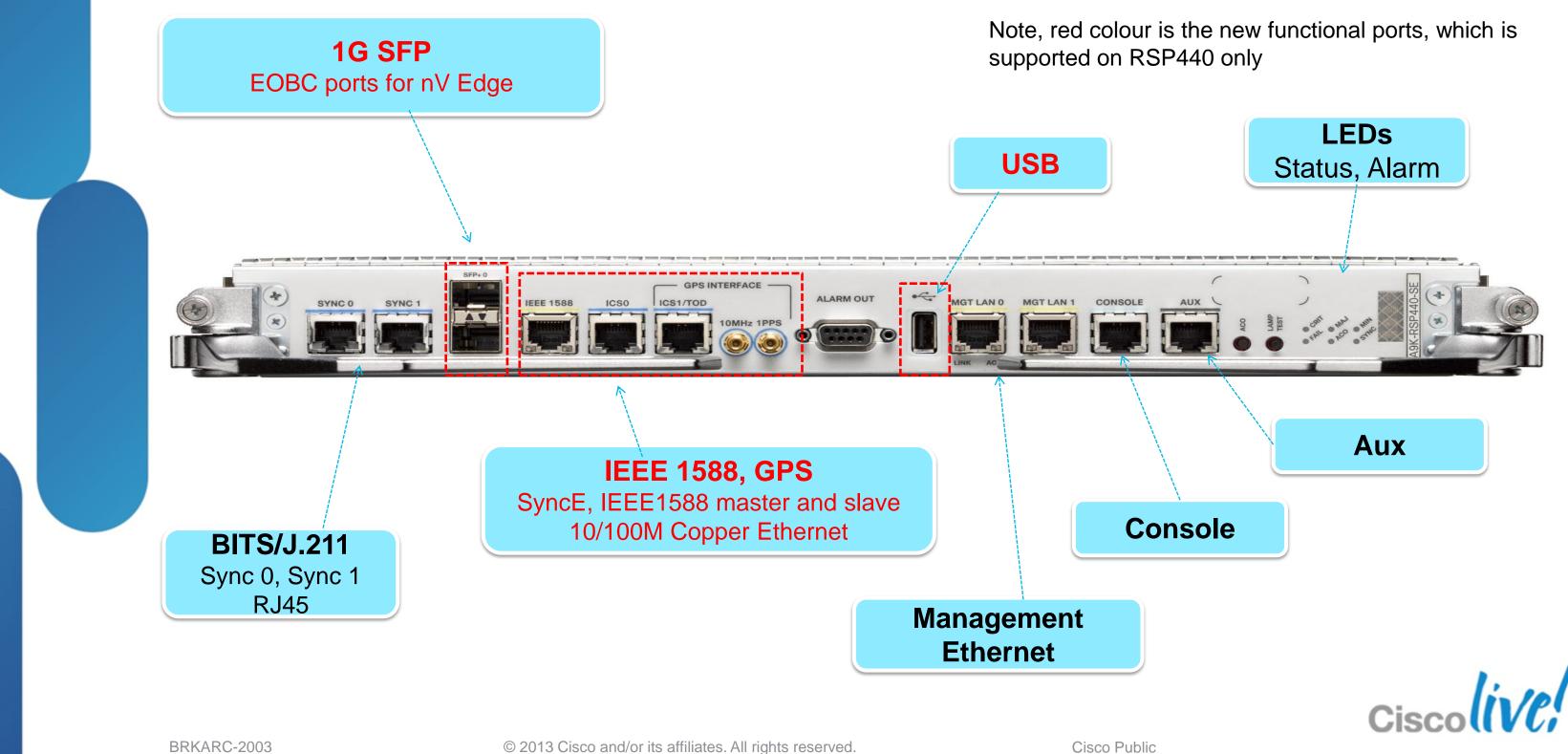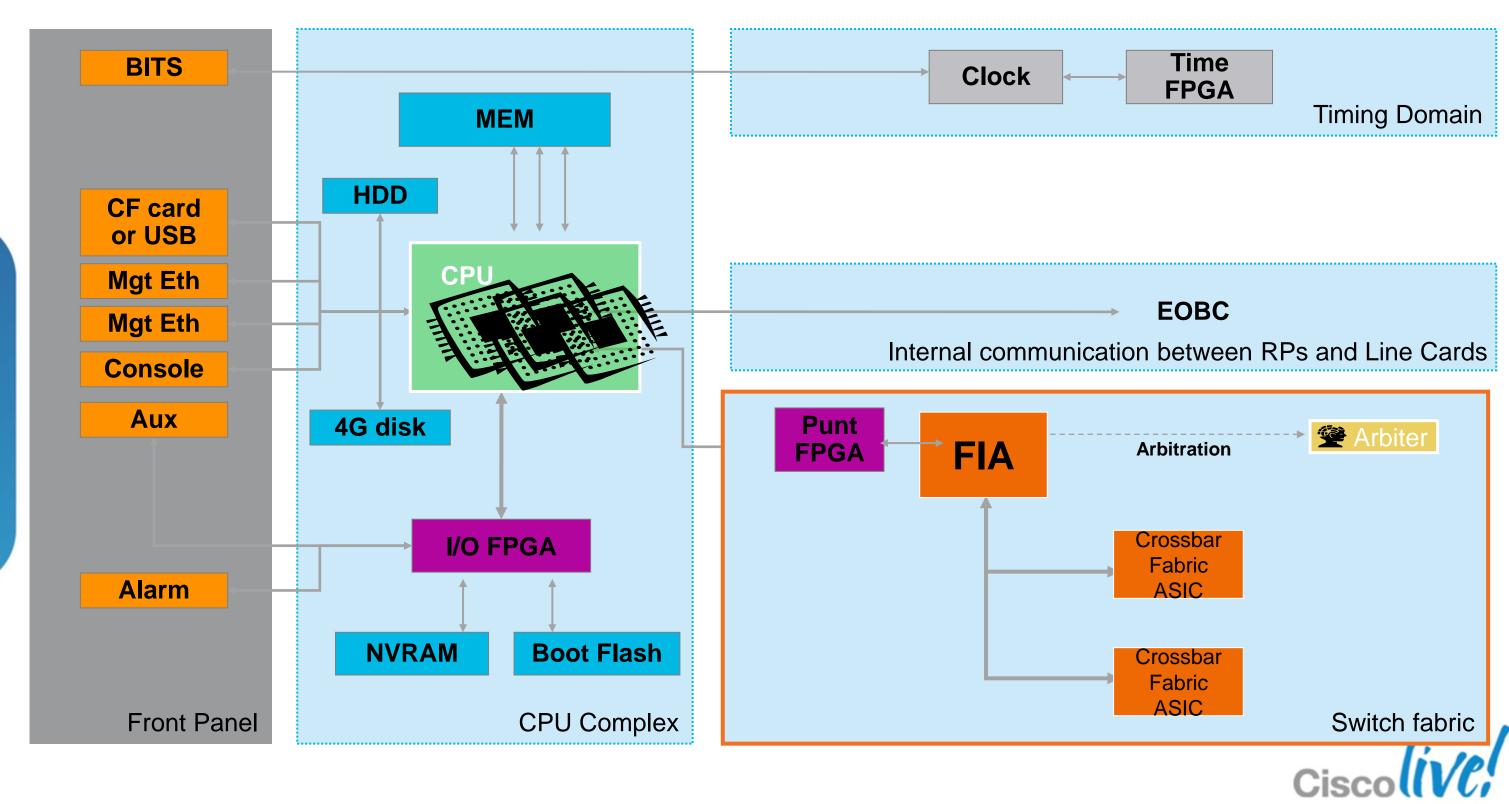
RSP440 – Front Ports

1G SFP
EOBC ports for nV Edge

Note, red colour is the new functional ports, which is supported on RSP440 only

USB

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

Cisco Public
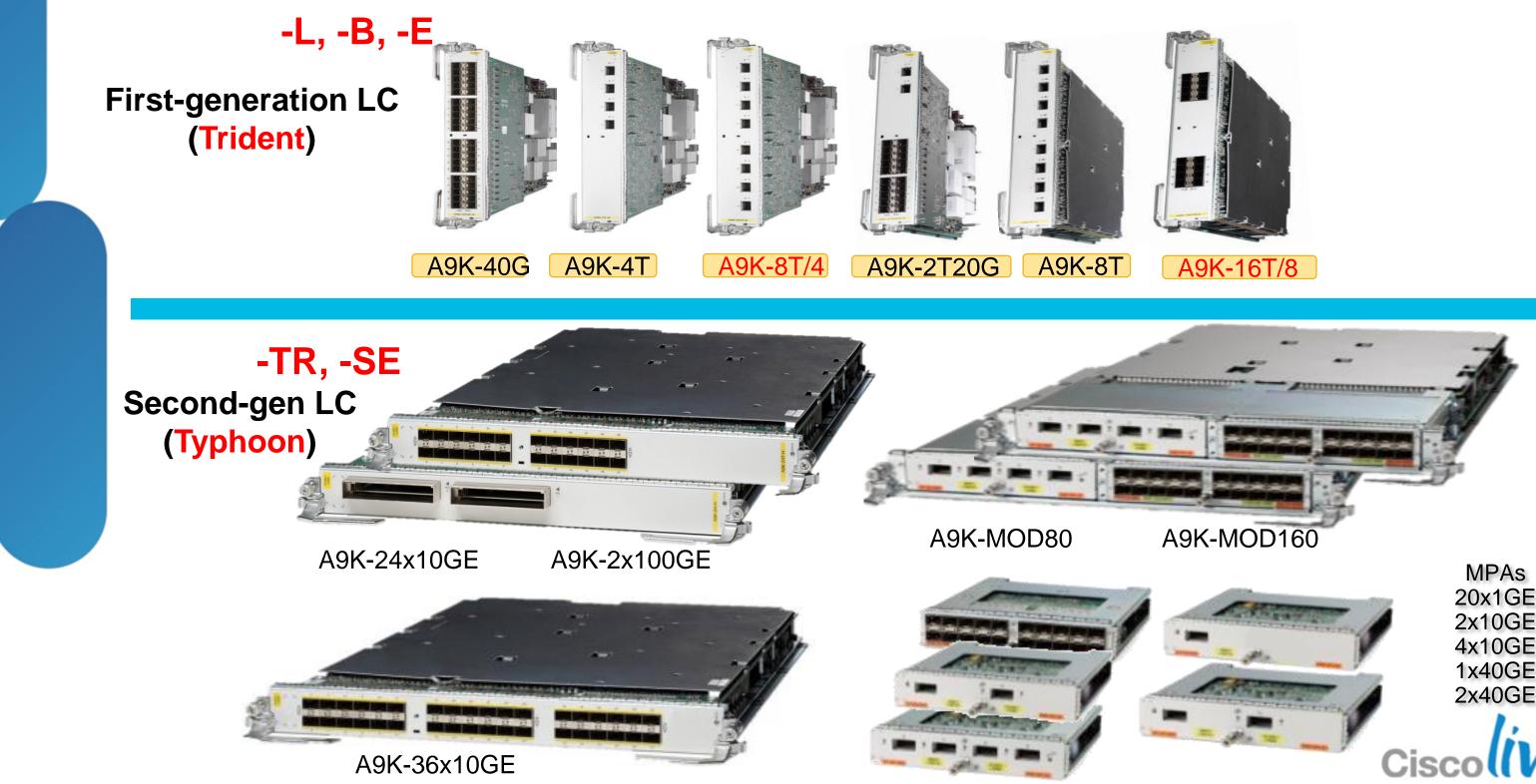
RSP Engine Architecture

Front Panel boxes:
- BITS
- CF card or USB
- Mgt Eth
- Mgt Eth
- Console
- Aux
- Alarm

CPU Complex:
- MEM
- HDD
- CPU
- 4G disk
- I/O FPGA
- NVRAM
- Boot Flash

Timing Domain:
- Clock
- Time FPGA

EOBC
Internal communication between RPs and Line Cards

Switch fabric:
- Punt FPGA
- FIA
- Arbiter
- Arbitration
- Crossbar Fabric ASIC
- Crossbar Fabric ASIC

Cisco Public

ASR 9000 Ethernet Line Card Overview

-L, -B, -E

First-generation LC (Trident)

| A9K-40G | A9K-4T | A9K-8T/4 | A9K-2T20G | A9K-8T | A9K-16T/8 |

-TR, -SE

Second-gen LC (Typhoon)

A9K-24x10GE   A9K-2x100GE
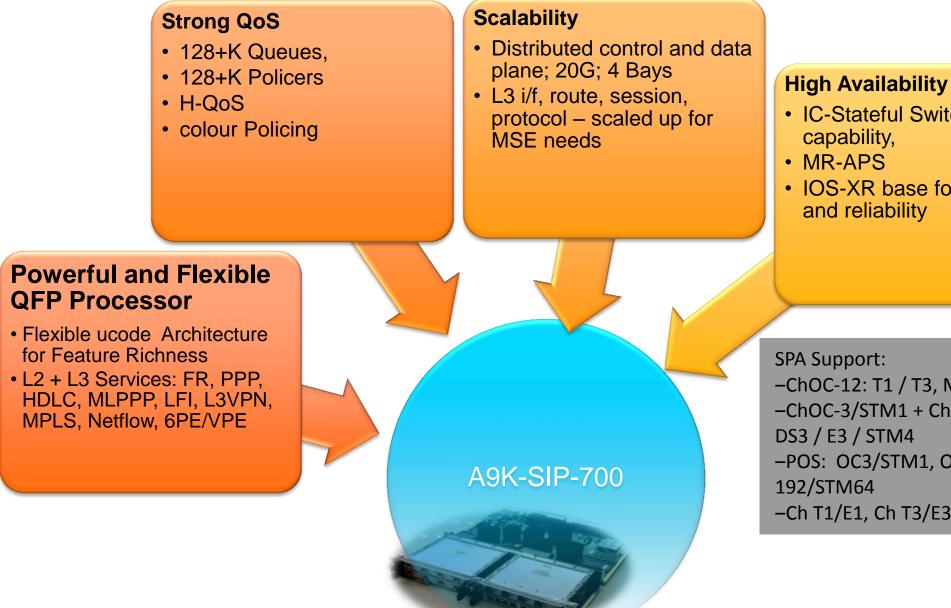
A9K-MOD80   A9K-MOD160

A9K-36x10GE

MPAs
20x1GE
2x10GE
4x10GE
1x40GE
2x40GE

Cisco Public

ASR 9000 SIP-700 and SPA
20Gbps, feature rich, high scale, low speed interfaces

SPA support list:

http://www.cisco.com/en/US/partner/prod/collateral/routers/ps9853/data_sheet_c78-573452.html

Strong QoS
- 128+K Queues,
- 128+K Policers
- H-QoS
- colour Policing

Scalability
- Distributed control and data plane; 20G; 4 Bays
- L3 i/f, route, session, protocol – scaled up for MSE needs

High Availability
- IC-Stateful Switch Over capability,
- MR-APS
- IOS-XR base for high scale and reliability

Powerful and Flexible QFP Processor
- Flexible ucode Architecture for Feature Richness
- L2 + L3 Services: FR, PPP, HDLC, MLPPP, LFI, L3VPN, MPLS, Netflow, 6PE/VPE

A9K-SIP-700

SPA Support:
- ChOC-12: T1 / T3, MLPPP, LFI, IC-SSO, MR-APS
- ChOC-3/STM1 + ChOC-12: DS0 / STM1 + ChOC-48: DS3 / E3 / STM4
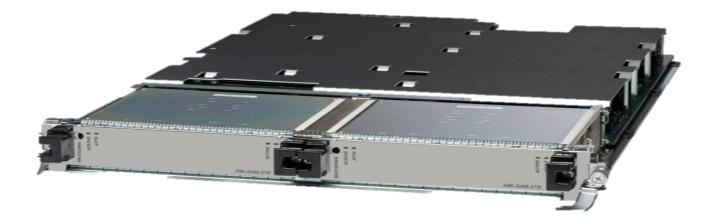- POS: OC3/STM1, OC-12/STM4, OC-48/STM16, OC-192/STM64
- Ch T1/E1, Ch T3/E3, CEoPs, ATM

ASR 9000 ISM (Integrated Service Module)

CDS Streaming:
TV and internet streaming
Error repair

CGN (carrier grade NAT):
NAT44, DS-Lite
NAT64

Feature	ASR 9000 ISM Capabilities
Applications	Ultra-Dense VoD, TV, Internet Streaming, Error Repair, CGv6
Bandwidth	30-40 Gbps streaming capacity ~3 Gbps cache fill rate
Compatibility	Works with all CDS appliances
Concurrent Streams	Up to 8,000 SD equivalent
Content Cache	3.2 TBytes at FCS - Modular Design
Video Formats	MPEG2 & AVC/H.264
Transport	MPEG over UDP / RTP
Session Protocols	RTSP / SDP
Environmental	NEBS / ETSI compliant

CDS: Manage 8,000 streams up to 40G per second
CGv6: 20M translations, 1M translations/sec., ~15Gbps throughput / ISM

Optical Networking Overview

Key concepts for high end optical routers/switches

- Optics are now a very significant part of system design.

- Mechanical design issues (space, power, heat) are becoming or in some cases more significant as packet forwarding

- The market for modular optics is exceptionally competitive and wildly variable in cost, quality, specifications

- This generates complexity and pain (we're sorry!) but ...

- lower costs and improves densities (you're welcome!)

