cisco Live!

Let's go

#CiscoLiveAPJC

The bridge to possible

Nexus 9000 Architecture

Nemanja Kamenica
Technical Marketing Engineer

BRKDCN-3939

Cisco Webex App

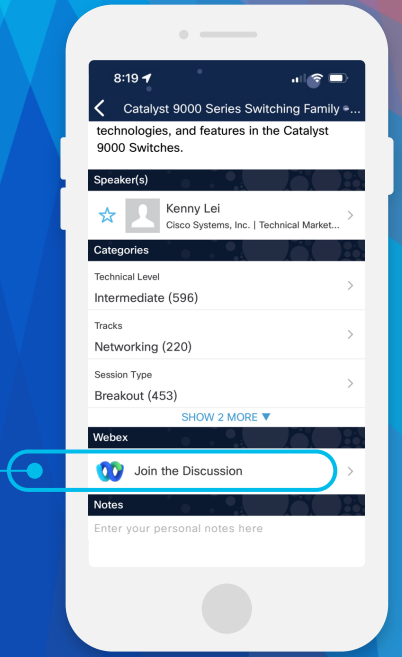
Questions?

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

How

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

Webex spaces will be moderated by the speaker until December 22, 2023.



https://ciscolive.ciscoevents.com/ciscolivebot/#BRKDCN-3939

What This Session Covers

- Latest generation of Nexus 9000 switches with Cloud Scale ASICs
- Nexus 9500 modular switches with Cloud Scale linecards
- Nexus 9300 Cloud Scale top-of-rack (TOR) switches
- System and hardware architecture, key forwarding functions, packet walks

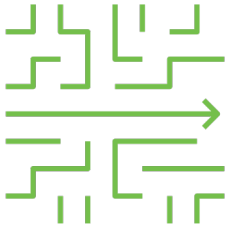
Not covered:

- First generation Nexus 9000 ASIC/platform architectures
- Nexus 9500 merchant-silicon based architectures
- Other Nexus platforms
- *Catalyst* 9000 platform

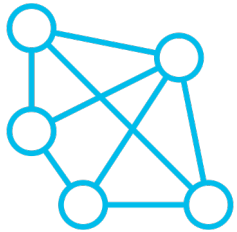
Agenda

- Data Center and Silicon Strategy
- Cloud Scale ASIC Architecture
- Cloud Scale Switching Platforms
- Packet Walks
- Key Takeaways

Cisco Data Center Network Strategy



Operational Simplicity



Sustainable Data Center Networking

Networking for AI/ML

Nexus 9000 Cloud Scale Switching Portfolio

Key Elements of the Data Center Strategy

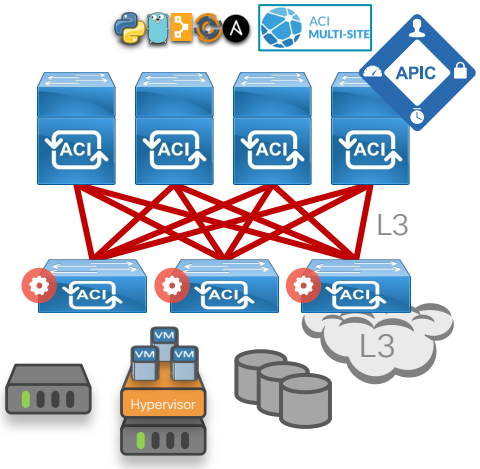
Nexus 9300-FX/FX2/FX3, GX/GX2, H1/H2R and 9408

Premier TOR platforms

Nexus 9500 with X9700-FX and X9700-GX Modules

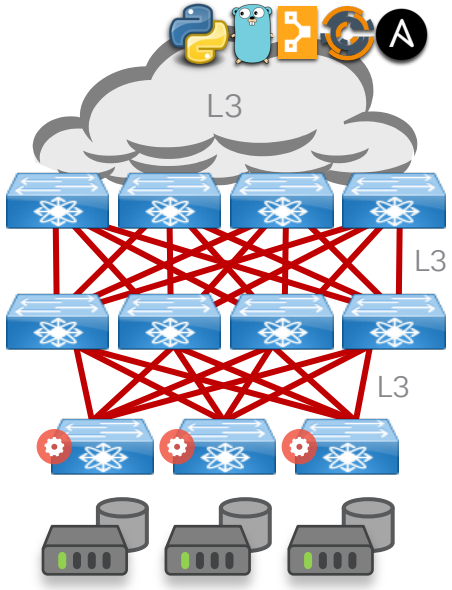
Flagship switching modules for Nexus 9500 modular chassis

Building Data Center Fabrics with Nexus 9000

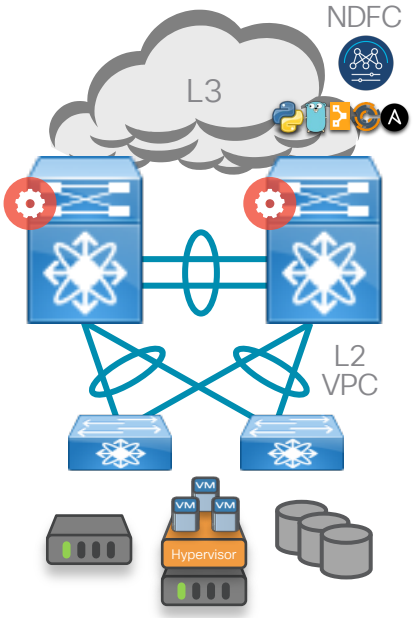


ACI – Turnkey Fabric

Standalone – Programmable Fabric with VXLAN+EVPN

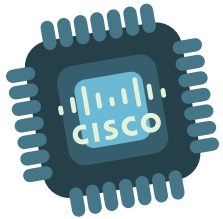


Standalone – Programmable IP Network



Standalone – Traditional Data Center Network

Why Custom Silicon?



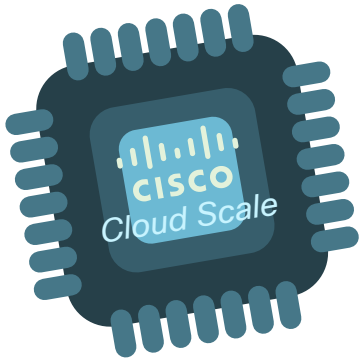
- Cisco competitive advantage – vehicle for differentiating innovations
  - ACI policy model
  - Flexible forwarding tiles
  - Single-pass tunnel encapsulations
  - In-built encryption technologies
  - Intelligent buffers
  - Streaming hardware telemetry
- Tight integration between hardware / software / marketing / sales / support
- Closely aligns hardware designs with software innovations, strategic product direction, competitive differentiators, serviceability

Agenda

- Data Center and Silicon Strategy
- Cloud Scale ASIC Architecture
- Cloud Scale Switching Platforms
- Packet Walks
- Key Takeaways

Cisco Cloud Scale ASIC Family

- **Ultra-high port densities** → Reduces equipment footprint, enables device consolidation, denser fabric designs
- **Multi-speed 100M/1/10/25/40/50/100G/400G** → Flexibility and future proofing
- **Rich forwarding feature-set** → ACI, Segment Routing, single-pass L2/L3 VXLAN routing
- **Flexible forwarding scale** → Single platform, multiple scaling alternatives
- **Intelligent buffering** → Shared egress buffer with dynamic, advanced traffic optimization
- **In-built analytics and telemetry** → Real-time network visibility for capacity planning, security, and debugging



Key Cloud Scale Family Members

LS25600GX2A – 64 x 400G

25.6T chip – 4 slice pairs of 8 x 400G
9300-GX2A TORs; 9408 centralized modular TOR

LS12800GX2B – 32 x 400G

12.8T chip – 2 slice pairs of 8 x 400G
9300-GX2B TOR

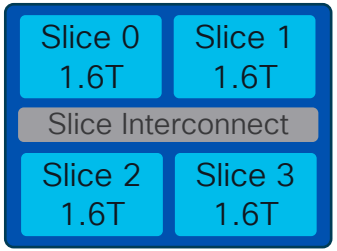
LS12800 H2R – 32 x 400G

12.8T chip – 2 slice pairs of 8 x 400G, 8GB HBM
9300-H2R TOR

LS6400H1 – 16 x 400G

3.6T chip – 2 slices of 8 x 400G
9300-H1 TORs

Key Cloud Scale Family Members



LS6400GX – 16 x 400G

6.4T chip – 4 slices of 4 x 400G
X9700-GX modular linecards; 9300-GX TORs

LS3600FX2 – 36 x 100G

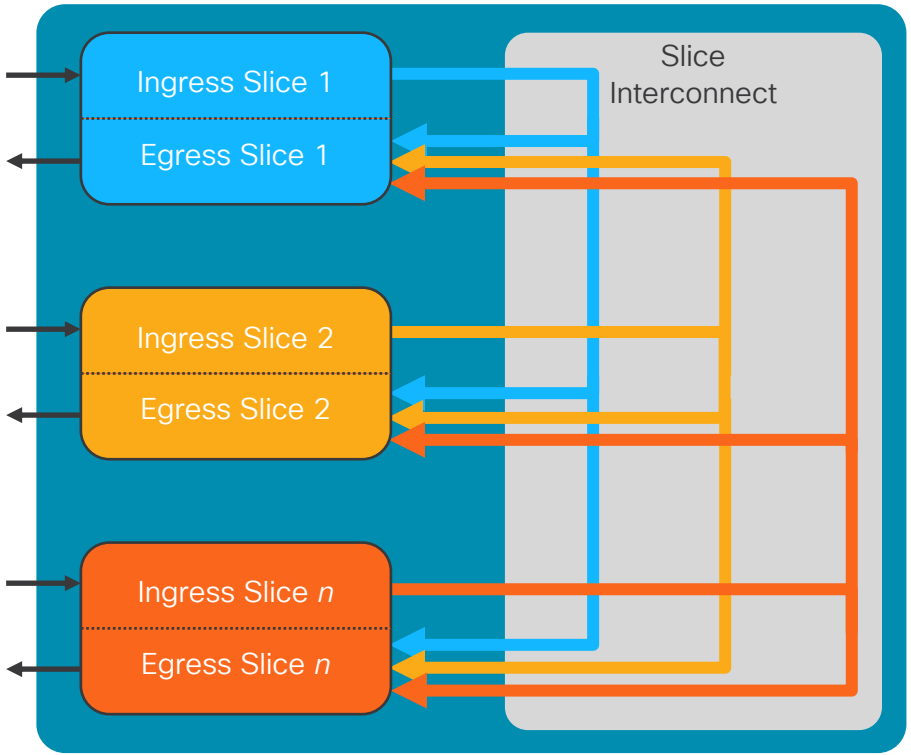
3.6T chip – 2 slices of 18 x 100G with MACSEC
9300-FX2 TORs