ASR 9000 Optical Interface Support

- Modular optics used everywhere
- GE: SFP interfaces (T, SX, LX, ZX)
- 10GE: SFP+ and XFP modules
  - XFP (for line card with less than 8x10GE ports) : SR, LR, ZR, DWDM, CWDM, Tunable
  - SFP+ (for line card with more than 8X10GE ports): SR, LR, ZR

XFP

SFP, SFP+

Cisco Public

ASR 9000 40/100GE Optics

- ASR9k 40GE: QSFP form factor
  - SR4
  - LR4
- ASR9k 100GE: CFP form factor
  - LR4
  - SR10
- DWDM support:
  - 40GE has g.709/FEC built-in at FCS
    needs Q-SFP DWDM optics (future)
  - 100GE is "gray" only and no FEC/g.709
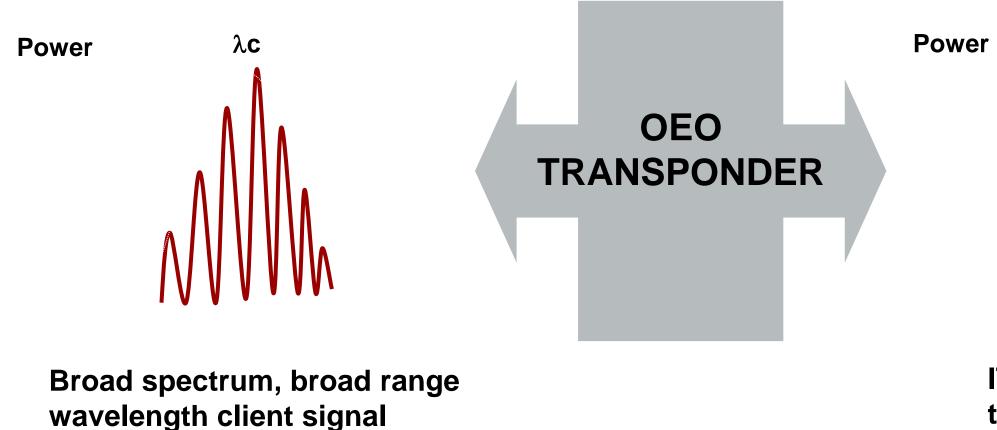   use external transponder

QSFP (40GE)

CFP

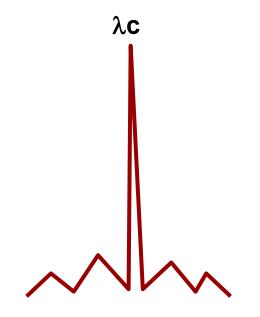**(100GE)**

IPoDWDM overview

- "IP over DWDM" is slightly mis-named....
- implies two things which are really orthogonal
  - Long-reach, high-precision, coloured optics
  - Enhanced error-correction algorithms to extend signal integrity ("transponder")
- (normally) requires some type of DWDM mux device (almost always dedicated HW)
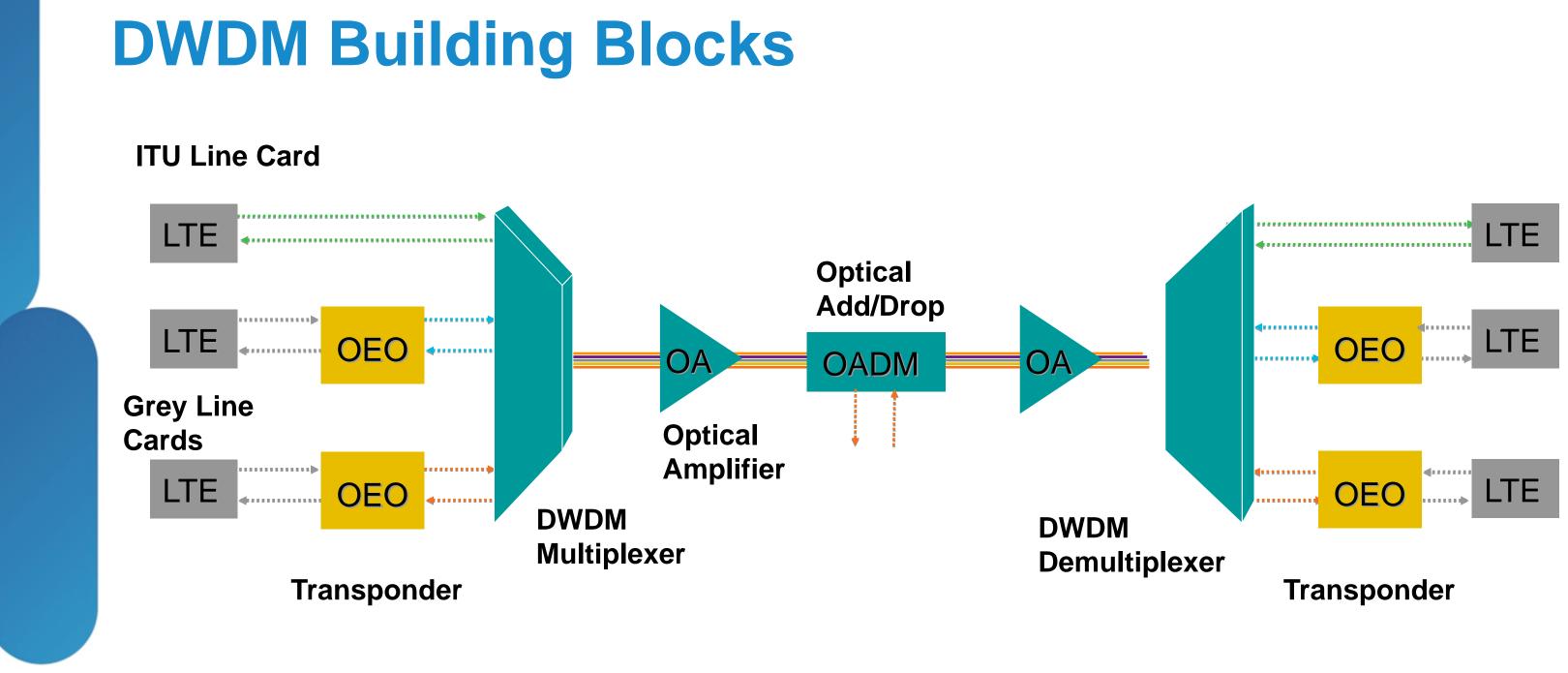
Transponder OEO Converter

- Converts the wide spectrum client laser input into a DWDM ITU compliant channel

- Converts the DWDM channel into a client compliant output signal

- Down side is that they add significant cost

Power λc

OEO TRANSPONDER

Power λc

Broad spectrum, broad range wavelength client signal

ITU Wavelength specific DWDM trunk channel

Cisco Public

DWDM Building Blocks

ITU Line Card

LTE

LTE

Grey Line Cards

LTE

OEO

OEO

Optical Add/Drop

OA

OADM

OA

DWDM Multiplexer

Optical Amplifier

DWDM Demultiplexer

Transponder

LTE

LTE

OEO

LTE

OEO

LTE

Transponder

LTE: line terminating equipment
OEO: optical-electrical-optical converter

ASR 9000 System Interconnect and Switch Fabric Architecture

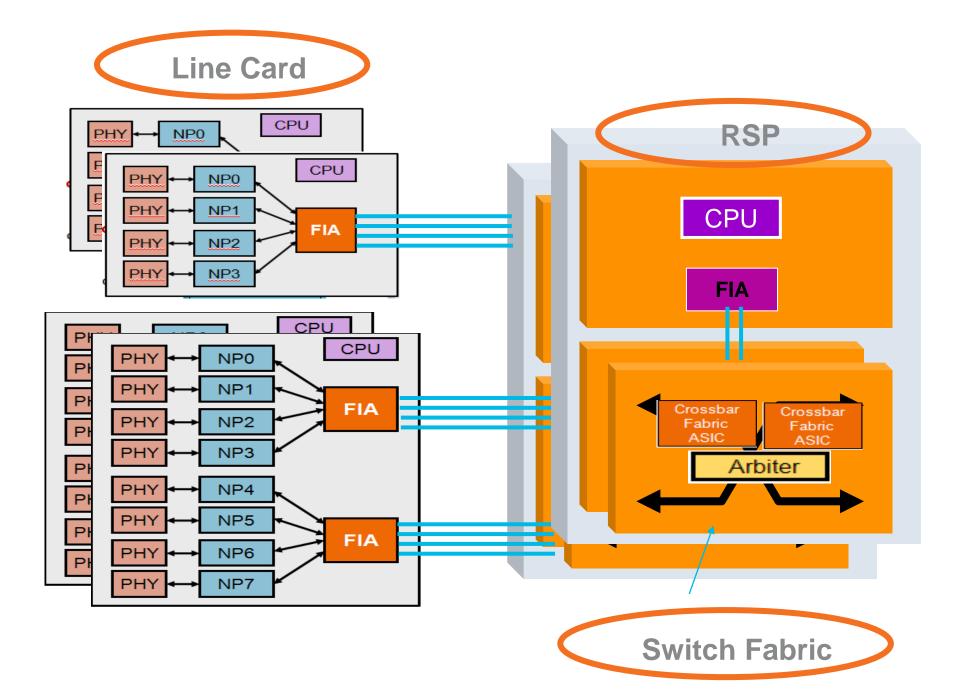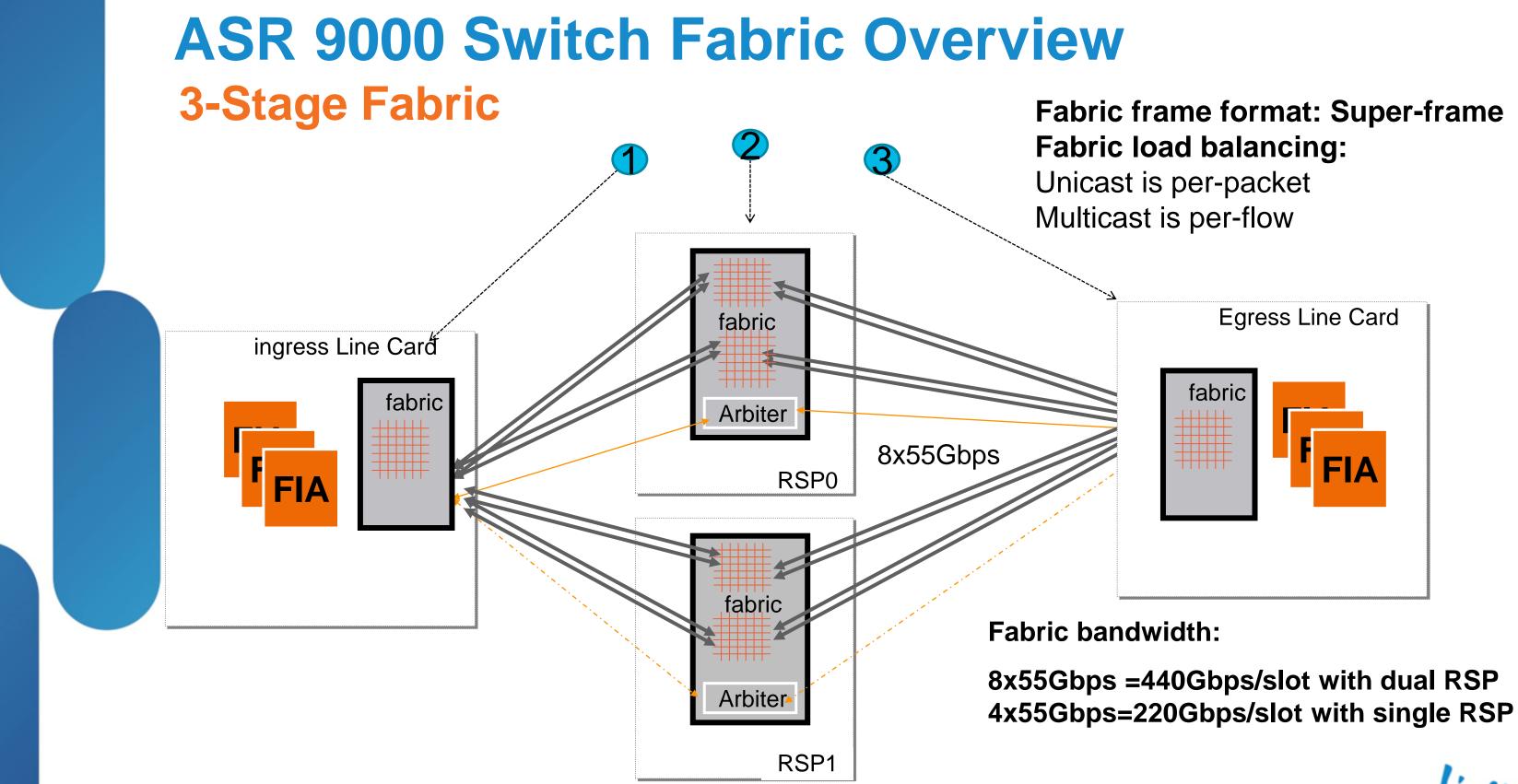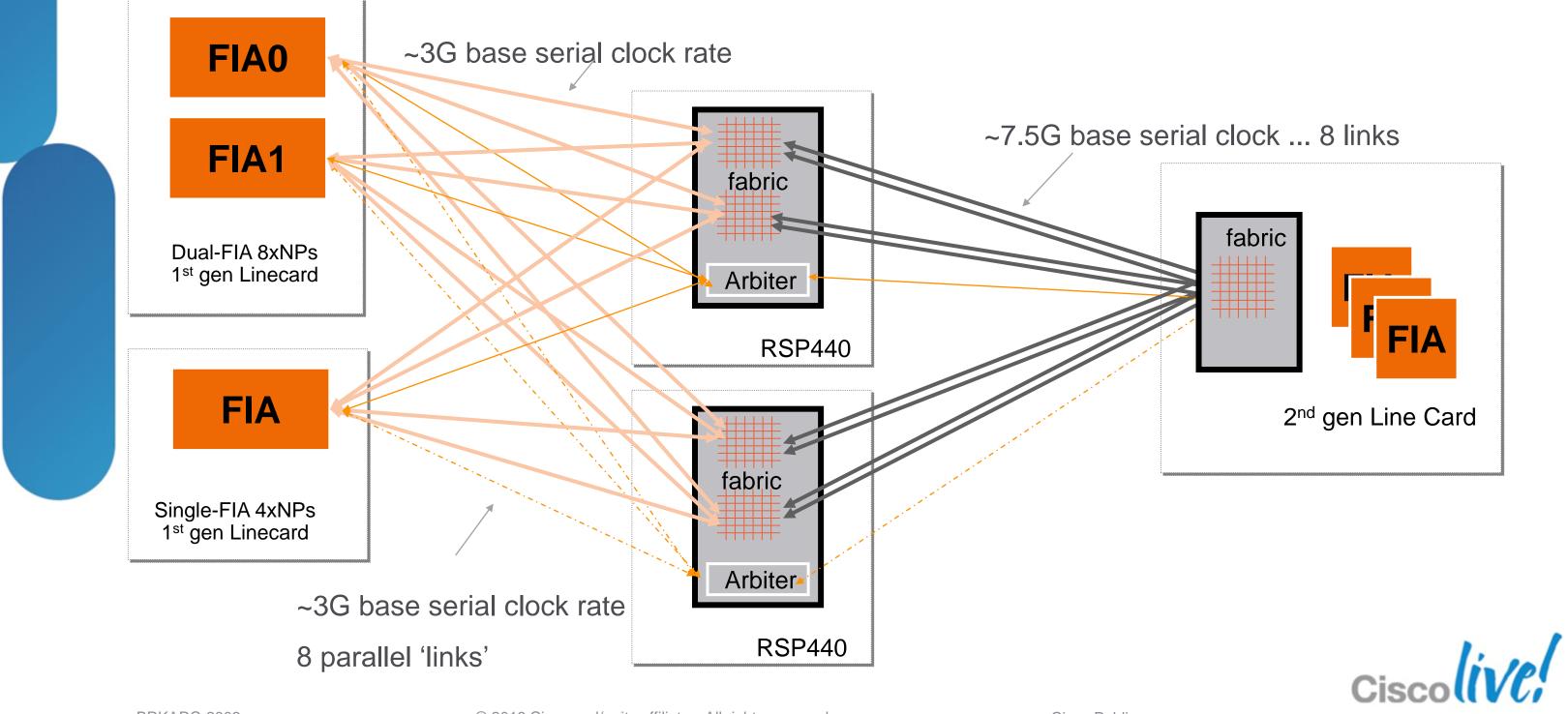# ASR 9000 System Architecture
"At-a-Glance"

Line Card

RSP

CPU

PHY — NP0 — CPU

PHY — NP0
PHY — NP1
PHY — NP2 — FIA
PHY — NP3

CPU

FIA

PHY — NP0
PHY — NP1
PHY — NP2 — FIA
PHY — NP3
PHY — NP4
PHY — NP5
PHY — NP6 — FIA
PHY — NP7

Crossbar Fabric ASIC   Crossbar Fabric ASIC

Arbiter

Switch Fabric

Cisco Public

ASR 9000 Switch Fabric Overview
3-Stage Fabric

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

ingress Line Card

FIA

fabric

① ② ③

fabric

Arbiter

RSP0

8x55Gbps

fabric

Arbiter

RSP1

Egress Line Card

fabric

FIA

Fabric bandwidth:

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

Cisco Public

Cisco live!