LS1800FX3 – 18 x 100G

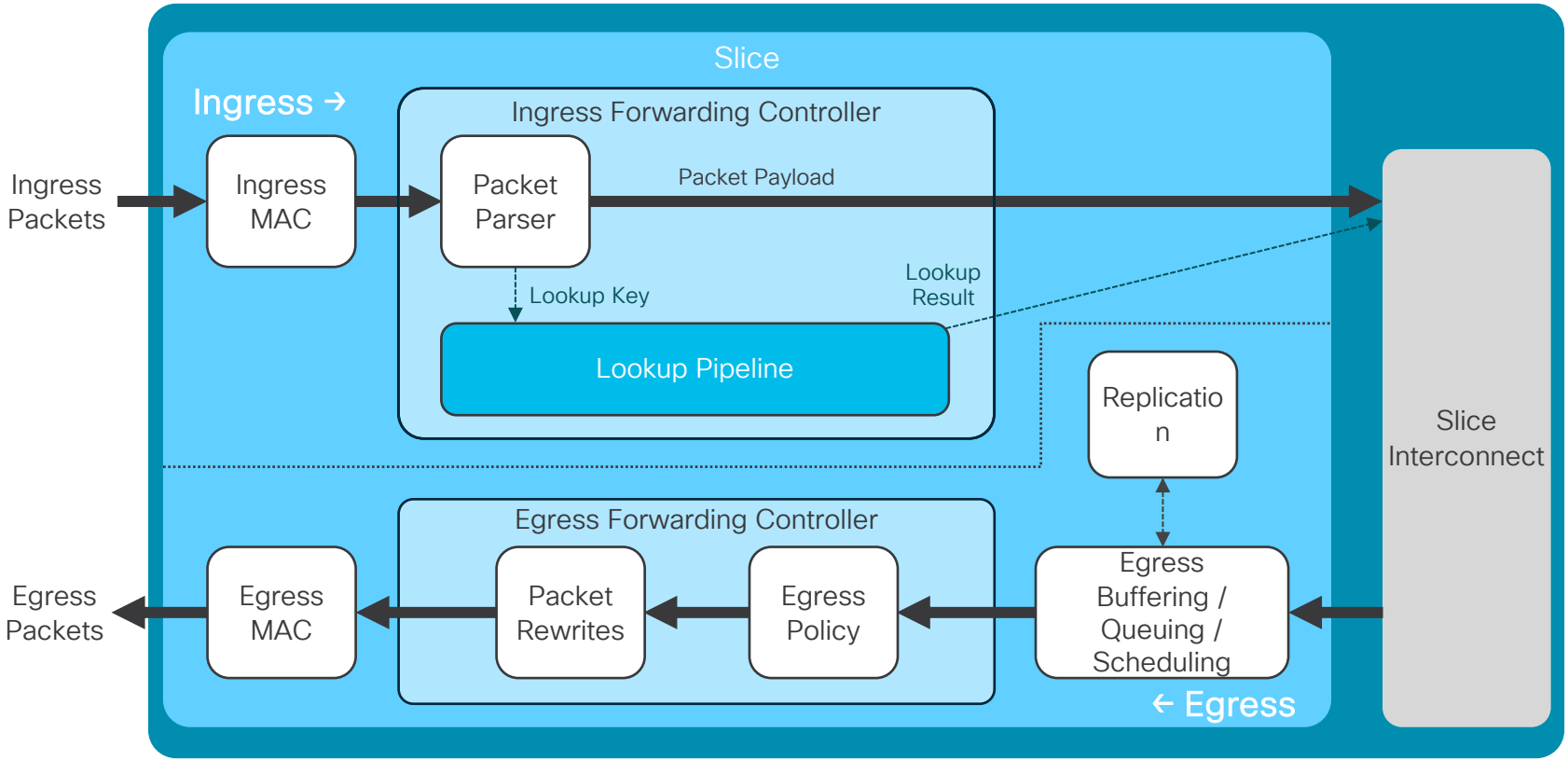
1.8T chip – 1 slice of 18 x 100G with MACSEC
X9700-FX modular linecards; 9300-FX/FX3 TORs

What Is a “Slice”?

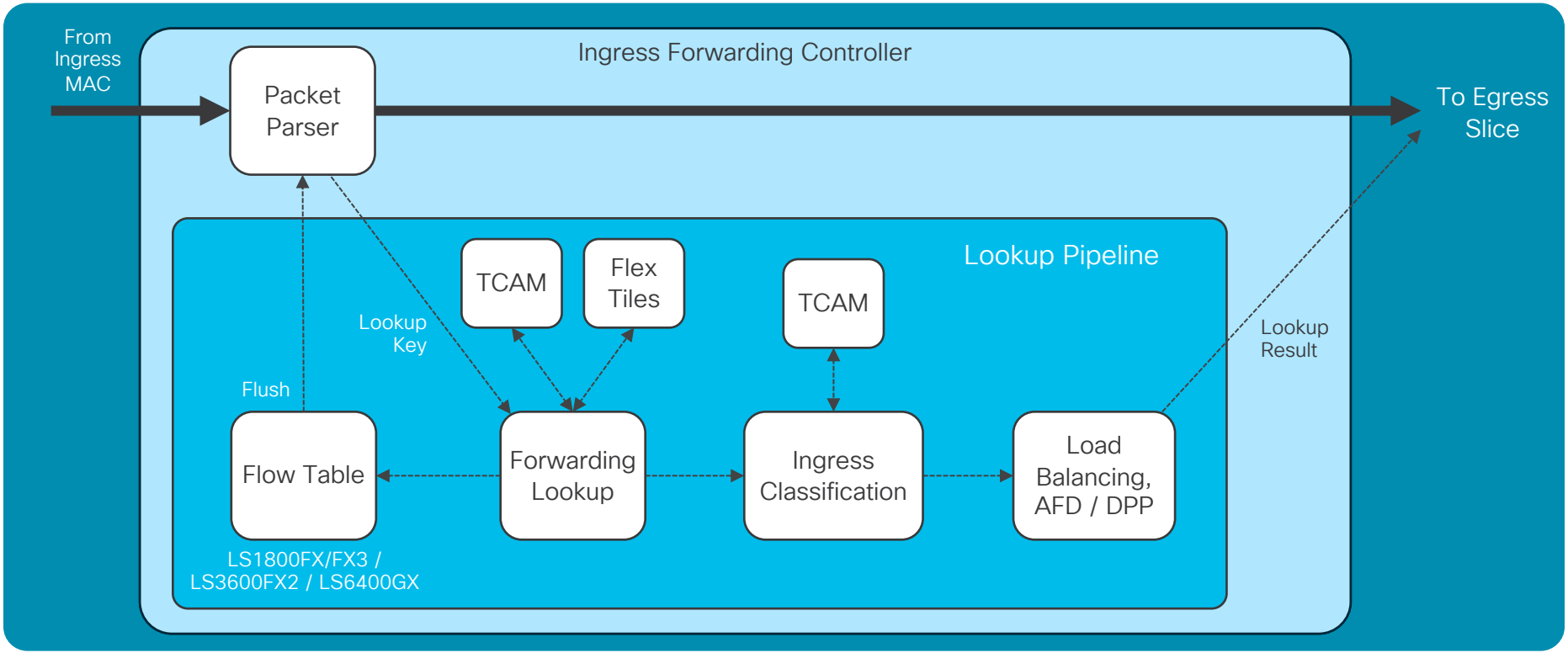
- Self-contained forwarding complex controlling subset of ports on single ASIC
- Separated into Ingress and Egress functions
- Ingress of each slice connected to egress of all slices
- Slice interconnect provides non-blocking any-to-any interconnection between slices



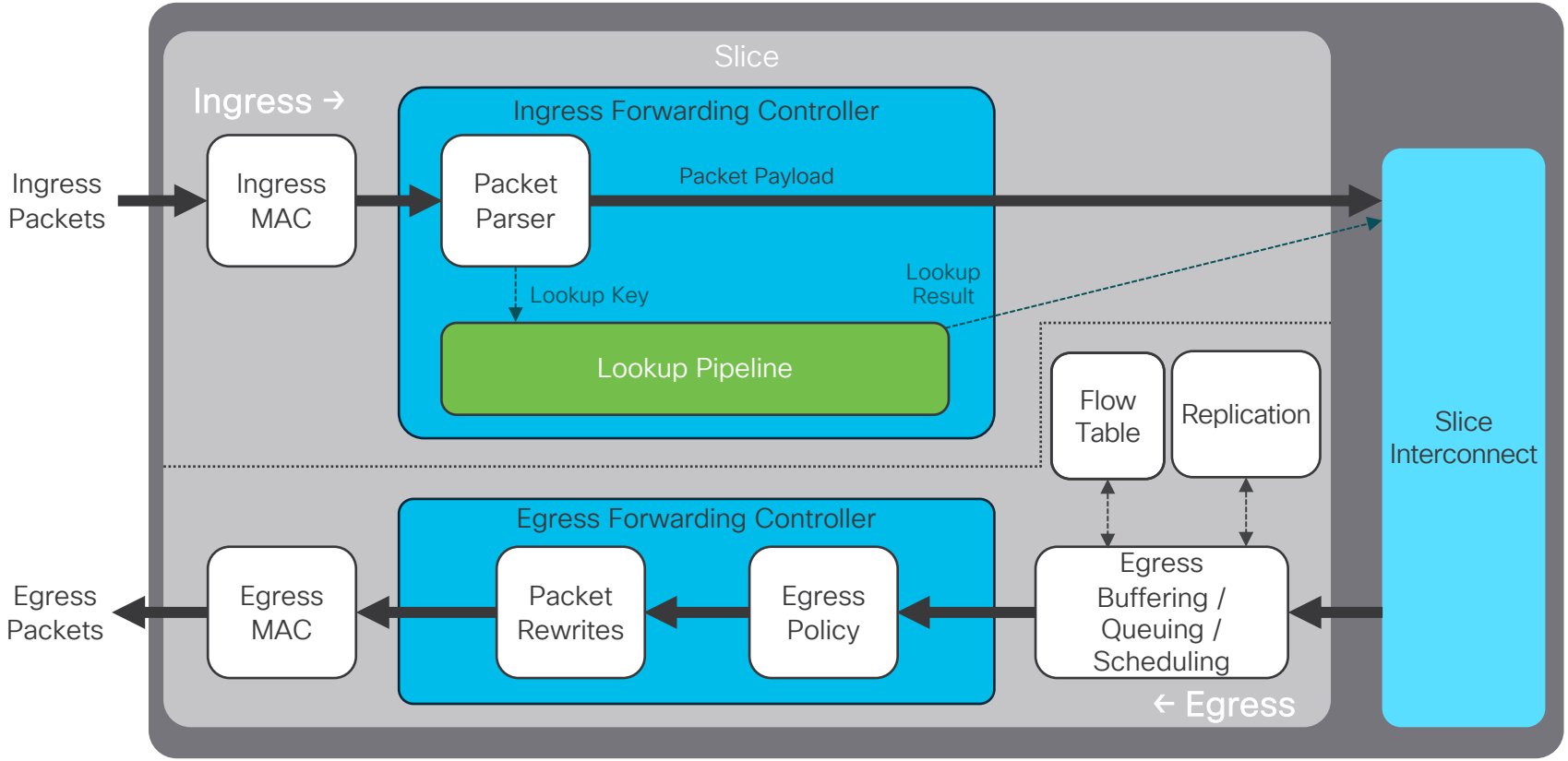
Slice Forwarding Path



Ingress Lookup Pipeline

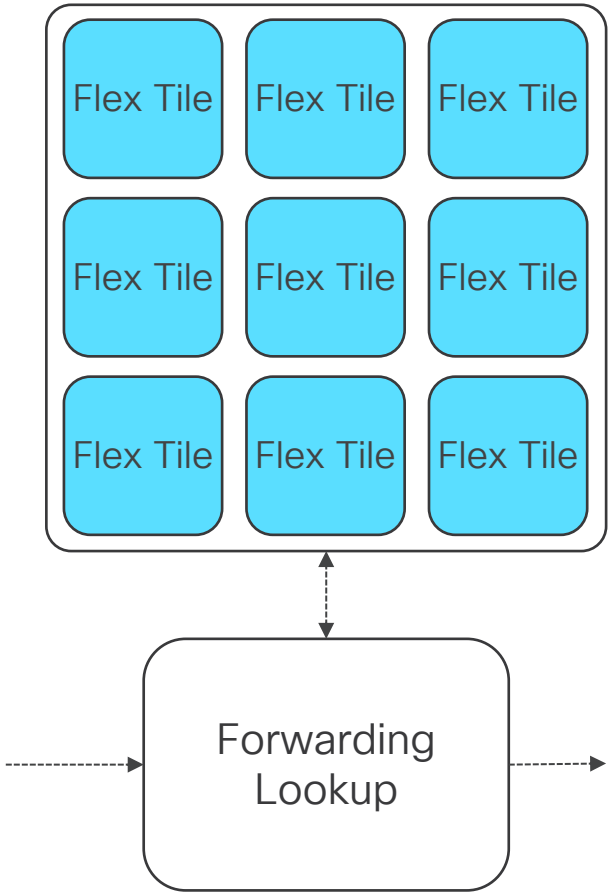


Slice Forwarding Path – 9300-GX2/H



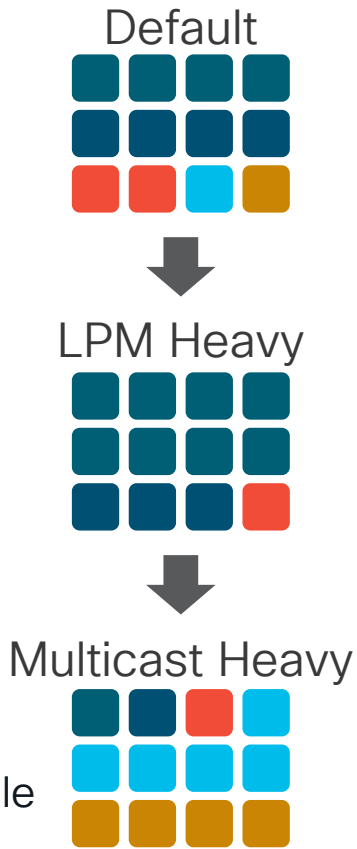
Flexible Forwarding Tiles

- Provide fungible pool of table entries for lookups
- Number of tiles and number of entries in each tile varies between ASICs
- Variety of functions, including:
  - IPv4/IPv6 unicast longest-prefix match (LPM)
  - IPv4/IPv6 unicast host-route table (HRT)
  - IPv4/IPv6 multicast (*,G) and (S,G)
  - MAC address/adjacency tables
  - ECMP tables
  - ACI policy



Flex Tile Routing Templates

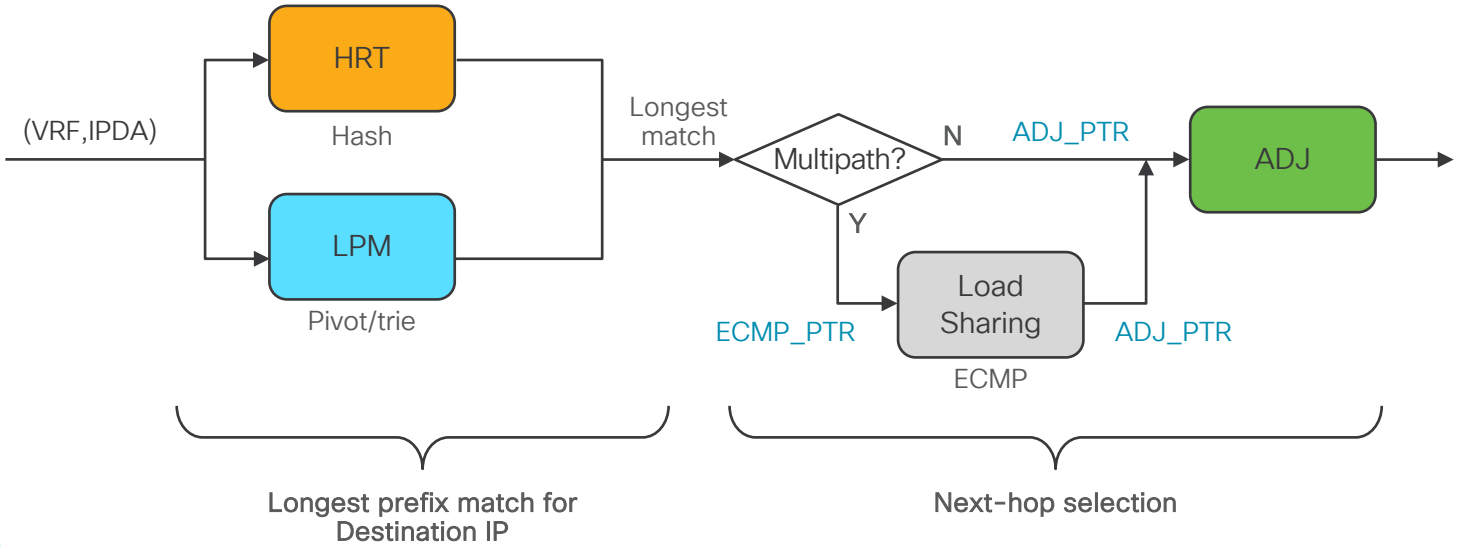
- Configurable forwarding templates determine flex tile functions
  - “system routing template” syntax
- Templates as of NX-OS 10.3(4a):
  - Default
  - Dual-stack host scale*†
  - Dual-stack multicast
  - Internet peering*
  - LPM heavy
  - L2-heavy
  - L3-heavy
  - MPLS heavy*
  - Multicast heavy
  - Multicast extra-heavy
  - Service provider
- Defined at system initialization – reboot required to change profile