1ˢᵗ/2ⁿᵈ Generation Switch Fabric Compatibility

FIA0

FIA1

Dual-FIA 8xNPs
1ˢᵗ gen Linecard

FIA

Single-FIA 4xNPs
1ˢᵗ gen Linecard

~3G base serial clock rate

fabric

Arbiter

RSP440

fabric

Arbiter

RSP440

~3G base serial clock rate

8 parallel 'links'

~7.5G base serial clock ... 8 links

fabric

FIA

2ⁿᵈ gen Line Card

Cisco Public

Fabric Forwards/Backwards Compatibility

- 1st generation linecards based on 3.25Ghz serial links

- 2nd gen linecards/fabric based on 7.5Ghz

- In 2nd generation, both the linecard fabric interface ASICs and the fabric crossbars can negotiate to higher or slower clock rate

- Provides forward/backwards compatibility in fabric

Cisco Public
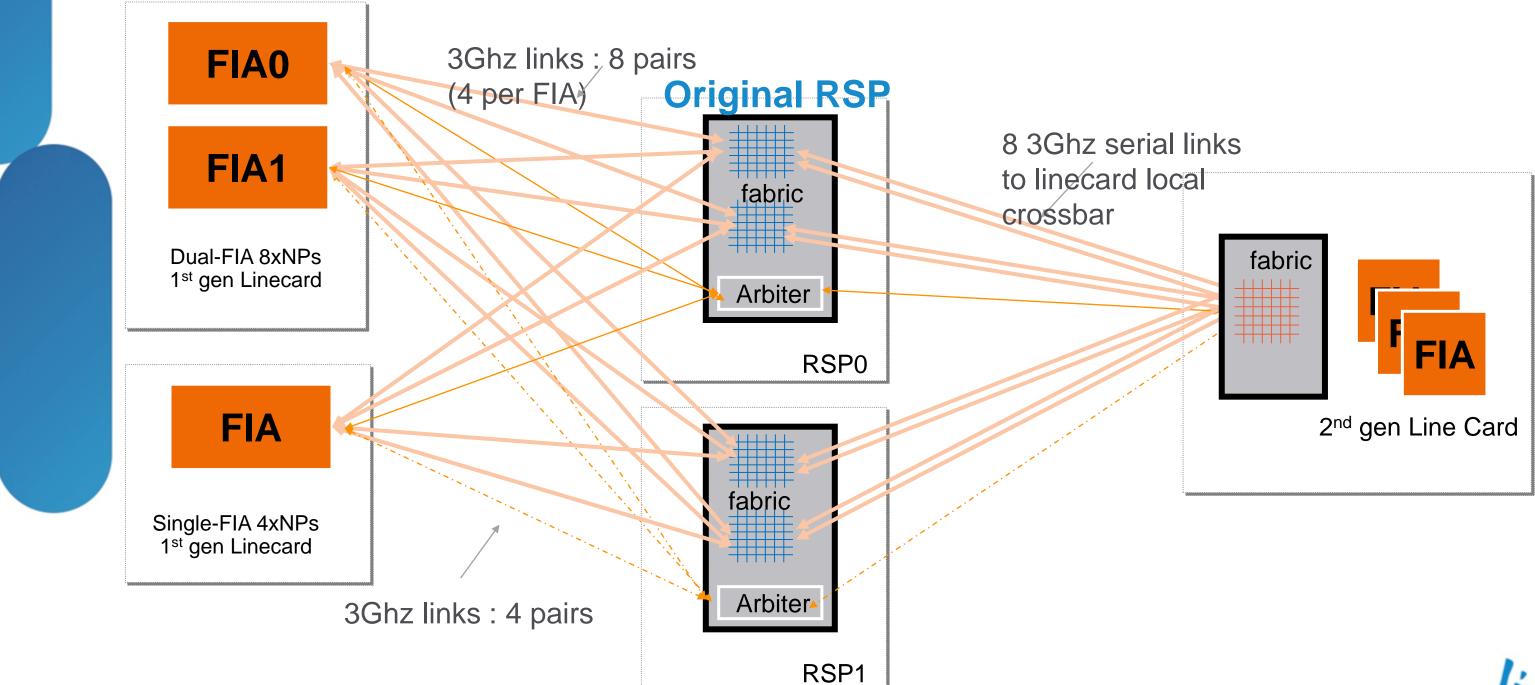
Switch Fabric Scalability in 3 Dimensions

- Base clock rate of the discrete serial components
  - 3.125Ghz, 6.25Ghz, 7.5Ghz, 15Ghz... ???
- Number of parallel serial links that make up a "fabric port" (usually either 4x or 8x on the asr9k)
- Number of parallel fabric "planes" in the system...

- Fabric planes provide redundancy (N+1) for switch fabric complex
- Example: 4 planes * 8 links * 3.125Ghz
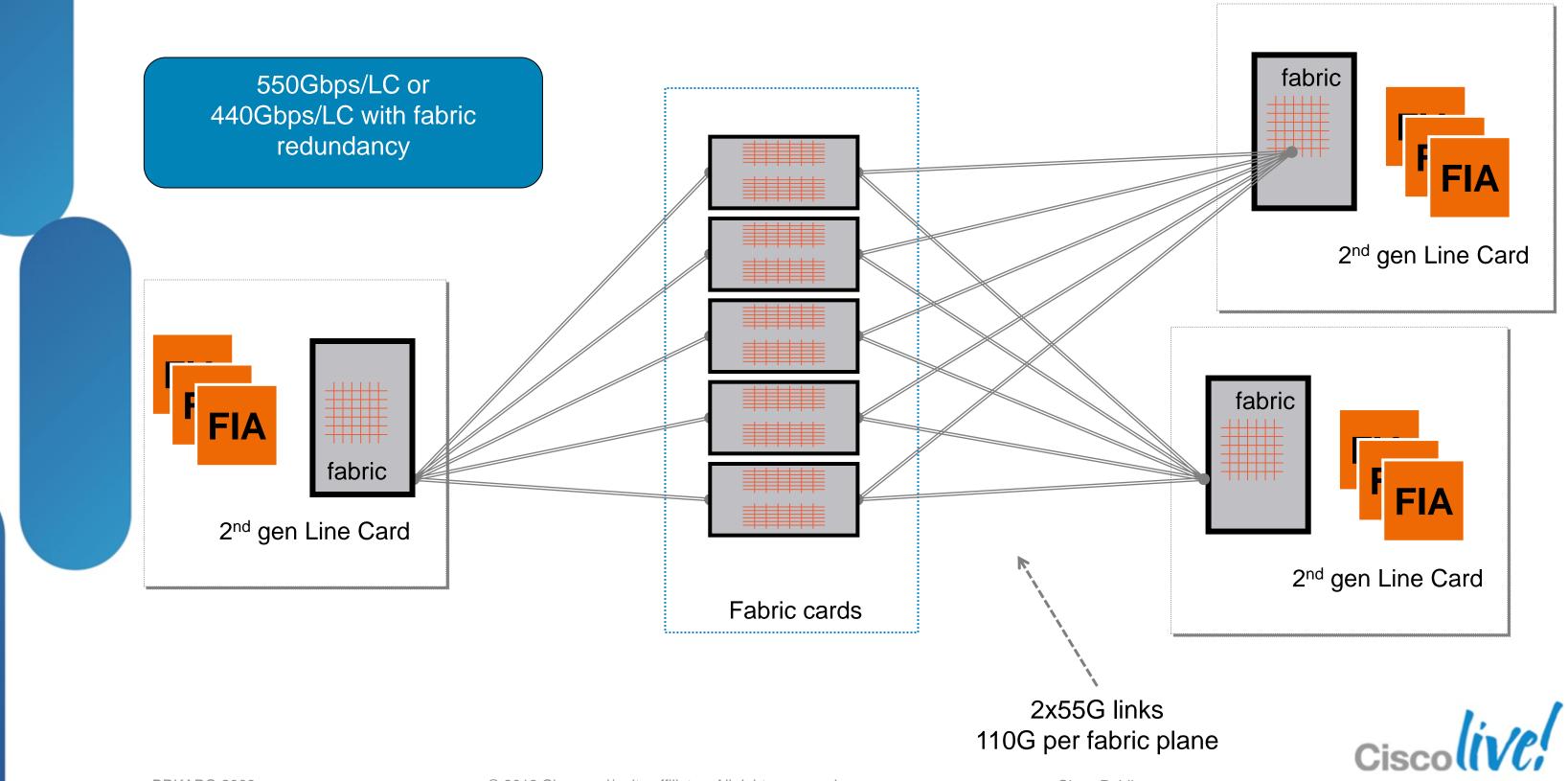 - Yields → ~92Gb (80g linecards @ 15% fab speedup)

1st/2nd Generation Switch Fabric Compatibility

Original RSP

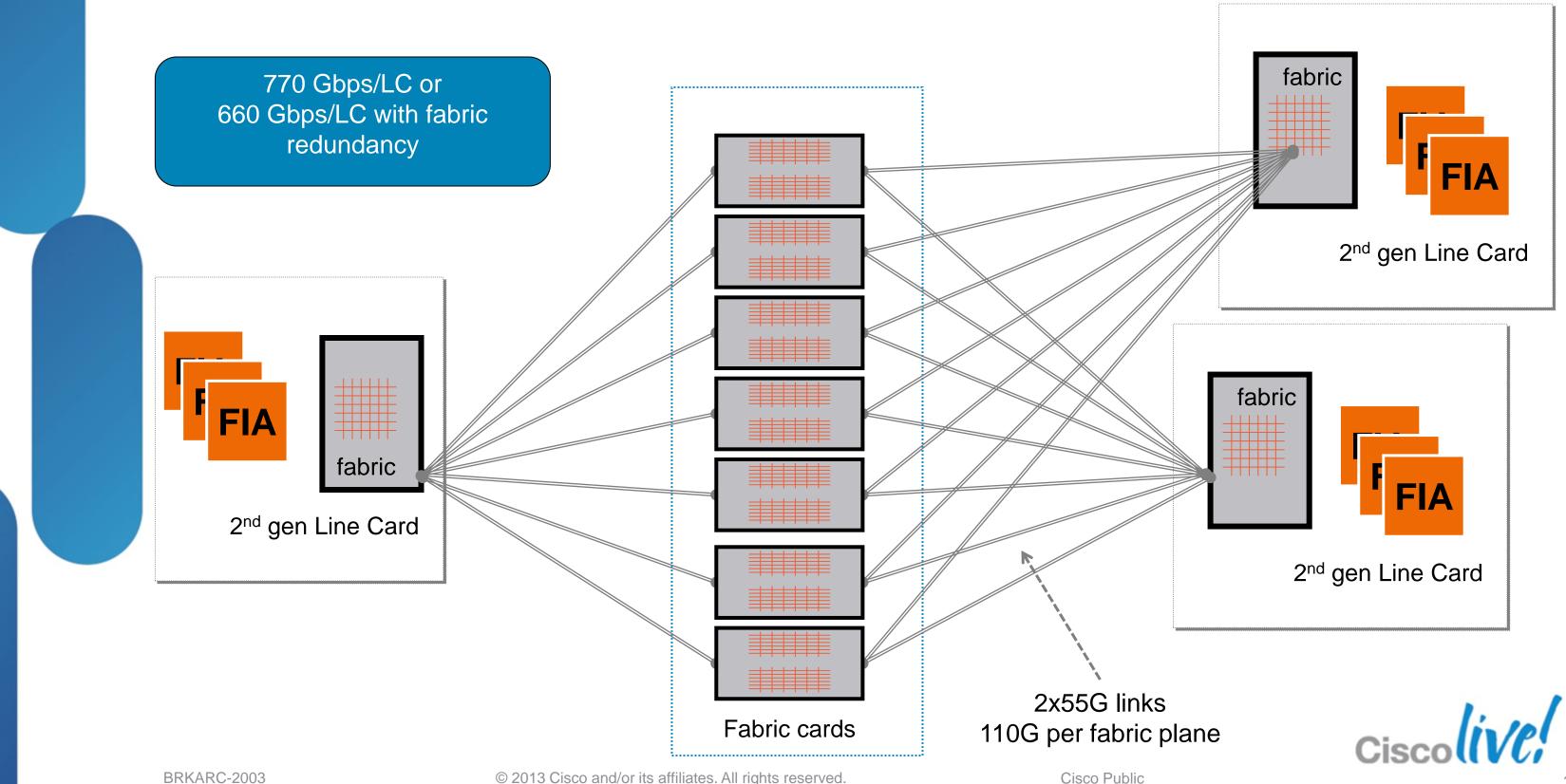FIA0

FIA1

Dual-FIA 8xNPs
1st gen Linecard

3Ghz links : 8 pairs
(4 per FIA)

fabric

Arbiter

RSP0

FIA

Single-FIA 4xNPs
1st gen Linecard

3Ghz links : 4 pairs

fabric

Arbiter

RSP1

8 3Ghz serial links
to linecard local
crossbar

fabric

FIA

2nd gen Line Card

Cisco Public

ASR 9922 Fabric Architecture : 5-plane System

550Gbps/LC or 440Gbps/LC with fabric redundancy

fabric

FIA

2nd gen Line Card

FIA

fabric

2nd gen Line Card

Fabric cards

fabric

FIA

2nd gen Line Card

2x55G links
110G per fabric plane

Cisco Public

ASR 9922 Fabric Architecture : 7-plane System

770 Gbps/LC or
660 Gbps/LC with fabric redundancy

fabric

FIA

2nd gen Line Card

FIA

fabric

2nd gen Line Card

fabric

FIA

2nd gen Line Card

Fabric cards

2x55G links
110G per fabric plane

Generic LC Architecture (1) – Components

Pluggable physical interfaces **PHY**
- speeds: GE, 10GE, 40GE, 100GE
- form factors: SFP, SFP+, XFP, QSFP, CFP
- media/reach: T, SR, LR, ZR, LR4, SR10
- colours: gray, CWDM, DWDM, Tunable
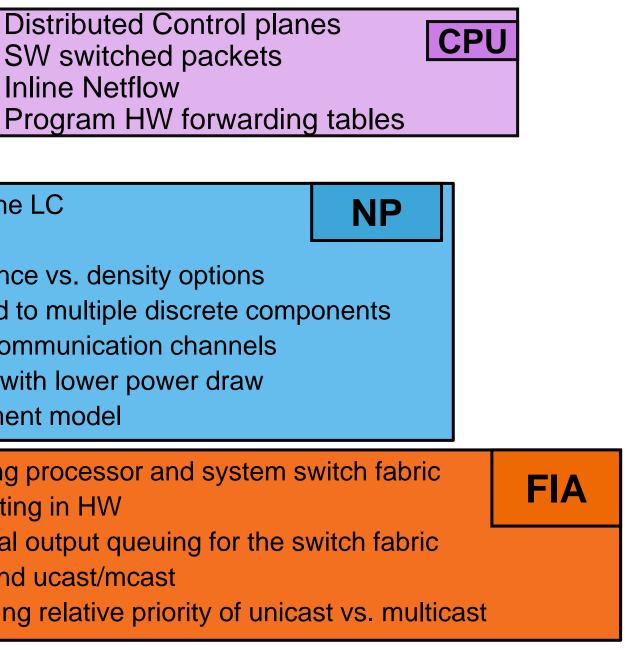
Distributed Control planes **CPU**
SW switched packets
Inline Netflow
Program HW forwarding tables

NP
- forwarding and feature engine for the LC
- scales bandwidth via multiple NPs
  - up to 8 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
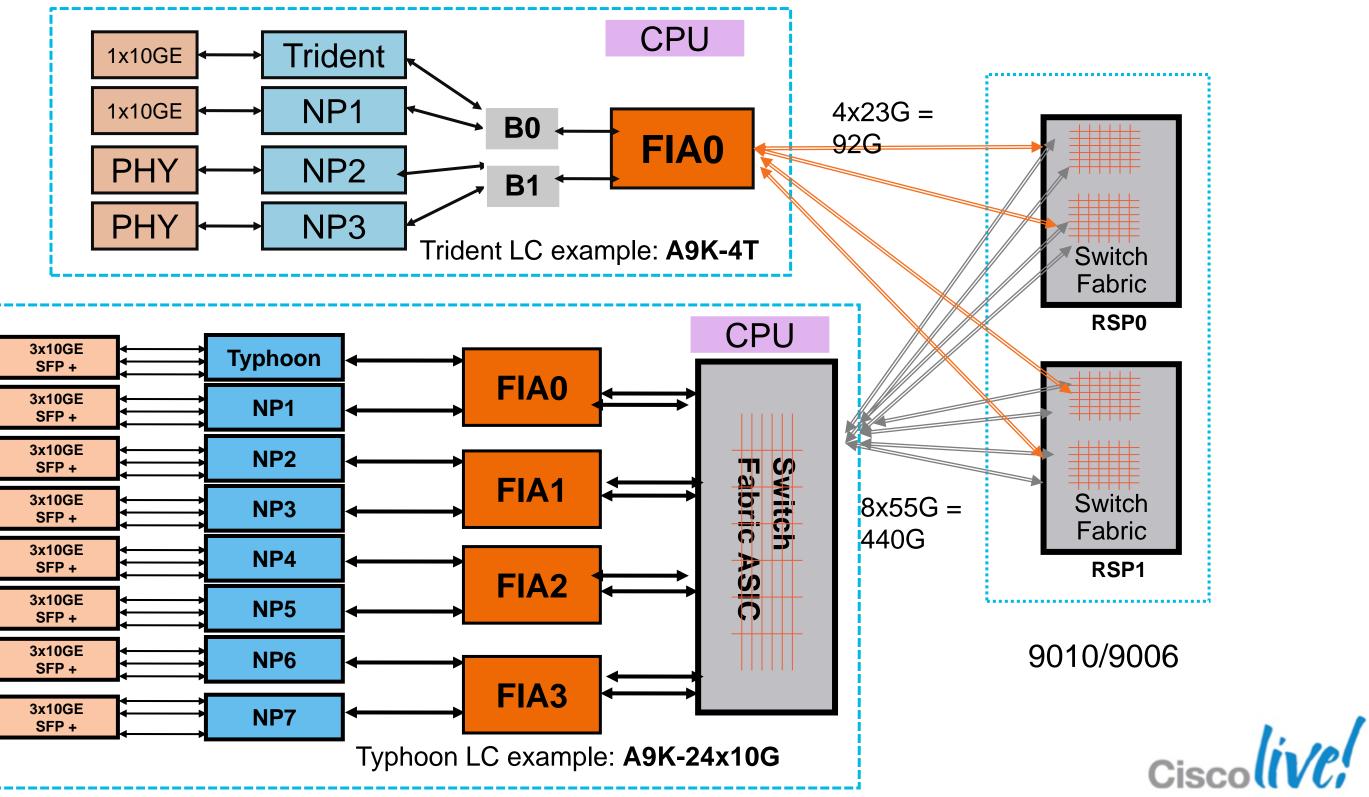
FIA
- interface between forwarding processor and system switch fabric
- arbitration, framing, accounting in HW
- provides buffering and virtual output queuing for the switch fabric
- QoS awareness for Hi/Lo and ucast/mcast
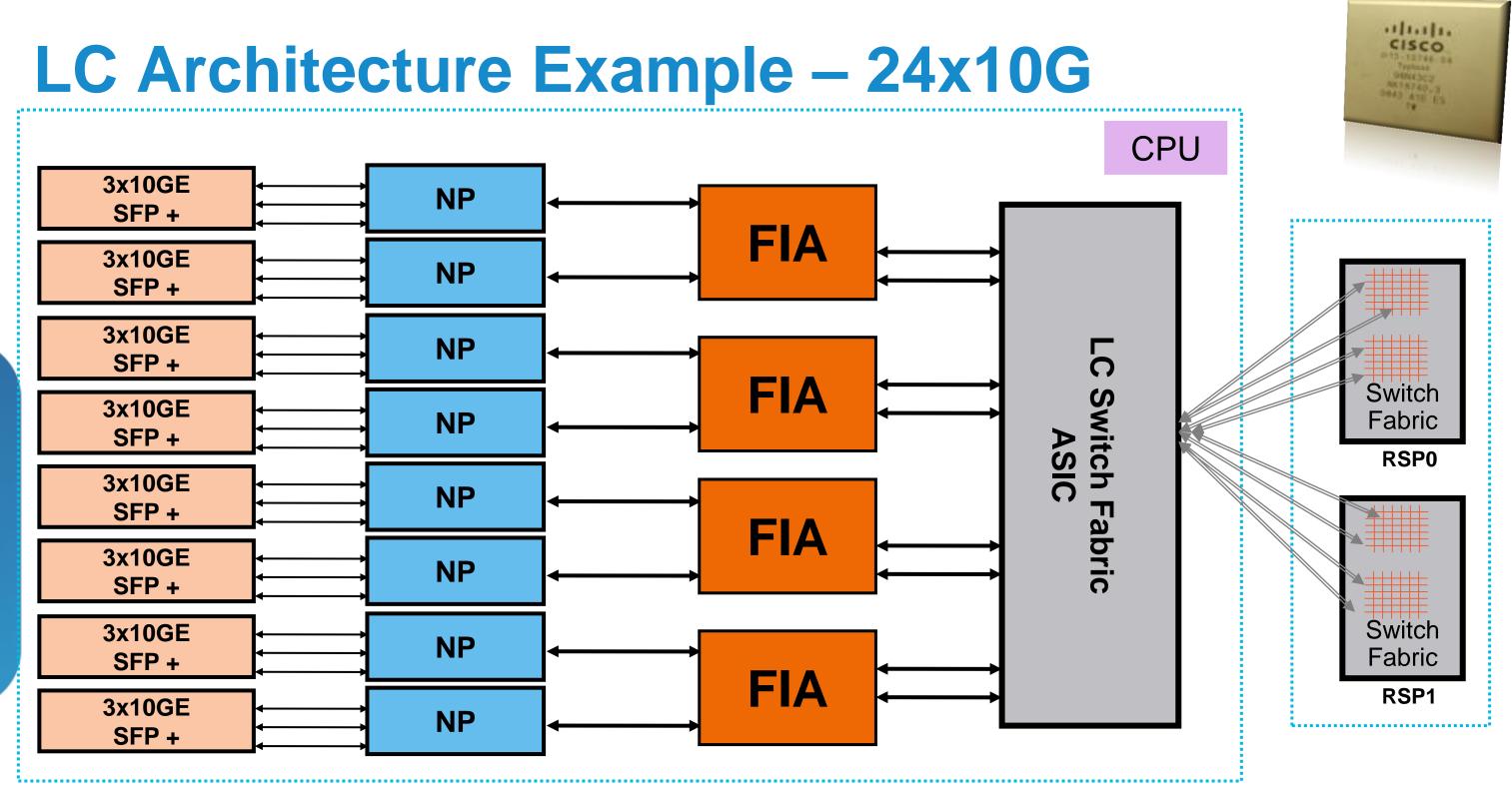 - total flexibility regarding relative priority of unicast vs. multicast