* Template does not support IP multicast
† Template not supported on modular Nexus 9500
** Template not supported on TORs

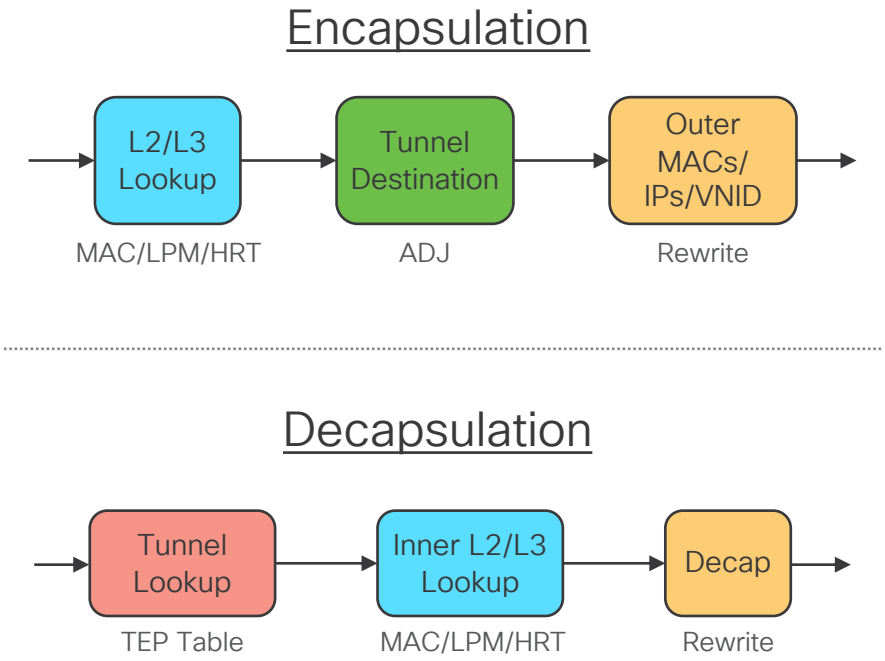
IP Unicast Forwarding

- Hardware lookup in flex tiles based on (VRF, IPDA)
- Longest-match from hash-based exact match (**HRT**) + pivot/trie match (**LPM**)
- Lookup result returns **adjacency** – directly or via **load-sharing** decision (ECMP)



VXLAN Forwarding

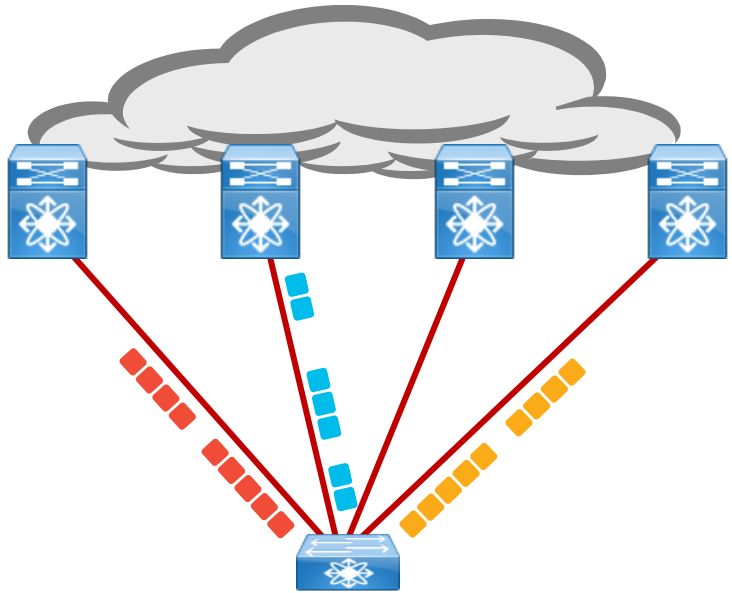
- VXLAN and other tunnel encapsulation/ decapsulation performed in single pass
- Encapsulation
  - L2/L3 lookup drives tunnel destination
  - Rewrite block drives outer header fields (tunnel MACs/IPs/VNID, etc.)
- Decapsulation
  - Outer lookup determines if tunnel is transit or terminated on local TEP
  - Inner lookup determines final output port and rewrites



Load Sharing

Equal-Cost Multipath (ECMP)

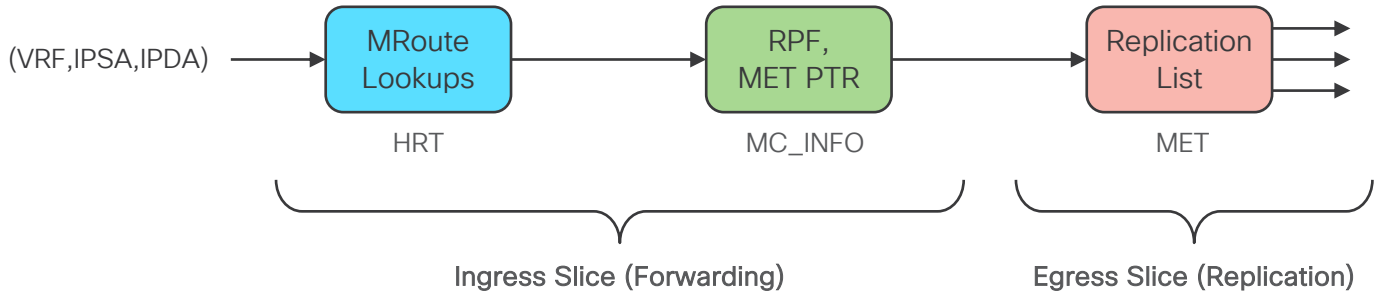
- Static flow-based load-sharing
- Picks ECMP next-hop based on hash of packet fields and universal ID
  - Source / destination IPv4 / IPv6 address (L3)
  - Source / destination TCP / UDP ports (L4)
  - L3 + L4 (default)
  - GRE key field
  - GTP TEID



ECMP

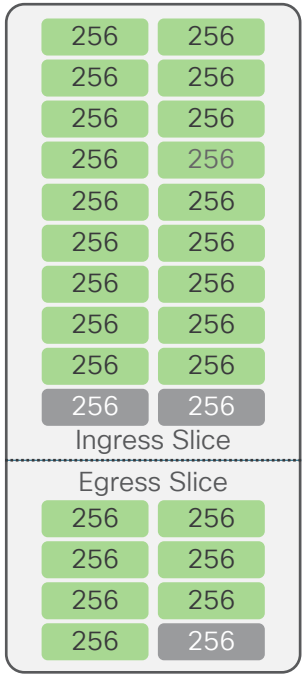
Multicast Forwarding

- Hardware performs multicast lookups in **HRT**
- Additional, secondary table for multicast also provisioned (“MC_INFO”) from flex tiles – **RPF check** and **MET pointer**
- **MET** in egress slice holds local output interface list (OIL)
- Replication is single copy, multiple reads



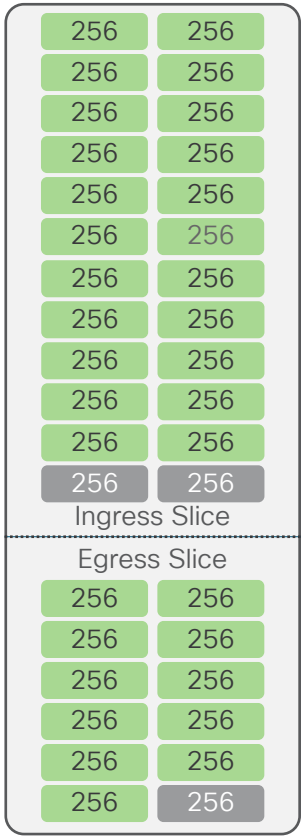
Classification TCAM

- Dedicated TCAM for packet classification
- Capacity varies depending on platform
- Leveraged by variety of features:
  - RACL / VACL / PACL
  - L2/L3 QOS
  - SPAN / SPAN ACL
  - NAT
  - COPP
  - Flow table filter



LS1800FX3 / LS3600FX2 / LS6400GX

5K ingress ACEs / 2K egress ACEs per slice

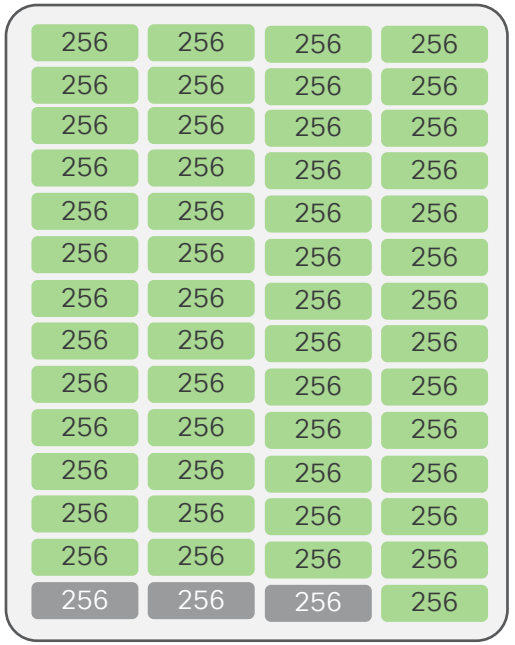


LS12800GX2B / LS26500GX2A

6K ingress ACEs / 3K egress ACEs per slice

Classification TCAM

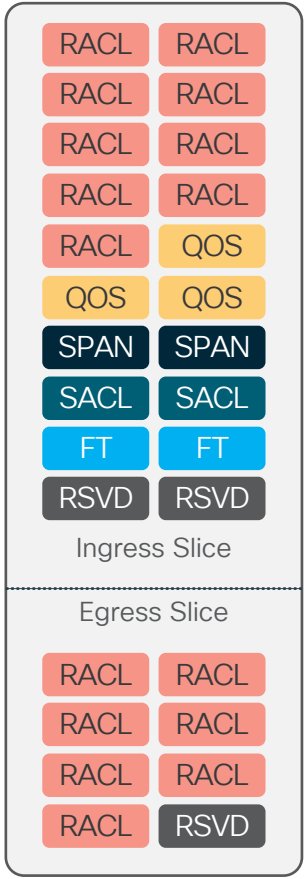
- Dedicated TCAM for packet classification
- Capacity varies depending on platform
- Leveraged by variety of features:
  - RACL / VACL / PACL
  - L2/L3 QOS
  - SPAN / SPAN ACL
  - NAT
  - COPP
  - Flow table filter



LS6400H1/LS12800H2R
14K shared ACEs per slice

TCAM Region Resizing

- Default carving allocates 100% of TCAM and enables:
  - Ingress / Egress RACL
  - Ingress QOS
  - SPAN
  - SPAN ACLs
  - Flow table filter
  - Reserved regions
- Based on features required, user can resize TCAM regions to adjust scale
  - To increase size of a region, some other region must be sized smaller
- Region sizes defined at initialization – changing allocation requires system reboot
  - Configure all regions to desired size (“hardware access-list tcam region”), save configuration, and reload



Cloud Scale Hardware Telemetry

Flow Table (FT)

- Captures full data-plane packet flow information, plus metadata

Flow Table Events (FTE)

- Triggers notifications based on thresholds / criteria met by data-plane packet flows

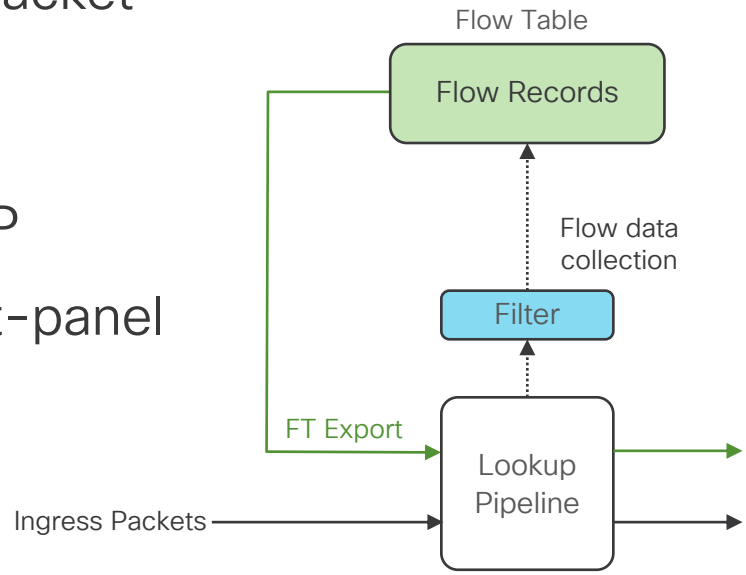
Data-Plane Flow Data

Flow Table

- Collects full flow information plus metadata
  - 5-tuple flow info
  - Interface/queue info
  - Flow start/stop time
  - Packet disposition (drop indicators)
  - Burst measurement
- 32K flow table entries per slice FX/FX2/FX3
- 64K flow table entries per slice on GX/GX2A
- 128K flow table entries per slice on GX2B/H1/H2R
- Direct hardware export
- Leveraged by Nexus Dashboard Insights, Netflow
- FX3 / FX2 / GX / GX2B / GX2A / H2R / H1 platforms support hardware flow table

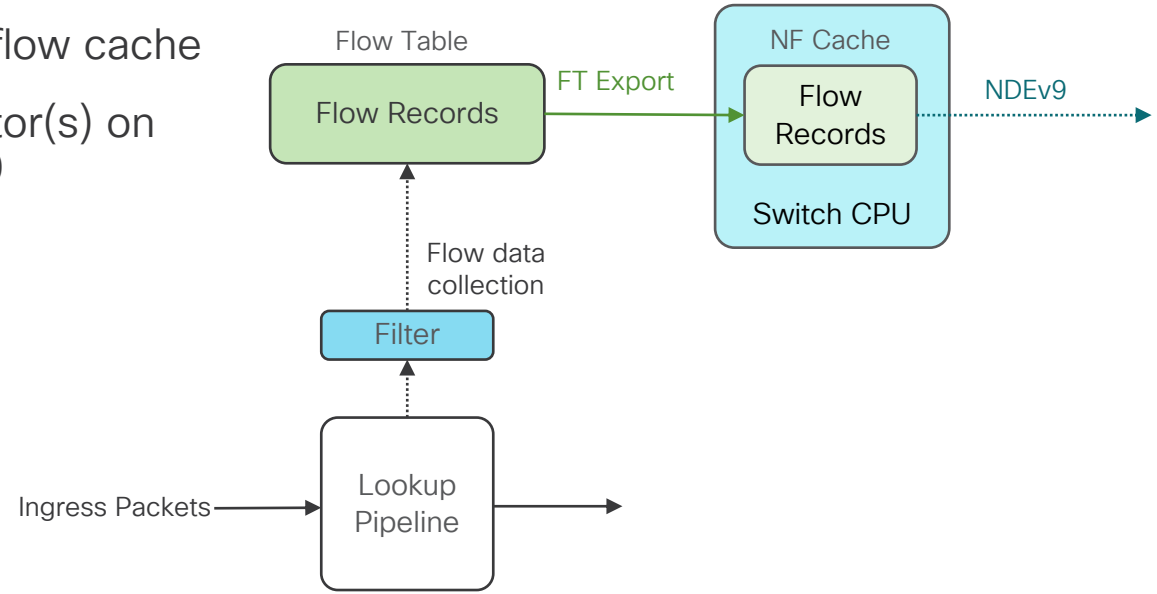
Flow Table Operation – ND Insights

- Determine if collection enabled for packet (filter TCAM)
- If so, install FT record
- Flush records, encapsulate in IP/UDP
- Perform lookup and forward on front-panel port



Flow Table Operation – Netflow

1. Install FT records as usual
2. Flush records to switch CPU
3. CPU builds traditional Netflow cache
4. NDEv9 exported to collector(s) on front-panel port or mgmt0

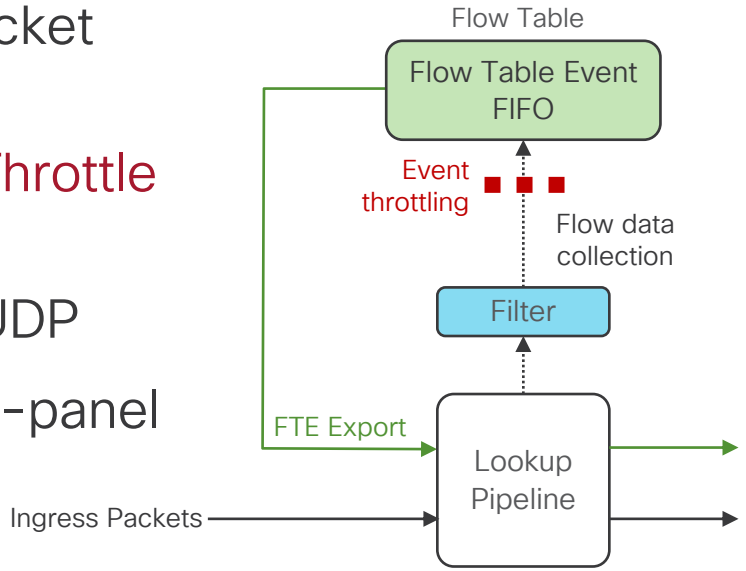


Flow Table Events

- Triggers notifications based on criteria / thresholds met by data-plane packet flows
- Collects full flow information plus metadata
  - 5-tuple flow info with timestamp
  - Interface/queue info
  - Buffer drop indication
  - Forwarding drop, ACL drop, policer drop indication
  - Latency/burst threshold exceeded indication
- Direct hardware export, with flow-level and global throttling
- FX3 / FX2 / GX / GX2B / GX2A / H2R / H1 platforms support hardware flow table events

Flow Table Events Operation

- Determine if event(s) enabled for packet (filter TCAM)
- If so, collect flow data in FTE FIFO; Throttle excess events
- Flush and encapsulate record in IP/UDP
- Perform lookup and forward on front-panel port

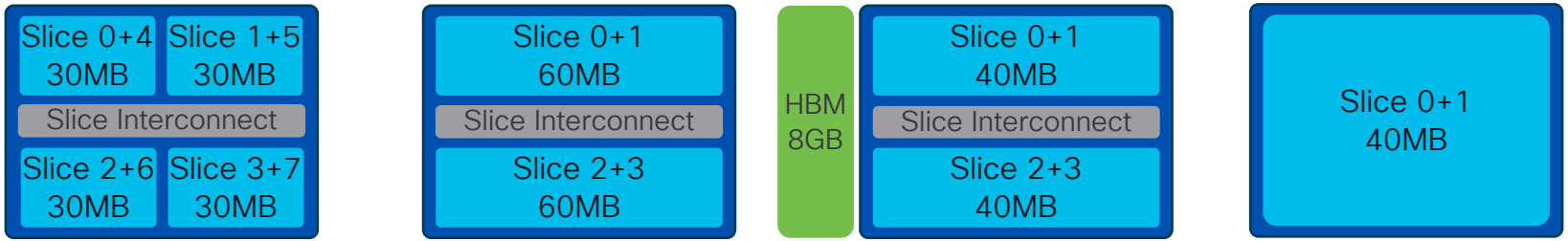


Hardware Telemetry Platform Support

Platform	ASIC	FT	FTE
9300-H2R	LS12800H2R	✓	✓
9300-H1	LS6400H1	✓	✓
9300-GX2A	LS25600GX2A	✓	✓
9300-GX2B	LS12800GX2B	✓	✓
9300-GX / X9700-GX	LS6400GX	✓	✓
9300-FX2	LS3600FX2	✓	✓
9300-FX3	LS1800FX3	✓	✓
9300-FX / X9700-FX	LS1800FX	✓	✓