Cisco Public

ASR 9000 Line Card Architecture Overview

Trident LC example: **A9K-4T**

Typhoon LC example: **A9K-24x10G**

4x23G = 92G

8x55G = 440G

RSP0

RSP1

9010/9006

LC Architecture Example – 24x10G

CPU

3x10GE SFP +	NP
3x10GE SFP +	NP	→ FIA
3x10GE SFP +	NP
3x10GE SFP +	NP	→ FIA
3x10GE SFP +	NP
3x10GE SFP +	NP	→ FIA
3x10GE SFP +	NP
3x10GE SFP +	NP	→ FIA

LC Switch Fabric ASIC

Switch Fabric
RSP0

Switch Fabric
RSP1
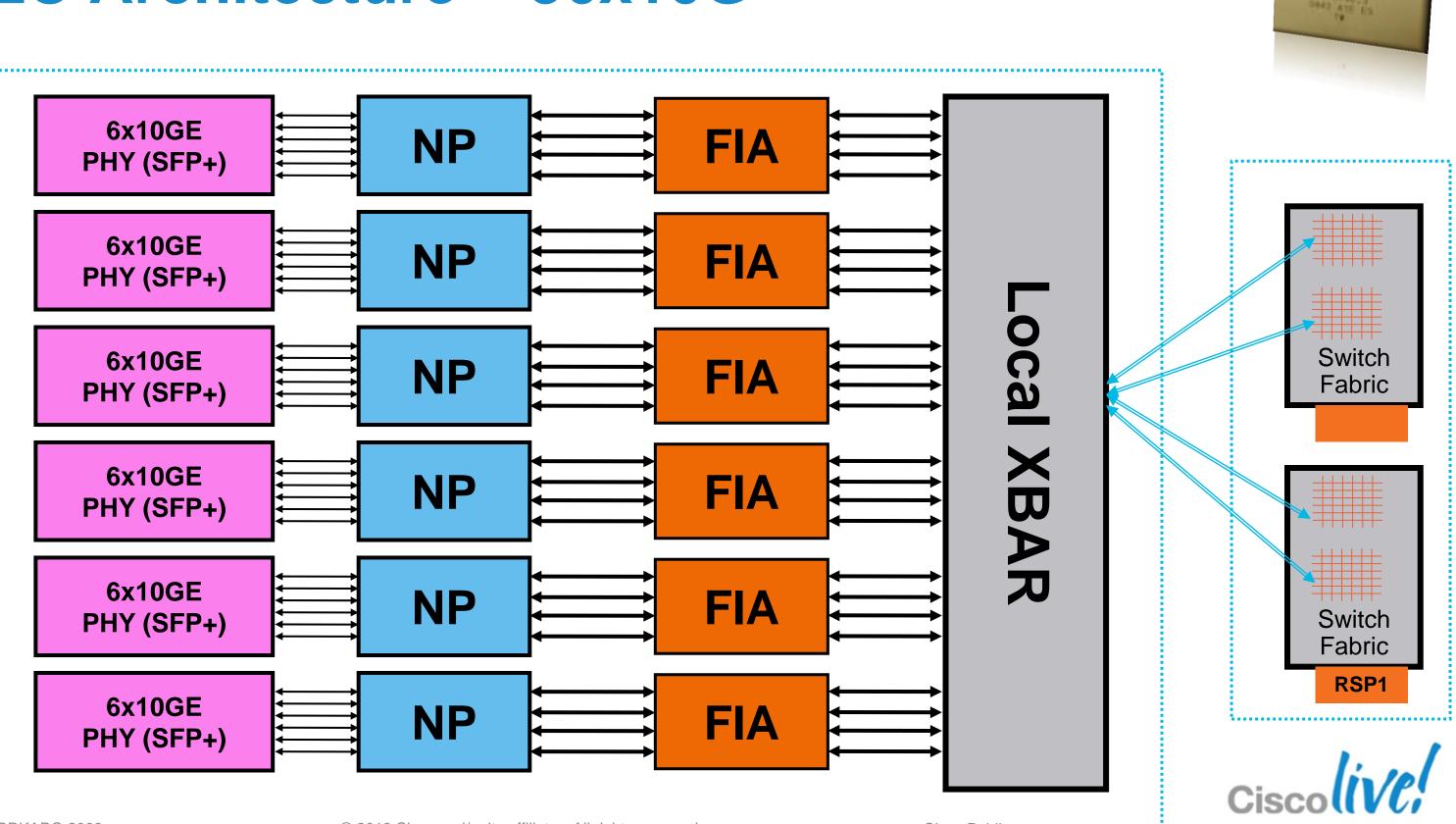
Each NP:
60Gbps bi-directional
120Gbps uni-directional
2x45Mpps

Each FIA:
60Gbps bi-directional
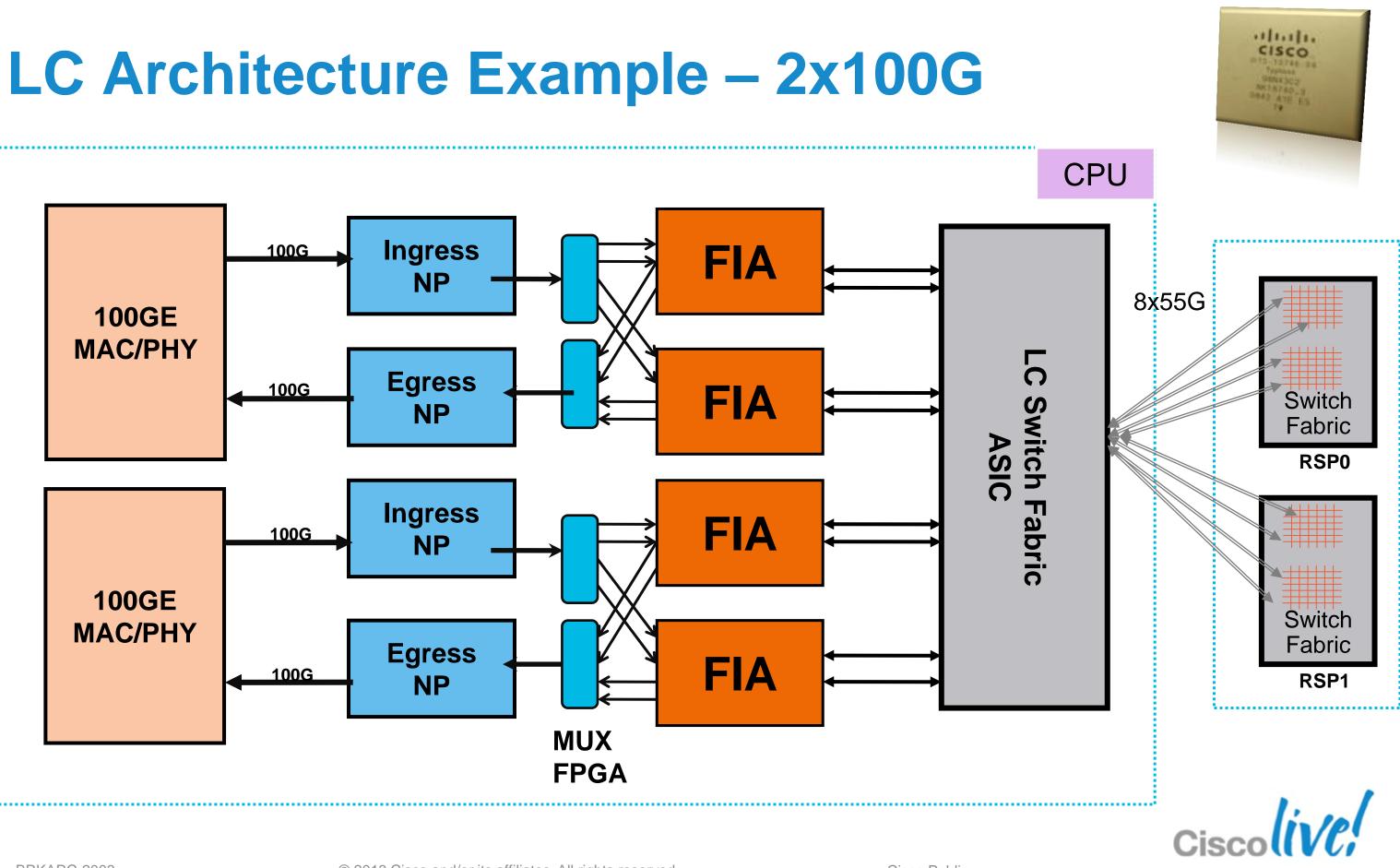
LC Architecture – 36x10G

Cisco Public

LC Architecture Example – 2x100G

 Cisco Public
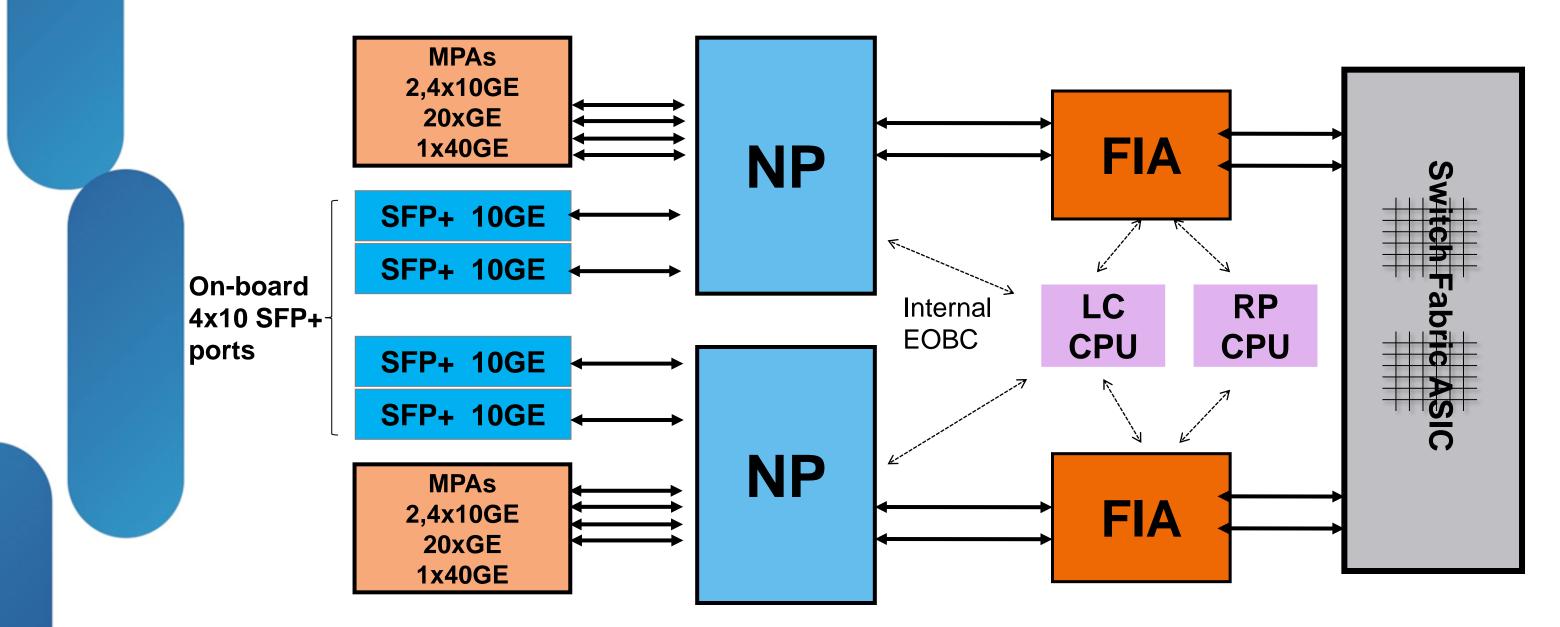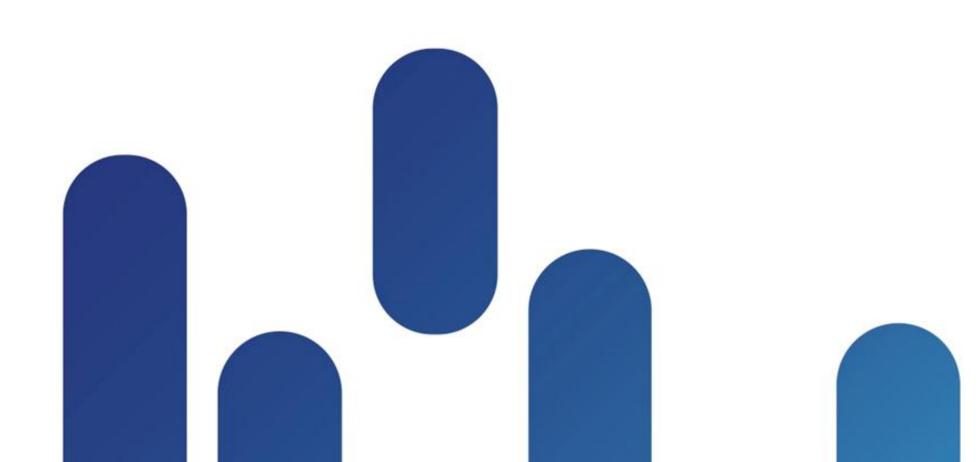
ASR 9001 Architecture Overview
Identical components and architecture

On-board 4x10 SFP+ ports

MPAs
2,4x10GE
20xGE
1x40GE

SFP+ 10GE

SFP+ 10GE

SFP+ 10GE

SFP+ 10GE

MPAs
2,4x10GE
20xGE
1x40GE

NP

NP

Internal EOBC

FIA

FIA

LC CPU

RP CPU

Switch Fabric ASIC

ASR 9000 Packet Forwarding Infrastructure

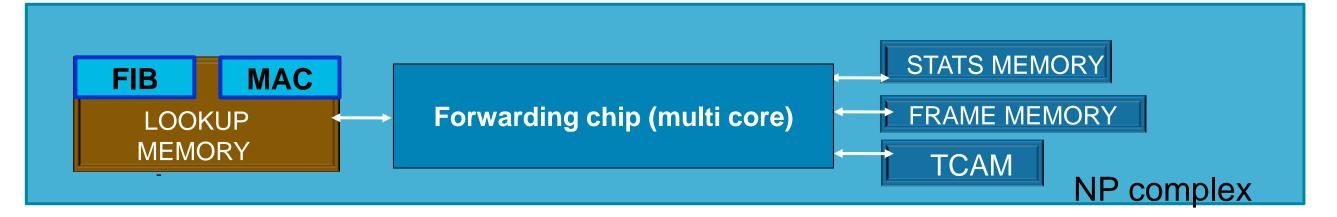# Architectural Choices: Forwarding Paradigms