Buffering

- Cloud Scale platforms implement shared-memory egress buffered architecture
- Slices share pool of buffer – ports on a slice pairs can use that buffer
- Dynamic Buffer Protection adjusts max thresholds based on class and buffer occupancy
- Intelligent buffer options maximize buffer efficiency



LS25600GX2A
30MB/slice pair
(120MB total)

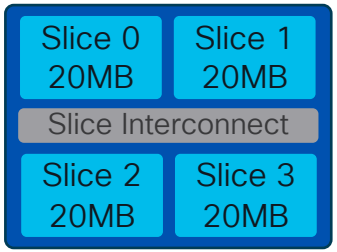
L12800GX2B
60MB/slice pair
(120MB total)

LS12800 H2R
60MB/slice pair
(120MB total)

LS6400H1
40MB/slice pair
(40MB total)

Buffering

- Cloud Scale platforms implement shared-memory egress buffered architecture
- Each ASIC slice has dedicated buffer – only ports on that slice can use that buffer
- Dynamic Buffer Protection adjusts max thresholds based on class and buffer occupancy
- Intelligent buffer options maximize buffer efficiency



LS6400GX
20MB/slice
(80MB total)

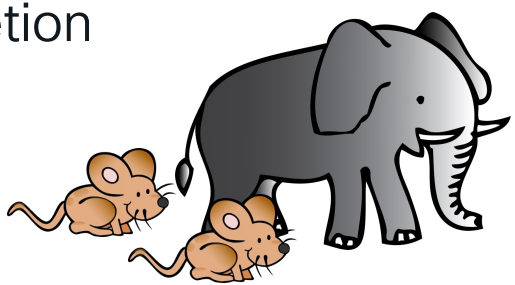
LS3600FX2
20MB/slice
(40MB total)

LS1800FX3
40MB/slice
(40MB total)

Intelligent Buffering

Innovative Buffer Management for Cloud Scale switches

- **Dynamic Buffer Protection** (DBP) – Controls buffer allocation for congested queues in shared-memory architecture
- **Approximate Fair Drop** (AFD) – Maintains buffer headroom per queue to maximize burst absorption
- **Dynamic Packet Prioritization** (DPP) – Prioritizes short-lived flows to expedite flow setup and completion



Images courtesy of:
https://www.clker.com/clipart-206333.html
https://www.clker.com/clipart-catroon-mouse.html

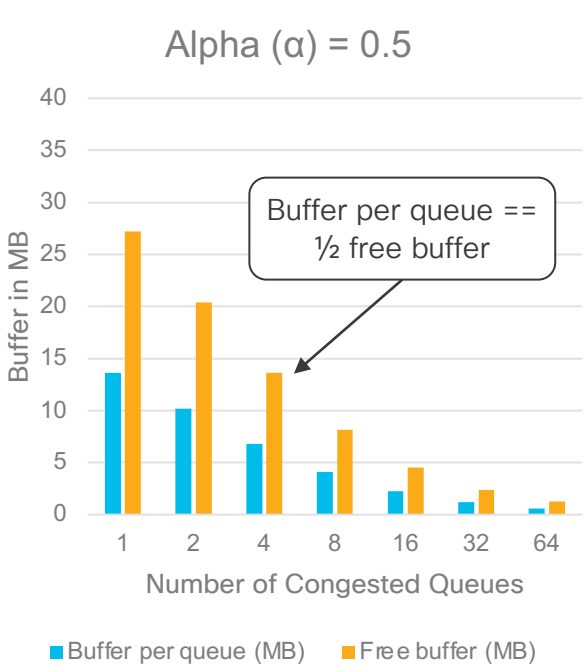
Miercom Report: Speeding Applications in Data Center Networks
http://miercom.com/cisco-systems-speeding-applications-in-data-center-networks/

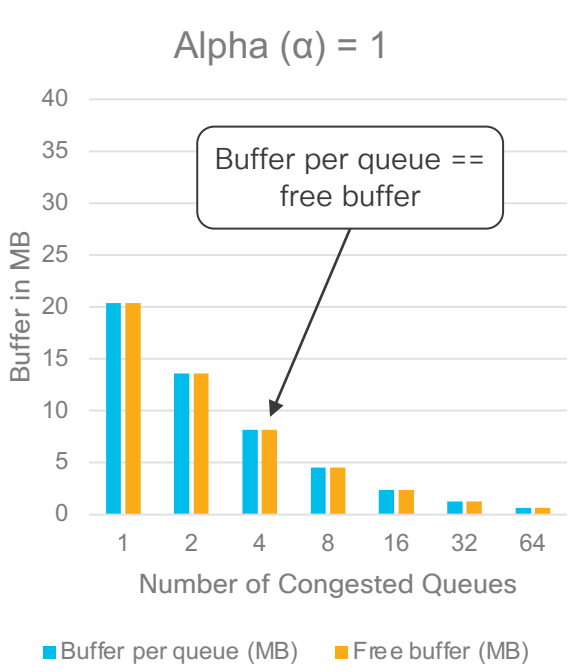
Dynamic Buffer Protection (DBP)

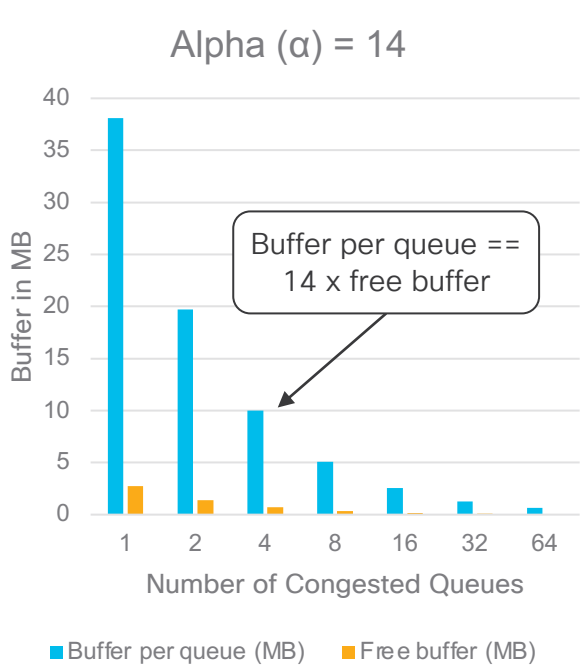
- Prevents any output queue from consuming more than its fair share of buffer in shared-memory architecture
- Defines dynamic max threshold for each queue
  - If queue length exceeds threshold, packet is discarded
  - Otherwise, packet is admitted to queue and scheduled for transmission
- Threshold calculated by multiplying free memory by configurable, per-queue **Alpha** (α) value (weight)
  - Alpha controls how aggressively DBP maintains free buffer pages during congestion events

Alpha Parameter Examples

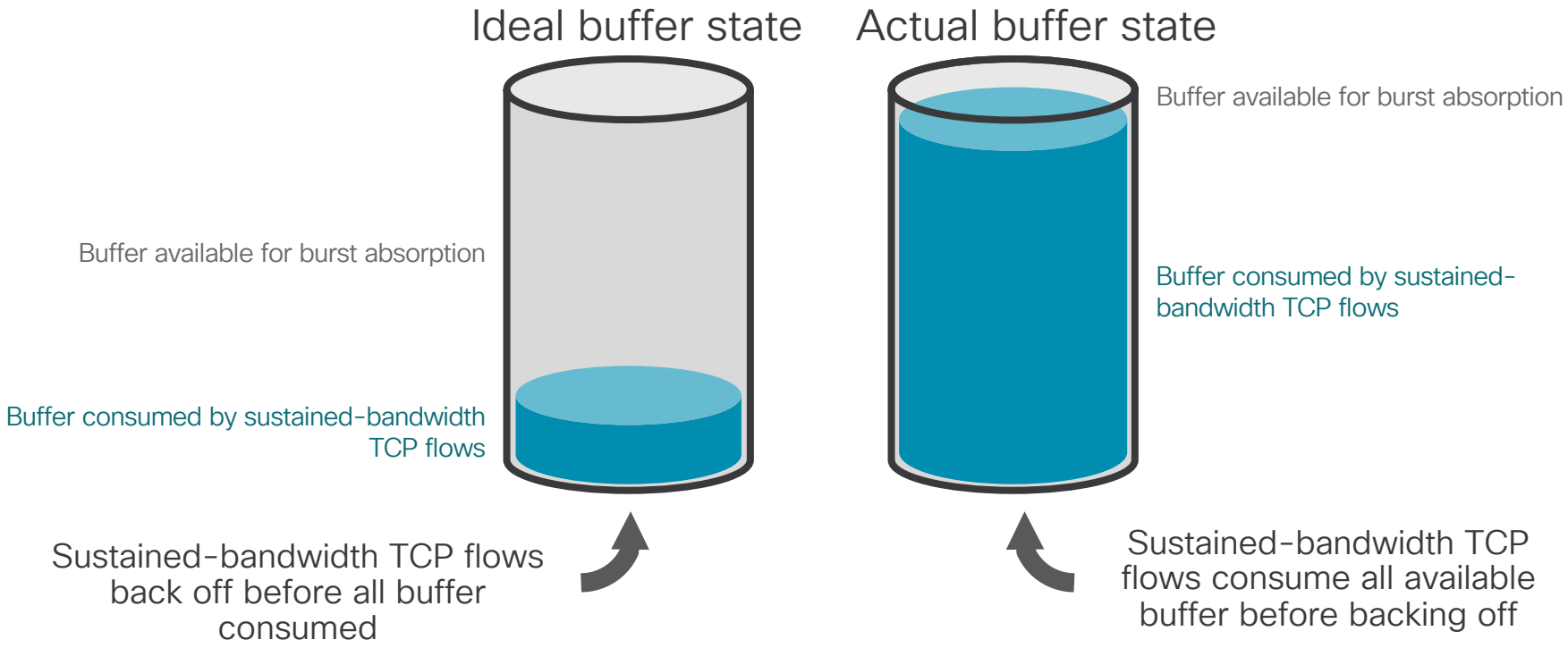
Default Alpha on Cloud Scale switches





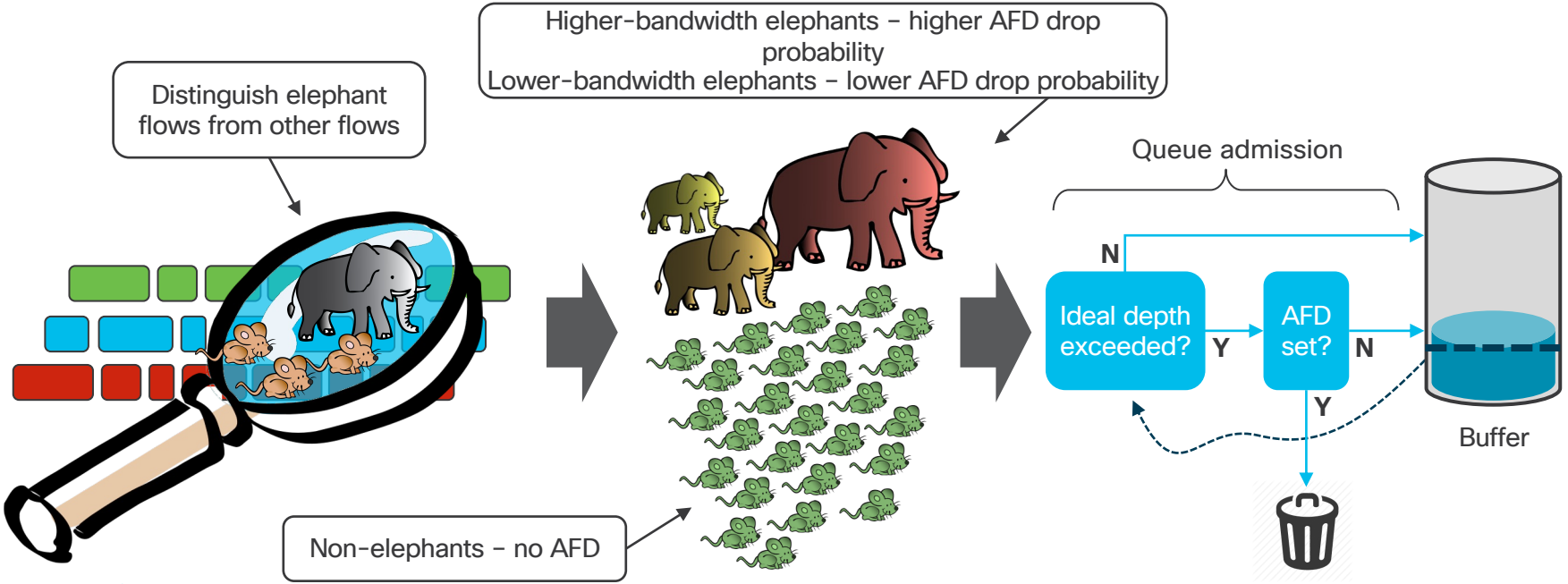


Buffering – Ideal versus Reality



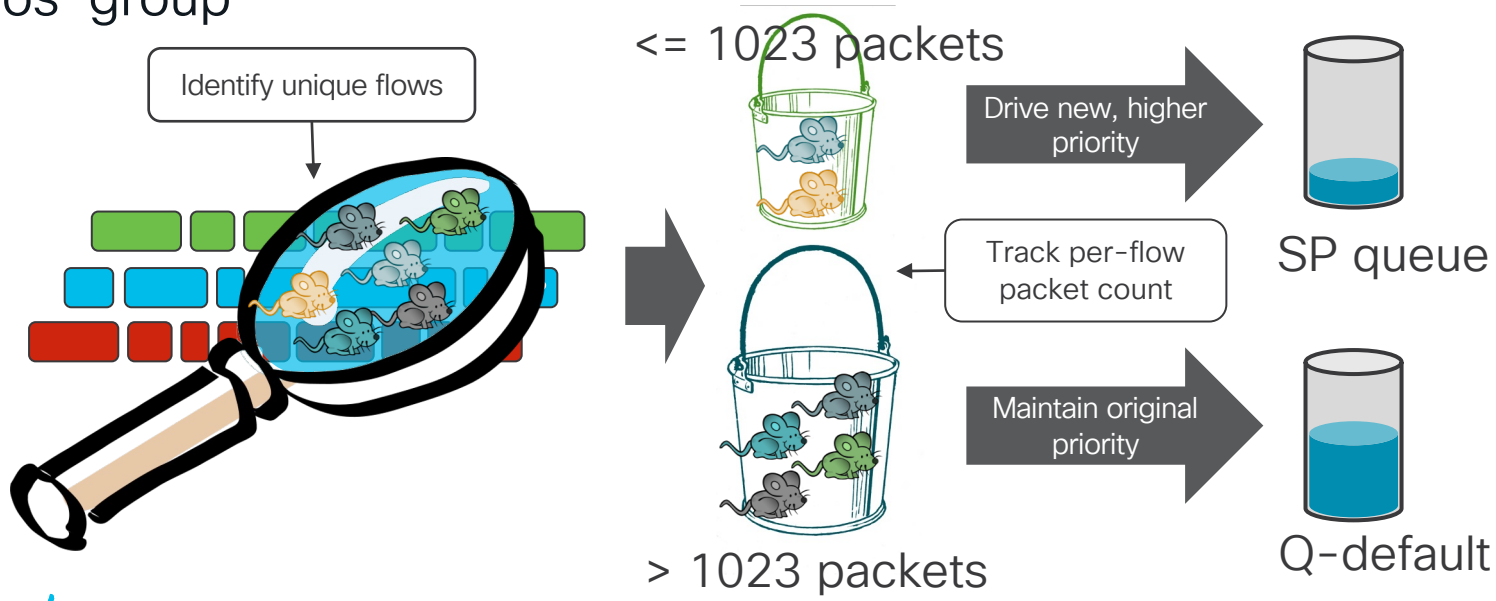
Approximate Fair Drop (AFD)

Maintain throughput while minimizing buffer consumption by elephant flows – **keep buffer state as close to the ideal as possible**

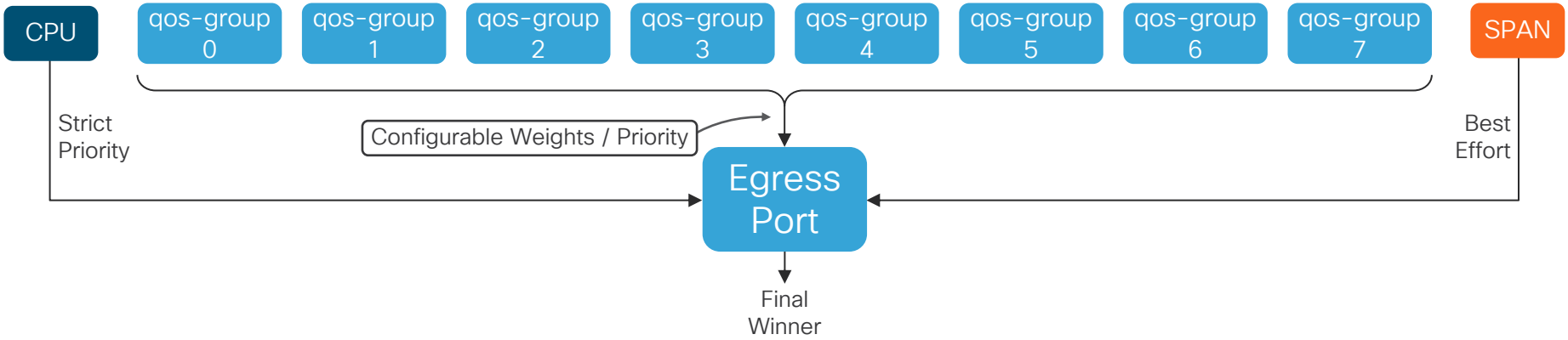


Dynamic Packet Prioritization (DPP)

- Prioritize initial packets of new / short-lived flows
- Up to first 1023 packets of each flow assigned to higher-priority qos-group