- **Three high level options**

- **Fully integrated ASIC**

  - Optimised for : highest performance, unit cost/power

  - Sacrifices: TTM, flexibility, extensibility

- **Network processor / customized microcode**

  - Optimised for : good flexibility with high performance / BW

  - Sacrifices: moderate cost, programming complexity

- **General purpose CPU / stack programming model**

  - Optimised: development, most flexible, TTM, complex code

  - Sacrifices: performance, bandwidth, unit cost/power

NP Architecture Details

FIB **MAC**

LOOKUP MEMORY

Forwarding chip (multi core)

STATS MEMORY

FRAME MEMORY

TCAM
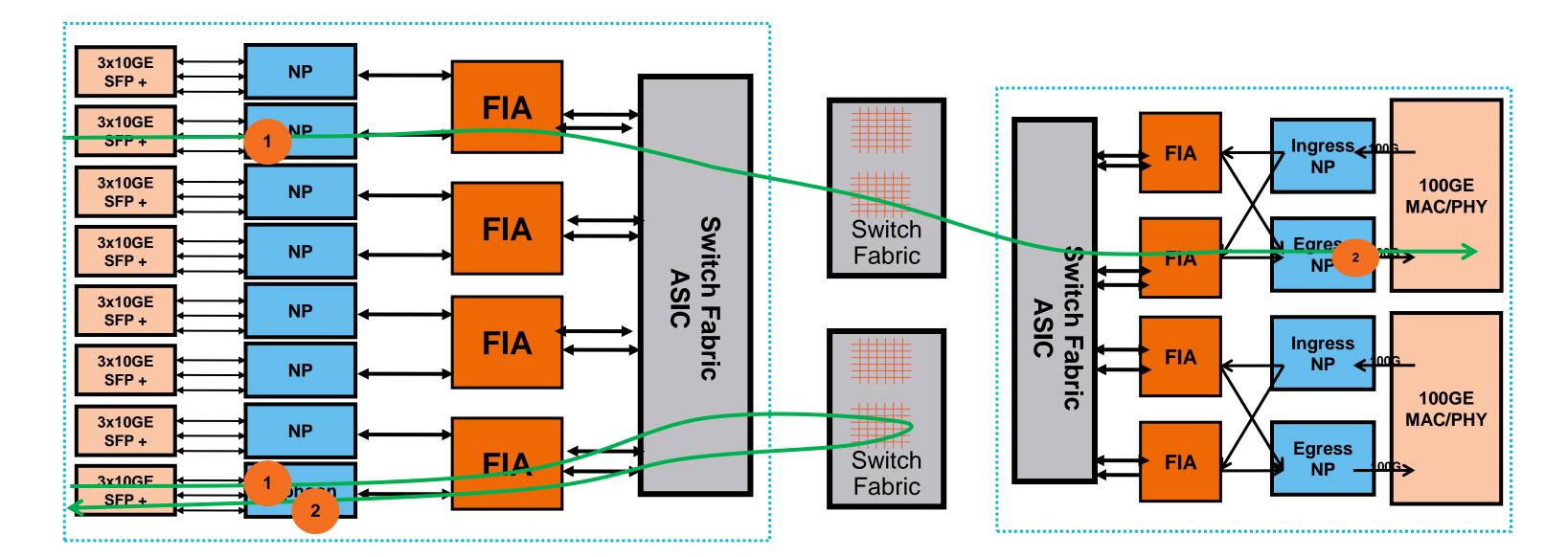
NP complex

- 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, -L/-B/-E**

  - **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**

IOS-XR 2-Stage Forwarding Overview
For high system scale

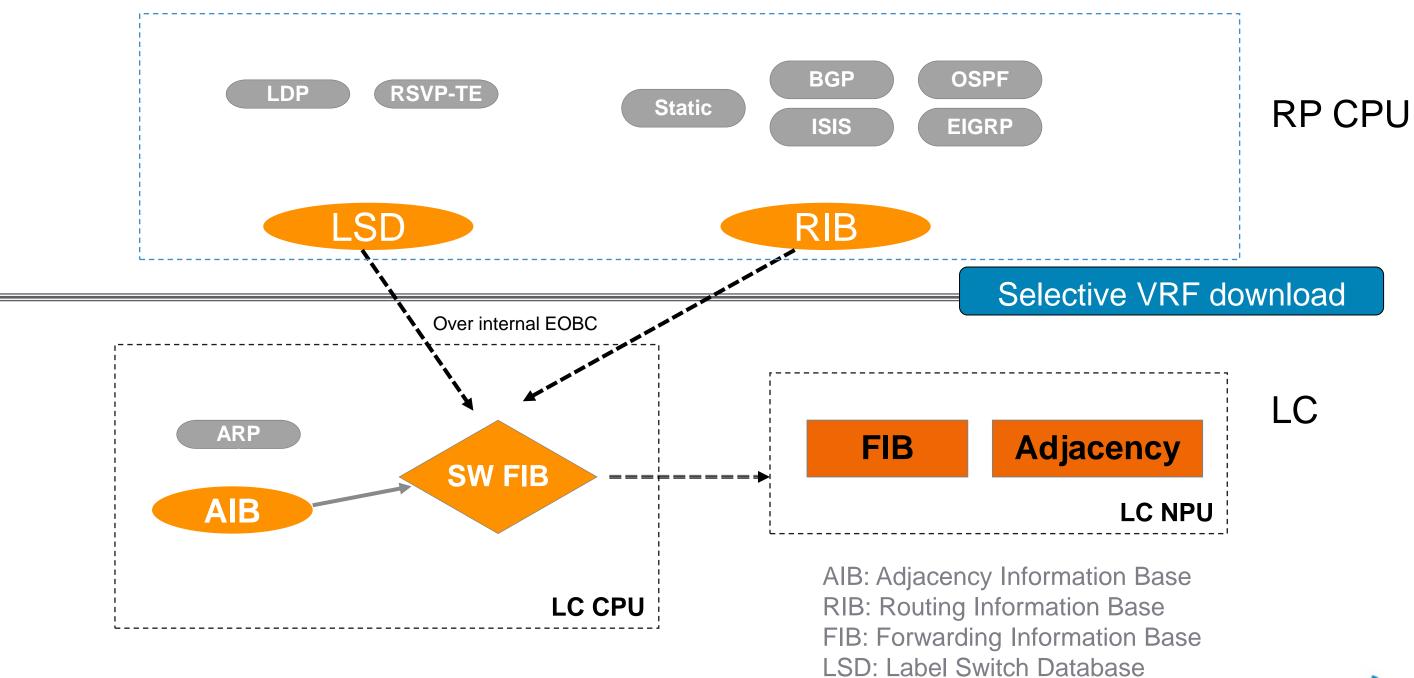Uniform packet flow for simplicity and predictable performance
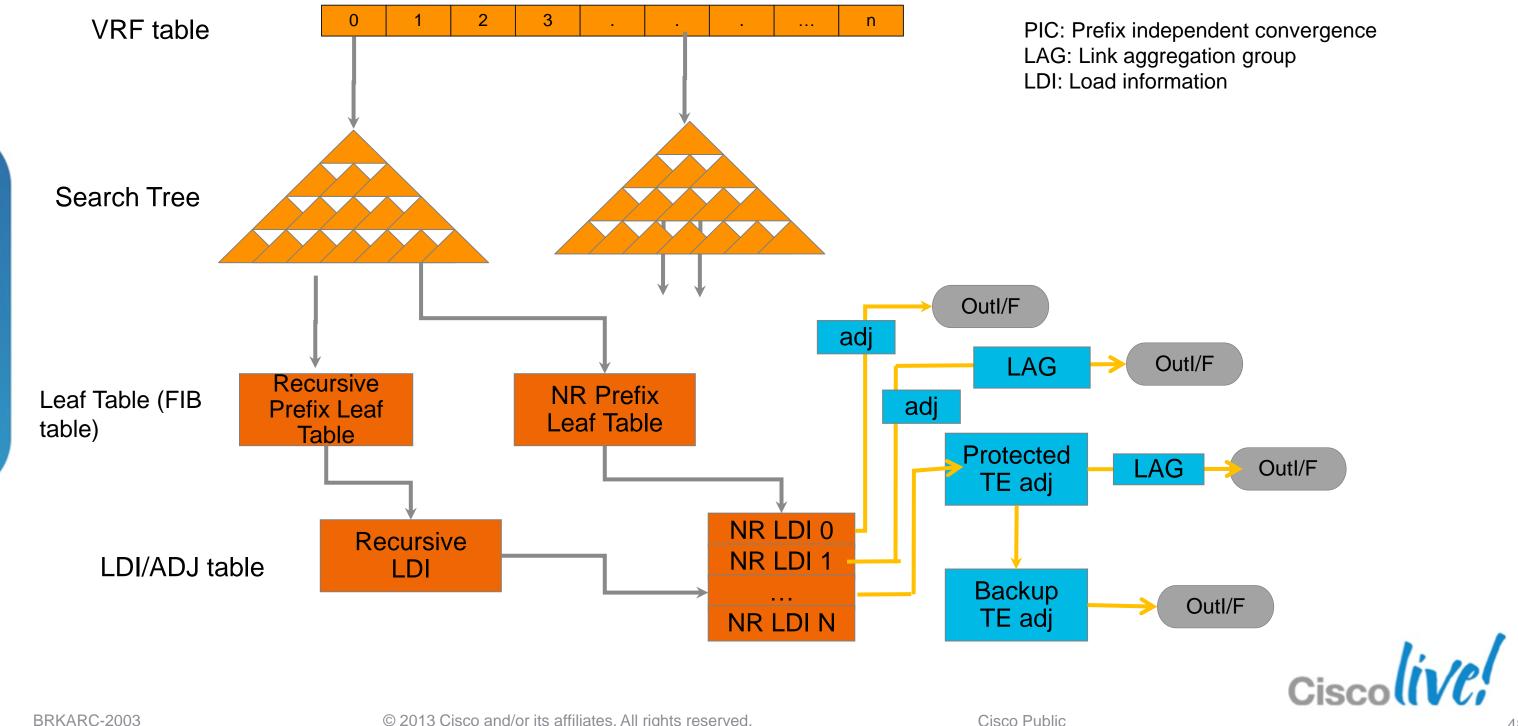
L3 Control Plane Overview

RP CPU

LDP   RSVP-TE   Static   BGP   OSPF
                                  ISIS   EIGRP

LSD   RIB

Selective VRF download

Over internal EOBC

LC

ARP

SW FIB   FIB   Adjacency

AIB   LC NPU

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

LC CPU

Cisco Public

L3 Forwarding – Hierarchical Data Structure
for prefix independent fast convergence: BGP, TE/FRR, LAG

VRF table

0	1	2	3	.	.	.	…	n

PIC: Prefix independent convergence
LAG: Link aggregation group
LDI: Load information

Search Tree

Leaf Table (FIB table)

Recursive Prefix Leaf Table

NR Prefix Leaf Table

adj

OutI/F

LAG → OutI/F

adj

LDI/ADJ table

Recursive LDI

NR LDI 0
NR LDI 1
…
NR LDI N

Protected TE adj → LAG → OutI/F

Backup TE adj → OutI/F

Cisco live!

MAC Learning and Sync

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
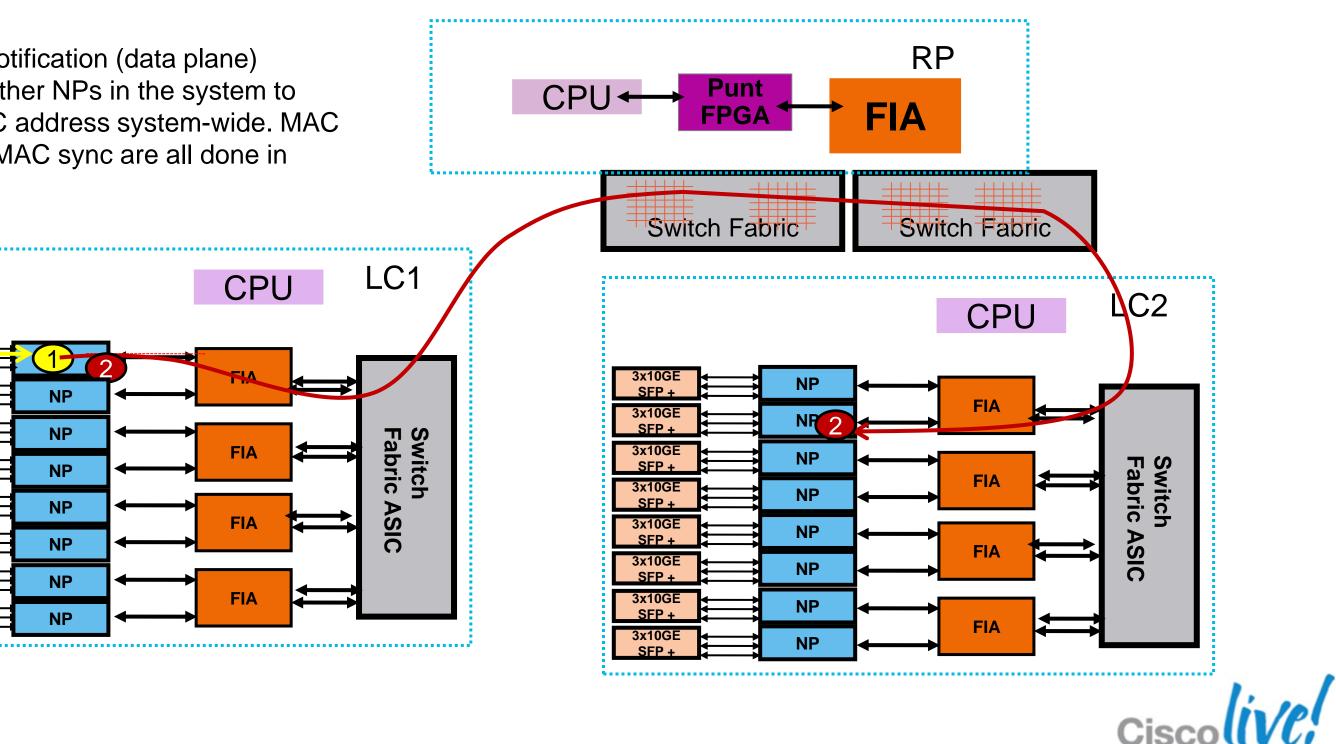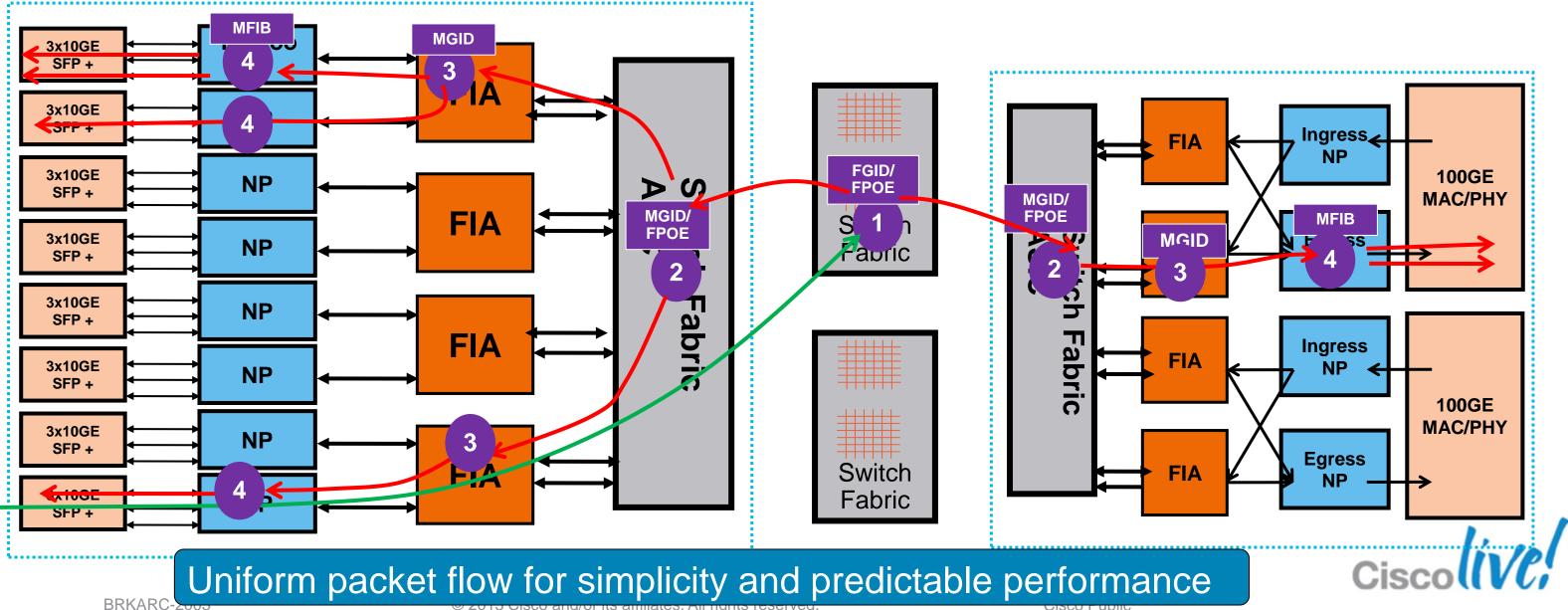
Internal Multicast Replication Overview

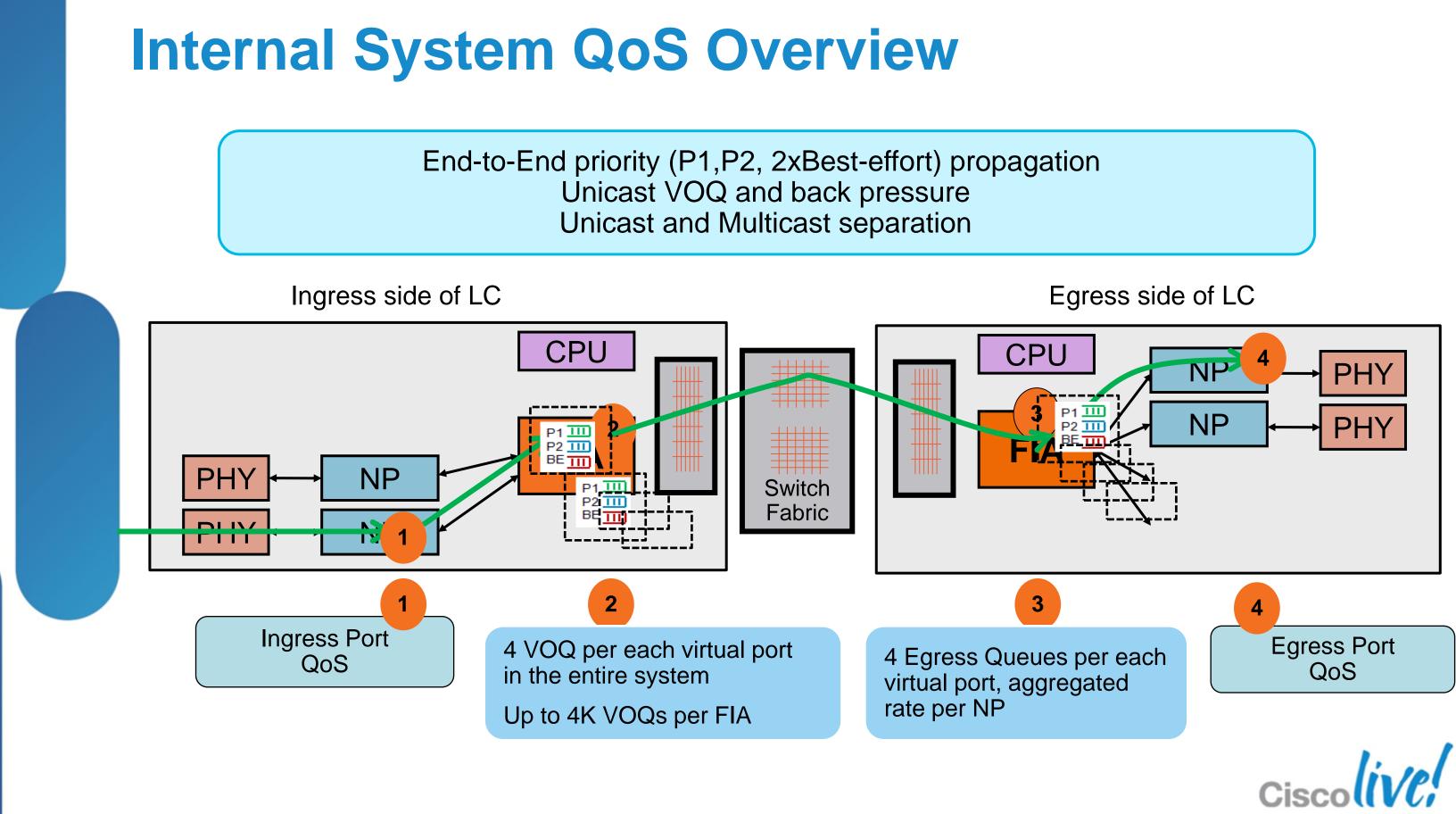1. Fabric to LC Replication
2. LC fabric to FIA Replication
3. FIA to NP Replication
4. NP to egress port Replication

FGID – Fabric Group ID
MGID – Multicast Group ID
MFIB – Multicast Forwarding Information Base

Efficiency: replicate if required
Line rate: for fully loaded chassis
Simple and clean architecture
Predictable performance

Uniform packet flow for simplicity and predictable performance

Internal System QoS Overview

End-to-End priority (P1,P2, 2xBest-effort) propagation
Unicast VOQ and back pressure
Unicast and Multicast separation

Ingress side of LC

Egress side of LC

CPU

CPU

PHY

NP

P1
P2
BE

P1
P2
BE

Switch
Fabric

FIA

P1
P2
BE

NP

PHY

NP

PHY

1

2

3

4

1

**Ingress Port
QoS**

2

4 VOQ per each virtual port
in the entire system

Up to 4K VOQs per FIA

3

4 Egress Queues per each
virtual port, aggregated
rate per NP

4

**Egress Port
QoS**

Cisco Public

Cisco live!

ASR 9000 nV: network Virtualisation
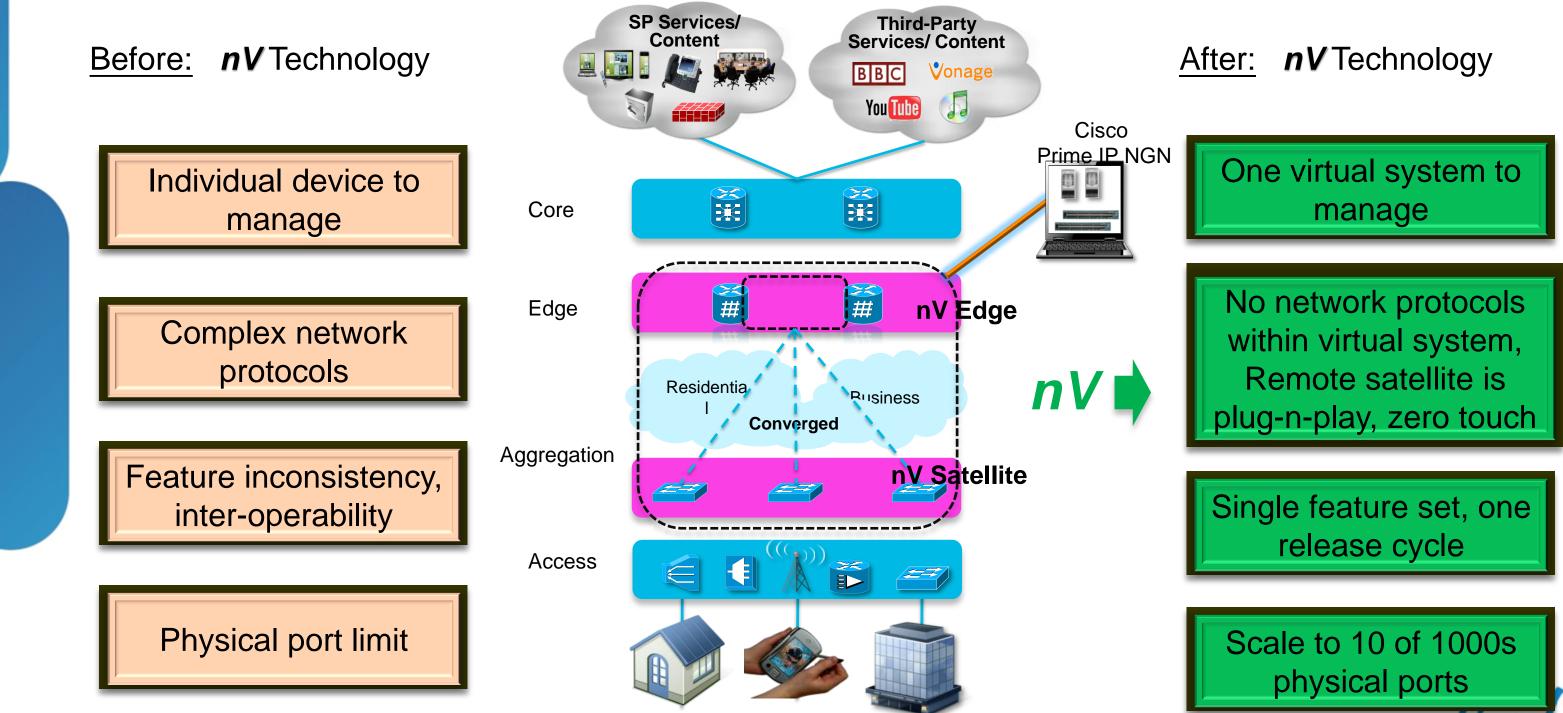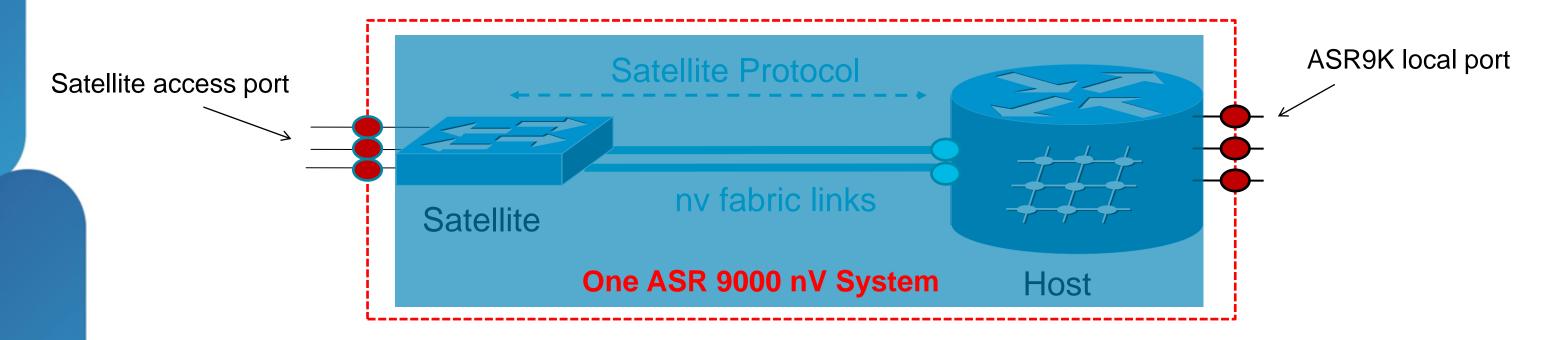
ASR 9000 nV Technology Overview

Before: **nV** Technology

After: **nV** Technology

SP Services/Content

Third-Party Services/ Content

BBC   Vonage   YouTube

Cisco Prime IP NGN

Before	After
Individual device to manage	One virtual system to manage
Complex network protocols	No network protocols within virtual system, Remote satellite is plug-n-play, zero touch
Feature inconsistency, inter-operability	Single feature set, one release cycle
Physical port limit	Scale to 10 of 1000s physical ports

Core

Edge — **nV Edge**

Residential   Business

Converged

Aggregation — **nV Satellite**

Access

nV

Cisco live!

ASR 9000 nV Satellite Overview
Plug-n-play, zero touch satellite access device

Satellite access port → ... Satellite Protocol ... ASR9K local port

nv fabric links

Satellite

One ASR 9000 nV System

Host

- Satellite and ASR 9000 Host run satellite protocol for auto-discovery, provisioning and management

- Satellite and Host could co-locate or in different location. There is no distance limit between satellite and Host

- The connection between satellite and host is called "nv fabric link", which could be L1 or over L2 virtual circuit (future)
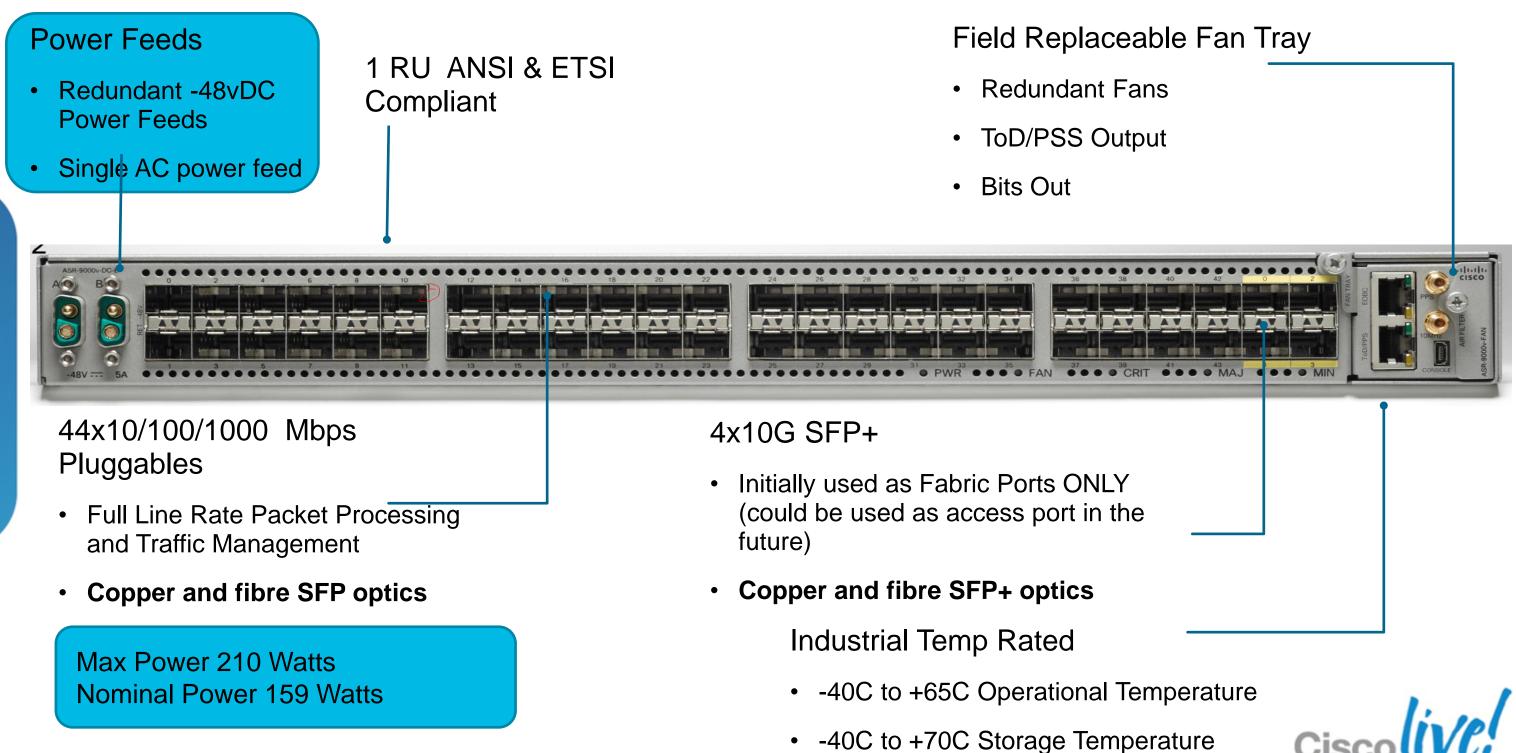
Satellite access port has the same feature parity as ASR9K local port, it works/feels just as local port

Cisco live!
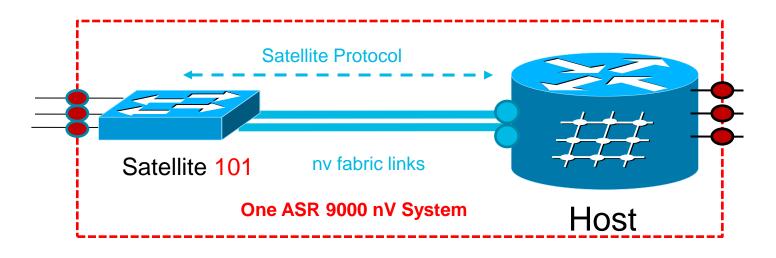
Satellite Hardware: ASR 9000v Overview

Power Feeds

- Redundant -48vDC Power Feeds
- Single AC power feed

1 RU ANSI & ETSI Compliant

Field Replaceable Fan Tray

- Redundant Fans
- ToD/PSS Output
- Bits Out

44x10/100/1000 Mbps Pluggables

- Full Line Rate Packet Processing and Traffic Management
- **Copper and fibre SFP optics**

Max Power 210 Watts
Nominal Power 159 Watts

4x10G SFP+

- Initially used as Fabric Ports ONLY (could be used as access port in the future)
- **Copper and fibre SFP+ optics**

Industrial Temp Rated

- -40C to +65C Operational Temperature
- -40C to +70C Storage Temperature

Satellite Operation – Initial Configuration


```
nv
  satellite 101  ← define satellite
    type asr9000v
    ipv4 address 10.0.0.101


interface TenGigE 0/2/0/2  ← configure satellite fabric port
   ipv4 point-to-point
   ipv4 unnumbered Loopback100
   nv
      satellite-fabric-link satellite 101
        remote-ports  ← satellite to fabric port mapping
           GigabitEthernet 0/0/0-9
```

Cisco Public

CLI example – Configure Satellite Access Port

One ASR 9000 nV System

Satellite Protocol

nv fabric links

Satellite

Host

Satellite access port configuration examples

Local port configuration examples

```
interface GigabitEthernet 101/0/0/1
  ipv4 address 1.2.2.2 255.255.255.0

interface TenGig 101/0/0/1.1
  encapsulation dot1q 101
  rewrite ingress tag pop 1 sym

interface Bundle-ethernet 200
  ipv4 address 1.1.1.1 255.255.255.0

interface GigabitEthernet 101/0/0/2
  bundle-id 200
```
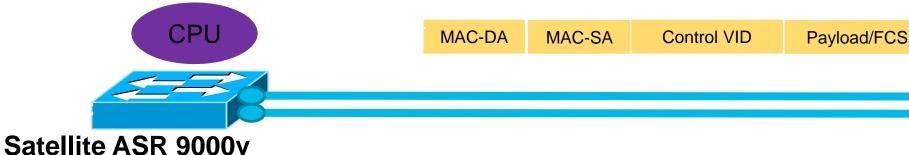
```
interface GigabitEthernet 0/0/0/1
  ipv4 address 2.2.2.2 255.255.255.0

interface TenGig 0/0/0/1.1
  encapsulation dot1q 101
  rewrite ingress tag pop 1 sym

interface Bundle-ethernet 100
  ipv4 address 1.1.1.1 255.255.255.0

interface GigabitEthernet 0/0/0/2
  bundle-id 100
```

Satellite – Host Control Plane
Satellite discovery and control protocol

MAC-DA	MAC-SA	Control VID	Payload/FCS

CPU

CPU

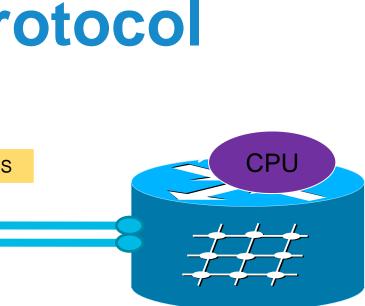**Satellite ASR 9000v**

ASR 9000 Host

Discovery Phase
• A CDP-like link-level protocol that discovers satellites and maintains a periodic heartbeat
• Heartbeat sent once every second, used to detect satellite or fabric link failures. CFM based fast failure detection plan for future release

Control Phase
• Used for Inter-Process Communication between Host and Satellite
• Cisco proprietary protocol over TCP socket for the time being. It could move to standard in the future
• Get/ Set style messages to provision the satellites and also to retrieve notifications from the satellite
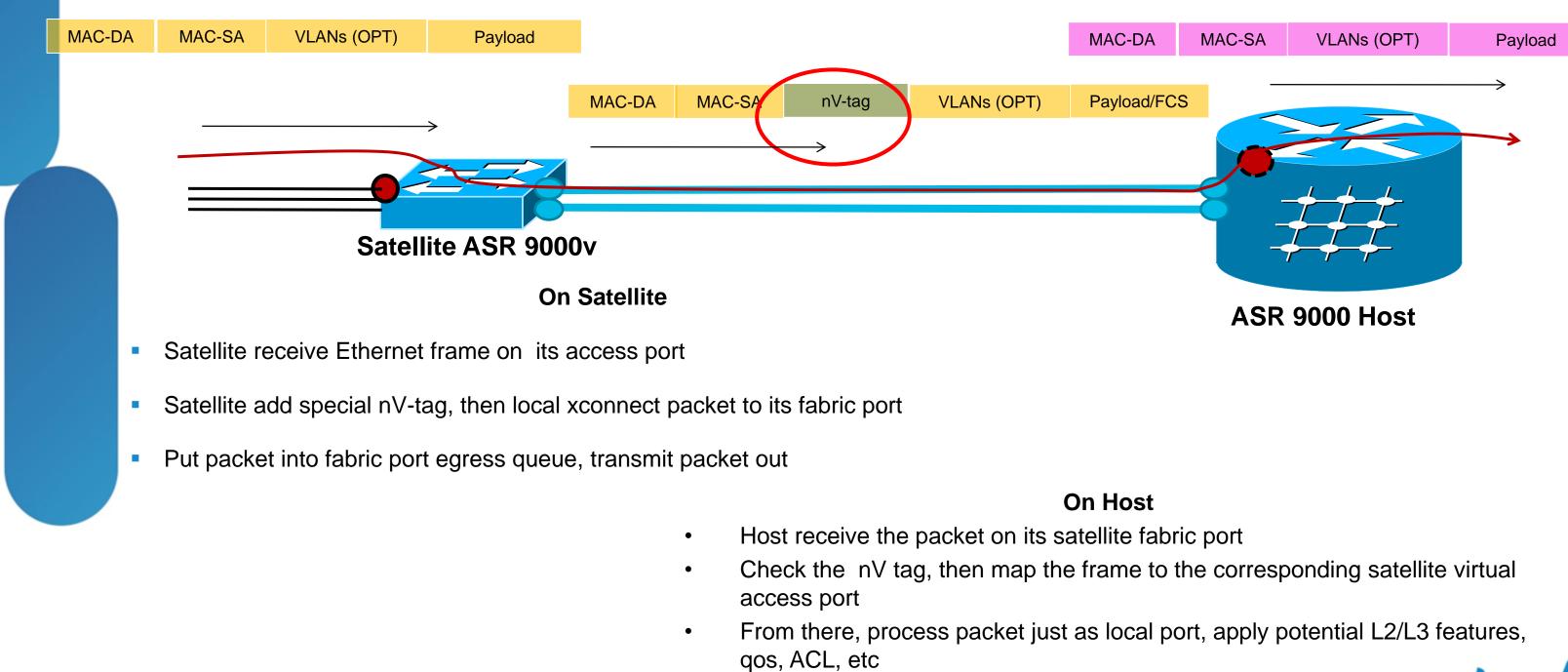
Standardisation is considered for future

Satellite – Host Data Plane Encapsulation

MAC-DA	MAC-SA	VLANs (OPT)	Payload

MAC-DA	MAC-SA	VLANs (OPT)	Payload

MAC-DA	MAC-SA	nV-tag	VLANs (OPT)	Payload/FCS

Satellite ASR 9000v

ASR 9000 Host

On Satellite

- Satellite receive Ethernet frame on its access port

- Satellite add special nV-tag, then local xconnect packet to its fabric port

- Put packet into fabric port egress queue, transmit packet out

On Host

- Host receive the packet on its satellite fabric port
- Check the nV tag, then map the frame to the corresponding satellite virtual access port
- From there, process packet just as local port, apply potential L2/L3 features, qos, ACL, etc
- Packet is forwarded out of local port, or satellite fabric port to same or different satellite

Cisco Public

Satellite Operation – Management

- Normal operation, like show CLIs are done on the Host directly, for example

  Satellite inventory reporting, environmental monitoring

  Interface counts, stats

  SNMP MIB

  NMS support, including ACT, ANA/ PRIME

- Low level debug could still be done directly on the satellite device

  User can telnet into satellite via out-of-band management console, or in-band from Host, and run regular show/debug CLIs

 Cisco Public

Satellite Operation – Image Upgrade

RP/0/RSP0/CPU0:R1#sh install active
  Node 0/RSP0/CPU0 [RP] [SDR: Owner]
    Boot Device: disk0:
    Boot Image: /disk0/asr9k-os-mbi-4.2.1.22K.CSCtz10483-0.0.4.i/0x100305/mbiasr9k-rsp3.vm
    Active Packages:
      disk0:asr9k-px-4.2.1.22K.CSCtz10483-0.0.4.i
      disk0:asr9k-satellite-px-4.2.1.22K ← satellite image PIE
      disk0:asr9k-mini-px-4.2.1.22K
      disk0:asr9k-mpls-px-4.2.1.22K
      disk0:asr9k-mcast-px-4.2.1.22K
      disk0:asr9k-fpd-px-4.2.1.22K

RP/0/RSP0/CPU0:R1#install nv satellite ?
  <100-65534> Satellite ID
  all   All active satellites

RP/0/RSP0/CPU0:R1#install nv satellite 100 ?
  activate Install a new image on the satellite, transferring first if necessary
  transfer Transfer a new image to the satellite, do not install yet

RP/0/RSP0/CPU0:R1#install nv satellite 100 active

> Golden satellite image is always there in satellite flash card for image fall back
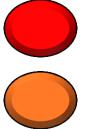
Satellite Operation – Rack Plug and Go

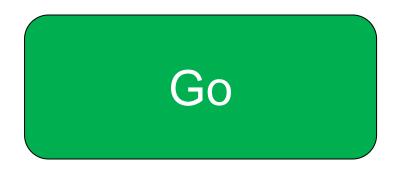| Initial satellite configuration (on ASR9K Host	→	Rack & Plug	→	Go

🔴 Critical Error LED ON → bad hardware, RMA

🟠 Major Error LED ON → Unable to connect to ASR9K host

- Missing the initial satellite configuration?

- L1 issue, at least one of the uplink port light green?
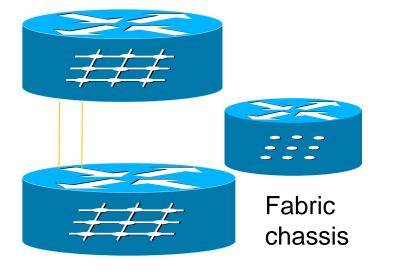
- Security check (optional), is the satellite SN# correct?
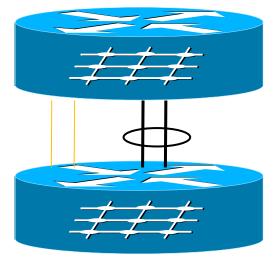
🟢 Status light green → ready to go, satellite is fully managed by Host

Cisco Public

ASR 9000 nV Edge Overview

Simple Network Dual-homing Resiliency & Extensiable Node Capacity

Fabric chassis

Leverage existing IOS-XR CRS multi-chassis SW infrastructure Simplified/Enhanced for ASR 9000 nV Edge

CRS Multi-Chassis

ASR 9000 nV Edge

Single control plane, single management plane, fully distributed data plane across two physical chassis → one virtual nV system
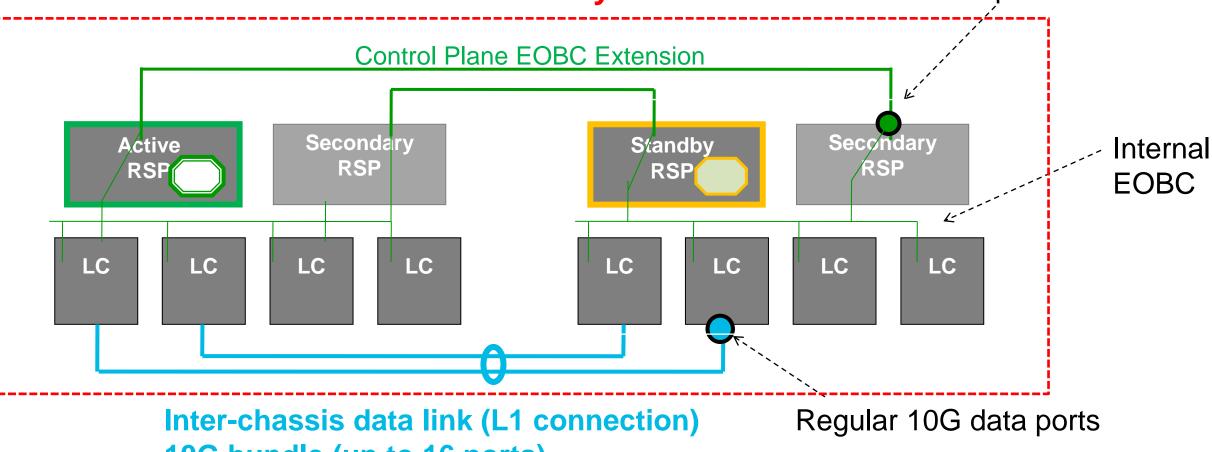
Cisco Public

Cisco live!

nV Edge Architecture Overview

One Virtual ASR 9000 nV System

RSP440 nV EOBC ports for control plane connection

Control Plane EOBC Extension

Active RSP

Secondary RSP

Standby RSP

Secondary RSP

Internal EOBC

LC LC LC LC LC LC LC LC

Inter-chassis data link (L1 connection)
10G bundle (up to 16 ports)

Regular 10G data ports

- **Control plane connection**: Active RSP and standby RSP are on the different chassis, they communicate via external EOBC links

- **Data plane connection**: bundle regular data links into special "nV fabric link" to simulate switch fabric function between two physical chassis for data packet

- **Flexible co-located or different location deployment** (upto 10msec latency)

Cisco Public

nV Edge Configuration

1   **Configure nv Edge globally**

```
nv

 edge-system
   serial FOX1437GC1R rack 1      ← static mapping of chassis serial# and rack#
   serial FOX1439G63M rack 0
```

2   **Configure the inter-chassis fabric(data plane) links**

```
    interface TenGigE1/2/0/0
      nv edge interface

    interface TenGigE0/2/0/0
      nv edge interface
```

NO need to configure the inter-chassis control plane EOBC ports. It's plug-and-play ☺

After this configuration, rack 1 will reload and then join cluster after it boot up
Now you successfully convert two standalone ASR 9000 into one ASR 9000 nV Edge
As simple as this !!!

   Cisco Public

IOS XR Software Overview

Cisco IOS-XR High Level Themes

- **Modular**—Runtime SW upgrade/downgrade support

- **Distributed**—scalable with multi chassis support

- **Platform Independent**—POSIX compliant

- **Management Interface**—Unified Data Model (XML)

- **High Availability**—Hot Standby and Process Restart

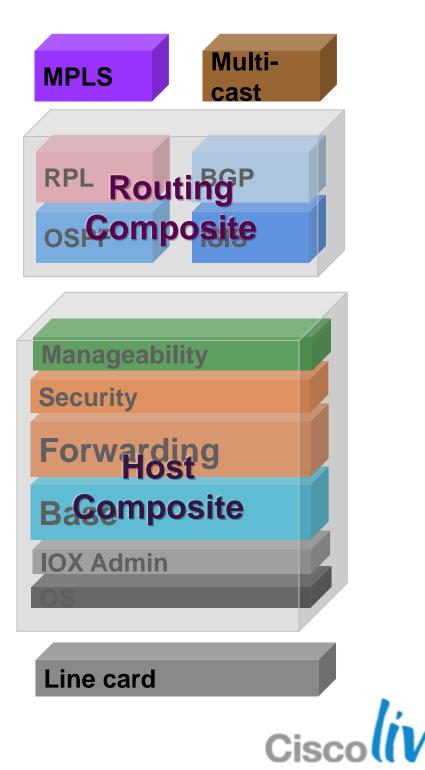- **Security**—Control, Data and Management Plane

Cisco IOS-XR Software Modularity

- Ability to upgrade independently MPLS, Multicast, Routing protocols and Line Cards

- Ability to release software packages async

- Ability to have composites into one manageable unit if desired

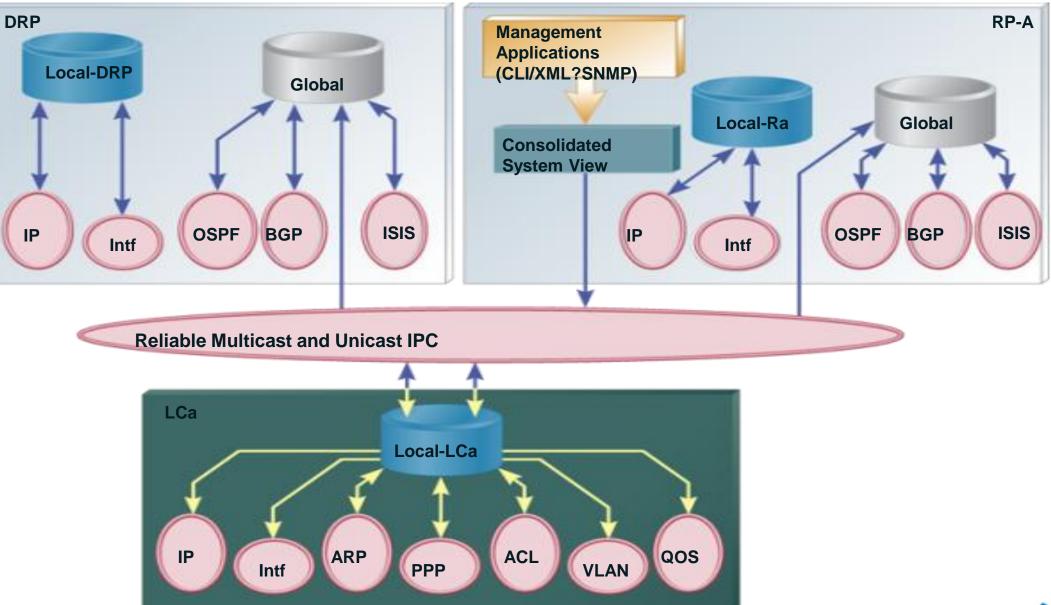- Notion of optional packages if technology not desired on device (Multicast, MPLS)

MPLS

Multi-cast

RPL   **Routing**   BGP
OSPF   **Composite**   ISIS

Manageability

Security

Forwarding   **Host**

Base   **Composite**

IOX Admin

OS

Line card

Distributed In-Memory Database (IMDB)

- Reliable Multicast IPC improves scale and performance

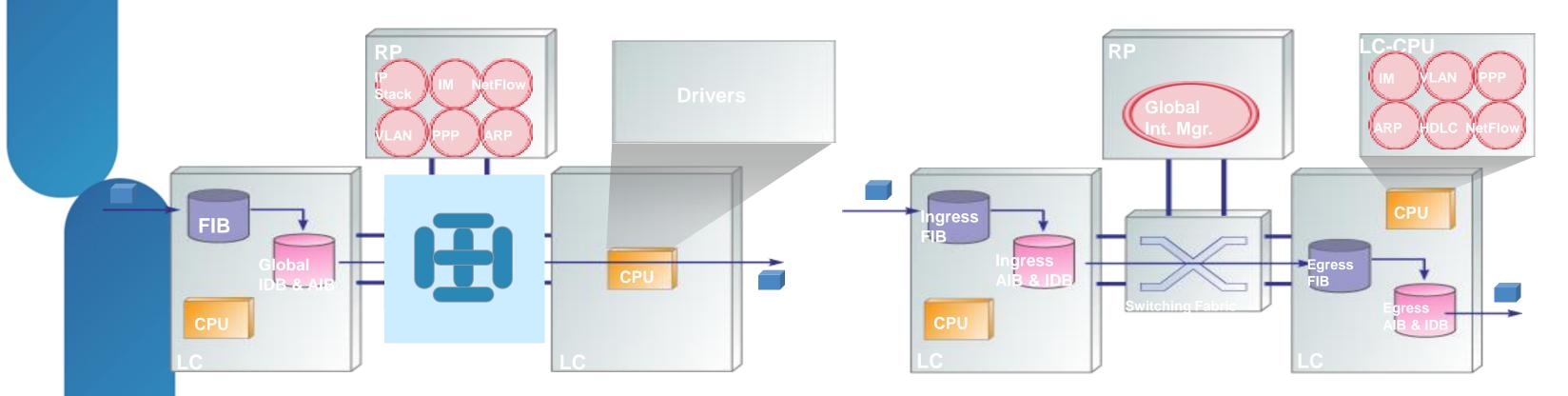- Distributed data management model improves performance and Scale

- Single Consolidated view of the system eases maintenance

IOS XR Distributed Forwarding Infrastructure

Single Stage Forwarding

- Single global Adjacency Information Base (AIB) distributed to all line cards

- Single global Interface Management DB distributed to all line cards

- Only Ingress FIB – forces forwarding features to be run in RP
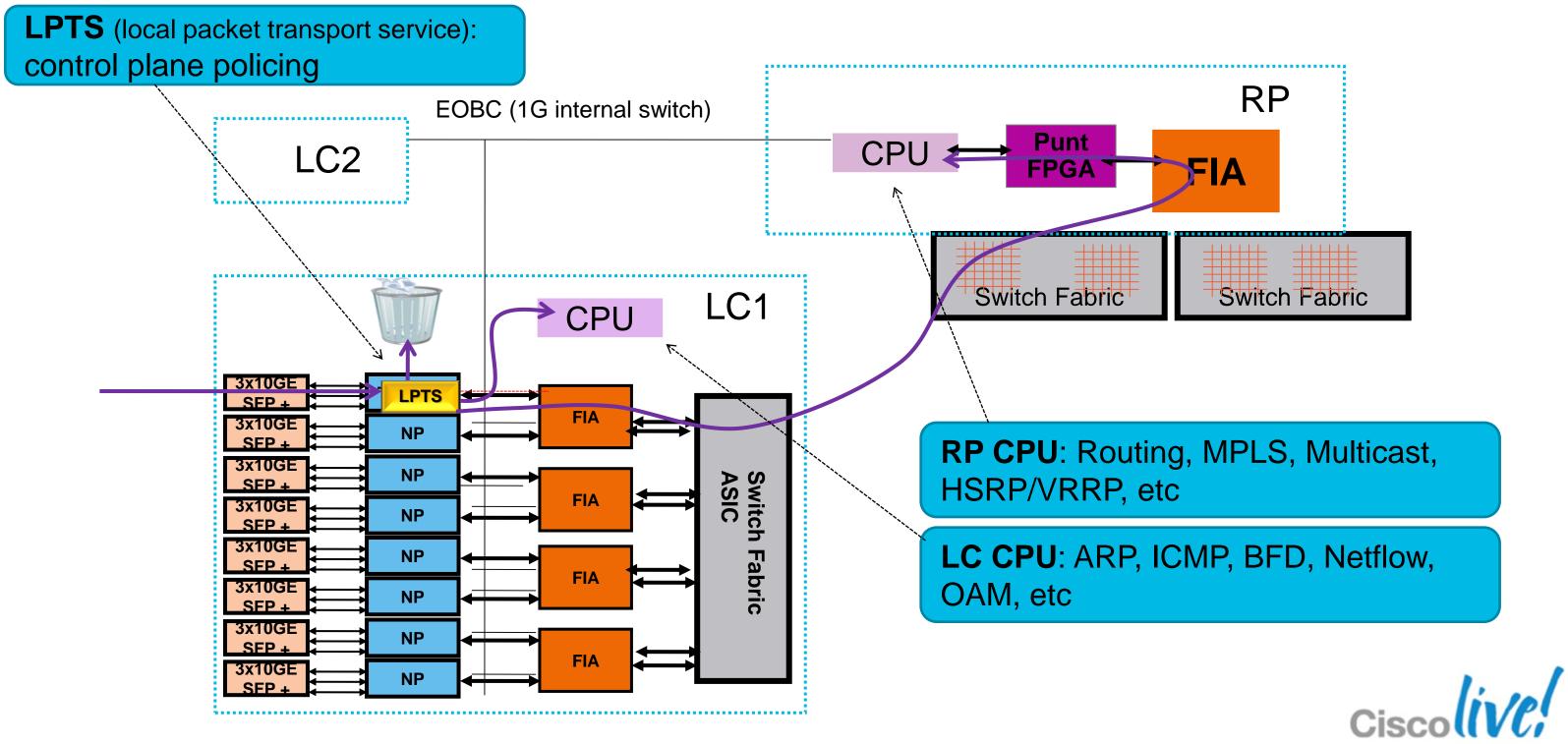
Two Stage Forwarding

- Each line card has independent AIB only for local interfaces

- Each line card has independent Interface DB for local interfaces

- Both Ingress and Egress FIB – allows forwarding features to be independently run in LCs

Cisco Public

ASR 9000 Fully Distributed Control Plane

LPTS (local packet transport service): control plane policing

EOBC (1G internal switch)

RP

LC2

CPU

Punt FPGA

FIA

CPU

LC1

Switch Fabric

Switch Fabric

3x10GE SFP +

LPTS

NP

FIA

3x10GE SFP +

NP

FIA

3x10GE SFP +

NP

3x10GE SFP +

NP

FIA

Switch Fabric ASIC

3x10GE SFP +

NP

3x10GE SFP +

NP

FIA

3x10GE SFP +

NP

3x10GE SFP +

NP

RP CPU: Routing, MPLS, Multicast, HSRP/VRRP, etc

LC CPU: ARP, ICMP, BFD, Netflow, OAM, etc

IOS XR HA Software Design Principles

NSR (GR)

Non-Stop Forwarding

Separate Control and Data Planes

RP/DRP Redundancy   Active/Standby Failover

Process Restartability: Active State Check pointing

All subsystems: Separate Address Spaces
memory faults affect only 1 process, recovery = restart process

Process Restart

- Microkernel includes minimal functionality
- Non-kernel processes can be restarted
- Critical to HA, ISSU, and MDR functions
- Restarting many processes can be tricky
    - Dependent processes may also need to restart

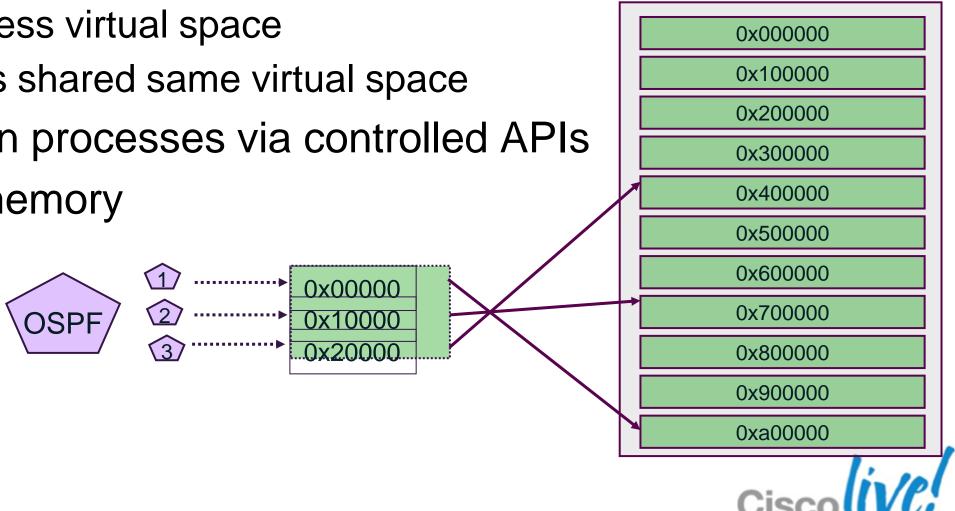# Software Maintenance Updates (SMU)

- Allows for software package installation/removal leveraging on Modularity and Process restart

- Redundant processors are not mandatory (unlike ISSU) and in many cases is non service impacting and may not require reload.

- Mechanism for

  delivery of software features (e.g. Multicast, MPLS)

  delivery of critical bug fixes without the need to wait for next maintenance release

Cisco Public

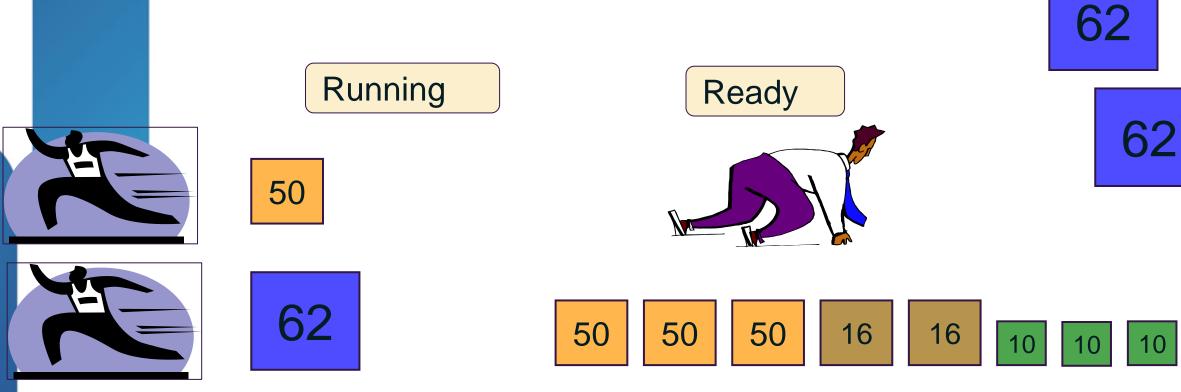# Protected Process Memory Space

- Each process has a virtual memory space
  - Kernel/MMU maps virtual address to physical address (at page level)
  - Threads share the memory space
- One process cannot corrupt another's memory
  - Process can only access virtual space
  - In IOS – all processes shared same virtual space
- Communication between processes via controlled APIs
- Limited use of shared memory

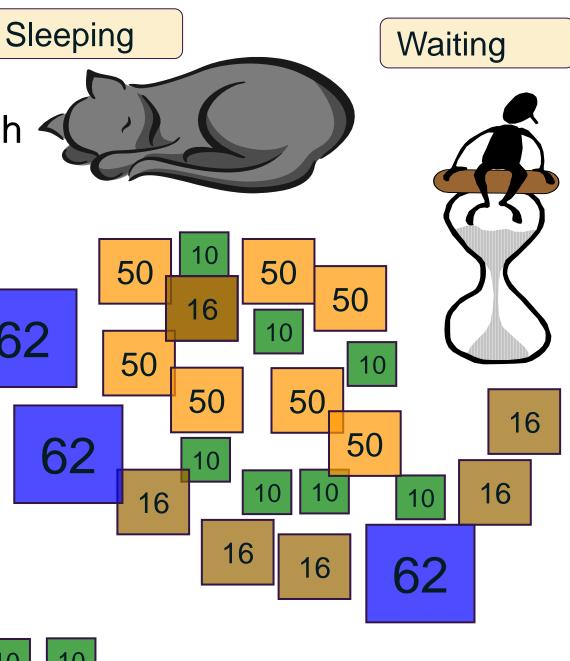| |
|---|
| 0x000000 |
| 0x100000 |
| 0x200000 |
| 0x300000 |
| 0x400000 |
| 0x500000 |
| 0x600000 |
| 0x700000 |
| 0x800000 |
| 0x900000 |
| 0xa00000 |

OSPF

1
2
3

0x00000
0x10000
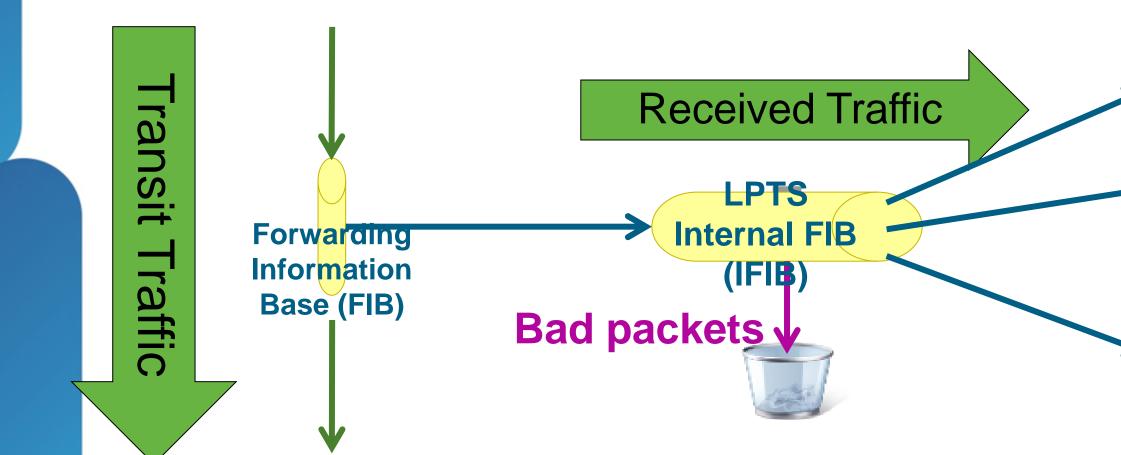0x20000

Preemptive Multitasking

- Default priority is 10
- Higher priority processes can interrupt
    In IOS, must wait for running process to finish
- FIFO within same priority
- Threads run while parent process is running

Sleeping

Waiting

Running

Ready

50

62

50   50   50   16   16   10   10   10

Cisco Public

Local Packet Transport Services (LPTS): Protection of Control Plane

Transit Traffic

Forwarding Information Base (FIB)

Received Traffic

LPTS Internal FIB (IFIB)

Bad packets

Application1 on RP

Application1 on RP

Local Stack on LC

- LPTS enables applications to reside on any or all RPs, DRPs, or LCs

   Active/Standby, Distributed Applications, Local processing
- IFIB forwarding is based on matching control plane flows

   Built in dynamic "firewall" for control plane traffic
- LPTS is transparent and automatic

Cisco Public

Local Packet Transport Service

- There is no longer a single RP

- IOS XR is a fully distributed operating system with applications running in multiple physical locations

- Filters and polices (in hardware) local 'receive' packets and sends them only to the nodes that need them

- Packet rate correlates with trust

- Defaults for 90% of customers, configurability for the rest

- Handles fragments, also checks TTL/hop count

- High Availability for NSR (Non-Stop Routing)

 Cisco Public

Local Packet Transport Service

- LPTS has Hardware policers on line cards to limit traffic sent to local or remote nodes

- LPTS entries in TCAM classifies packets to select a policer to apply

- The policer value can be tuned to 0 (to drop all packet matching classification criteria)

- Polices on protocol (BGP, OSPF, SSH) and flow state (BGP established, BGP configured, and BGP listen)

- Policing done on the LC Hardware ASIC before packets hit RP/LC CPU

- All filters are automatically and dynamically installed by the IOS XR infrastructure

Cisco live!

Summary
Cisco ASR 9000: Next-gen SP+DC Edge Routing

Carrier Class

IOS-XR true modular OS

Superior Multicast replication

Advanced internal system QoS

And more …

Scalable System Architecture
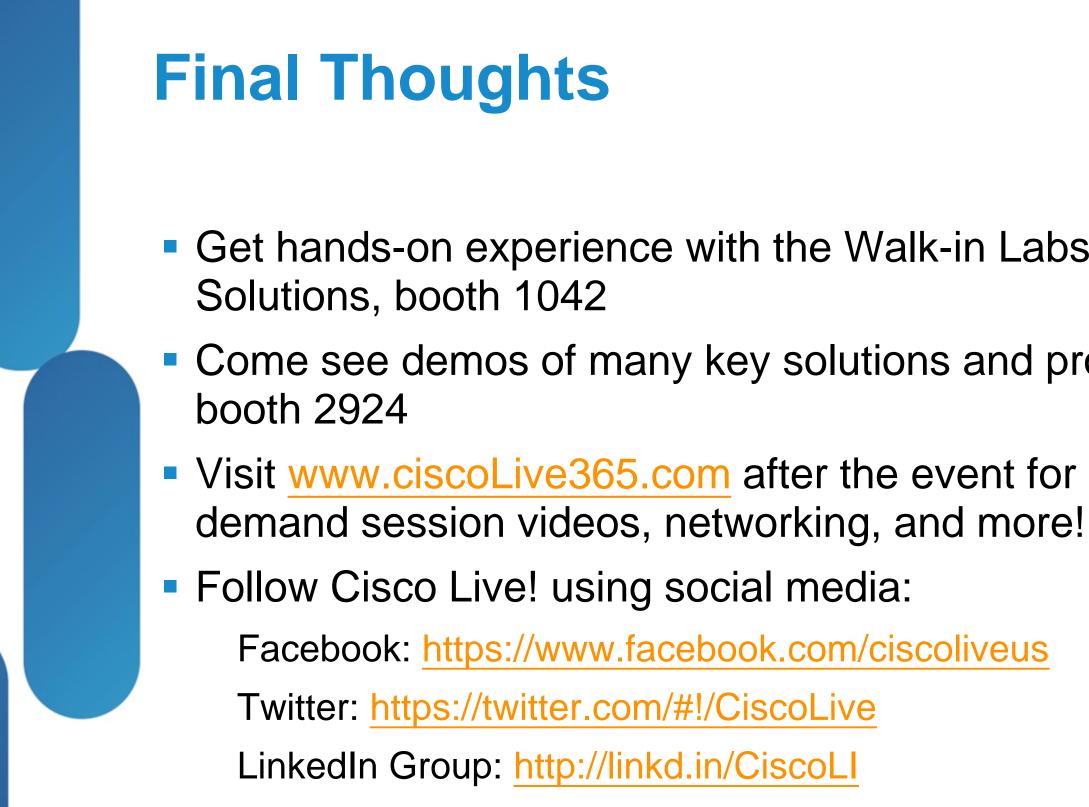
Full HW portfolio: chassis, LC, port

nV – expand the system capacity and physical ports

Fully distributed IOS-XR control plane

IOS-XR 2-stage forwarding

HW MAC learning, prefix independent fast convergence, and more …

Cisco Public

Final Thoughts

- Get hands-on experience with the Walk-in Labs located in World of Solutions, booth 1042

- Come see demos of many key solutions and products in the main Cisco booth 2924

- Visit www.ciscoLive365.com after the event for updated PDFs, on-demand session videos, networking, and more!

- Follow Cisco Live! using social media:

    Facebook: https://www.facebook.com/ciscoliveus

    Twitter: https://twitter.com/#!/CiscoLive

    LinkedIn Group: http://linkd.in/CiscoLI

Cisco Public

Q & A

Complete Your Online Session Evaluation

Give us your feedback and receive a Cisco Live 2013 Polo Shirt!

Complete your Overall Event Survey and 5 Session Evaluations.

- Directly from your mobile device on the Cisco Live Mobile App
- By visiting the Cisco Live Mobile Site www.ciscoliveaustralia.com/mobile
- Visit any Cisco Live Internet Station located throughout the venue

Polo Shirts can be collected in the World of Solutions on Friday 8 March 12:00pm-2:00pm

Cisco live! 365

Don't forget to activate your Cisco Live 365 account for access to all session material, communities, and on-demand and live activities throughout the year. Log into your Cisco Live portal and click the "Enter Cisco Live 365" button.

www.ciscoliveaustralia.com/portal/login.ww

Cisco live!

Cisco Public