Queuing and Scheduling

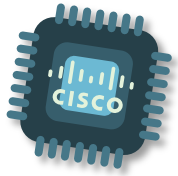


- 8 qos-groups per output port – shared by unicast and multicast traffic
- Egress queuing policy defines priority and weights
- Dedicated classes for CPU traffic and SPAN traffic

Agenda

- Data Center and Silicon Strategy
- Cloud Scale ASIC Architecture
- Cloud Scale Switching Platforms
- Packet Walks
- Key Takeaways

Cloud Scale Platforms



Nexus 9300-FX/FX2/FX3, GX/GX2, H1/H2R and 9408

- Premier TOR platforms
- Full Cloud Scale functionality
- ACI leaf / standalone leaf or spine
- FX2 option with key enhancements using LS3600FX2 silicon
- GX option with 400G and SRv6
- GX2 high density 400G
- H1 key enhancement
- H2 deep buffer and advanced timing

Nexus 9500 with X9700-FX and X9700-GX Modules

- Switching modules for Nexus 9500 modular chassis
- Full Cloud Scale functionality
- ACI spine / standalone aggregation or spine
- FX option with MACSEC using LS1800FX silicon
- GX option with MACSEC

Nexus 9300-FX3 Cloud Scale TOR Switches

48-port 10/25G SFP28 + 6-port 100G QSFP28
N9K-C93180YC-FX3 – LS1800FX3-based
ACI: 5.1(3)
NX-OS: 9.3(7)

48-port 1/10GBASE-T + 6-port 100G QSFP28
N9K-C93108TC-FX3P – LS1800FX3-based
ACI: 5.1(3)
NX-OS: 9.3(7)

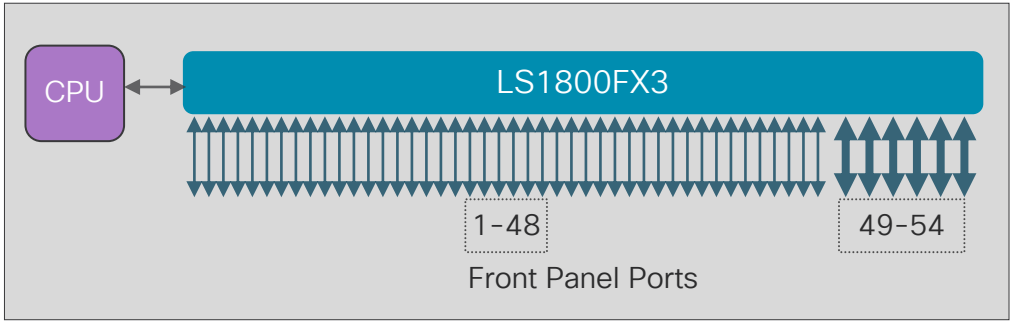
48-port 10M/100M/1GBASE-T + 4-port 10G/25G + 2-port 100G QSFP28
N9K-C9348GC-FX3 – LS1800FX3-based
ACI: 6.1(1)
NX-OS: 10.4(1)

40-port 10/100M/1GBASE-T + 8-port 10/100M Half-Duplex + 4-port 10G/25G + 2-port 100G QSFP28
N9K-C9348GC-FX3PH – LS1800FX3-based
ACI: 6.1(1)
NX-OS: 10.4(1)

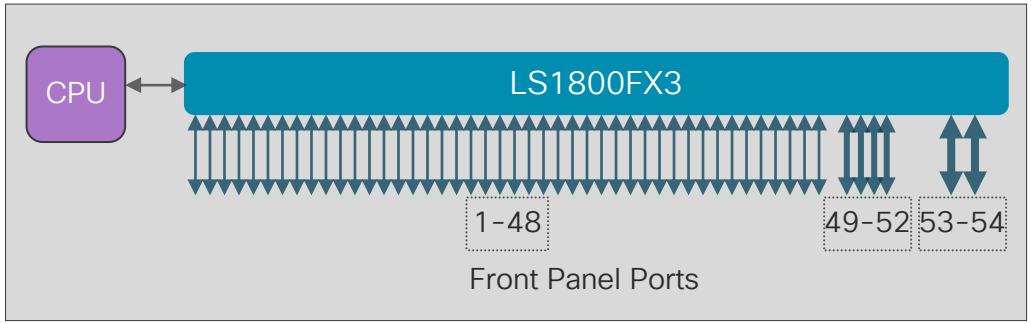
Key Features

- Dual capability – ACI and NX-OS mode
- Flexible port configurations – 100M/1/10/25/40/50/100G
- **Flow Table** for ND Insights, Netflow
- **MACSEC on all ports**
- Smart buffer capability (AFD / DPP)
- Telecom PTP and SyncE - N93180YC-FX3

Nexus 9300-FX3 Switch Architectures



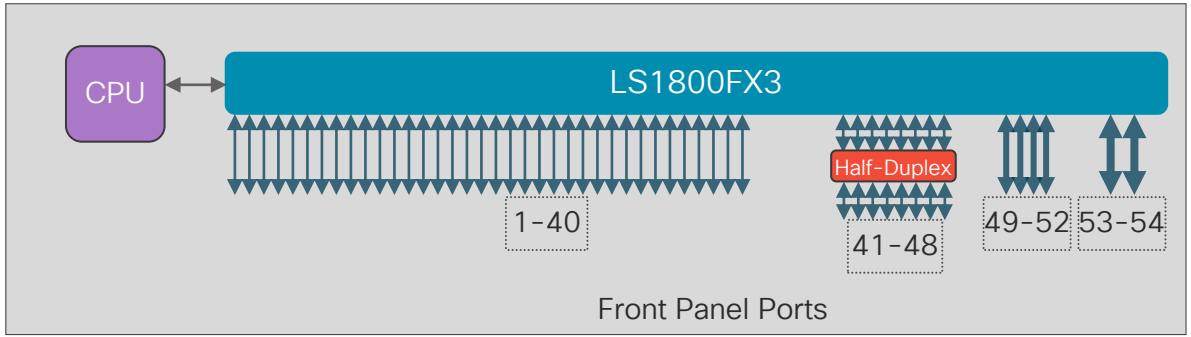
C93180YC-FX3 (10/25G + 100G) /
C93108TC-FX3P (10G + 100G)



C9348GC-FX3 (10M/100M/1G + 10/25G + 100G)

Slice 0

Nexus 9300-FX3 Switch Architectures



C9348GC-FX3H (10M/100M/1G + 10/25G + 100G)

Slice 0

Nexus 9300-FX2 Cloud Scale TOR Switches

36-port 100G QSFP28
N9K-C9336C-FX2 – LS3600FX2-based
ACI: 3.1(2)
NX-OS: 7.0(3)I7(3)

48-port 10/25G SFP28 + 12-port 100G QSFP28
N9K-C93240YC-FX2 – LS3600FX2-based
ACI: 4.0(1)
NX-OS: 7.0(3)I7(3)

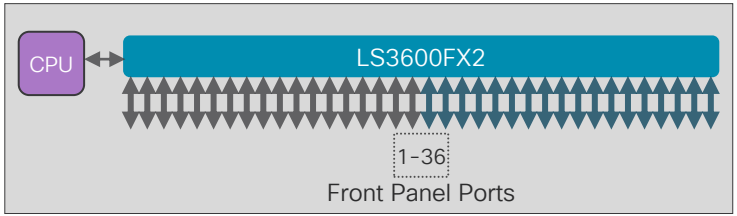
96-port 10/25G SFP28 + 12-port 100G QSFP28
N9K-C93360YC-FX2 – LS3600FX2-based
ACI: 4.1(2)
NX-OS: 9.3(1)

96-port 1/10GBASE-T + 12-port 100G QSFP28
N9K-C93216TC-FX2 – LS3600FX2-based
ACI: 4.1(2)
NX-OS: 9.3(1)

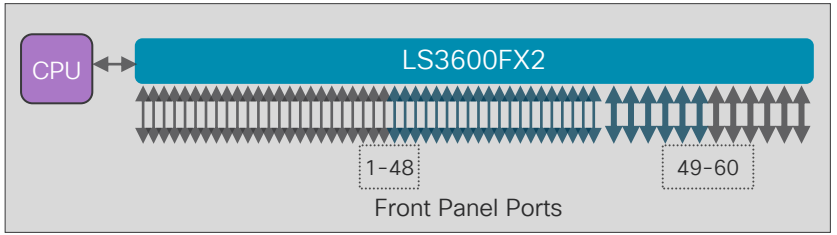
Key Features

- Dual capability – ACI and NX-OS mode
- Versatile standalone 100G switch (9336C)
- High-performance 100G ACI leaf switch (9336C)
- 100G/50G/40G/10G with breakout capability
- 2RU copper/fiber options for high density racks
- **Flow Table** for Network Insights, Netflow
- **MACSEC on all ports**
- Smart buffer capability (AFD / DPP)

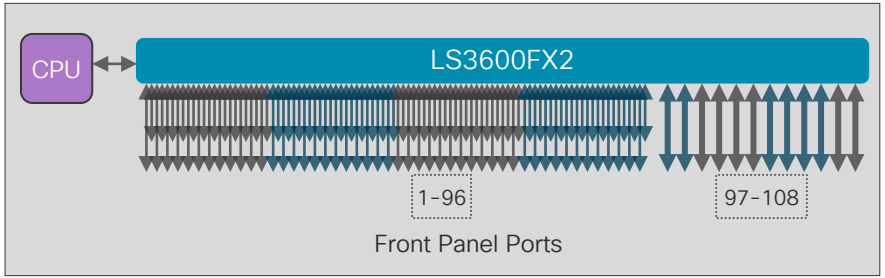
Nexus 9300-FX2 Switch Architecture



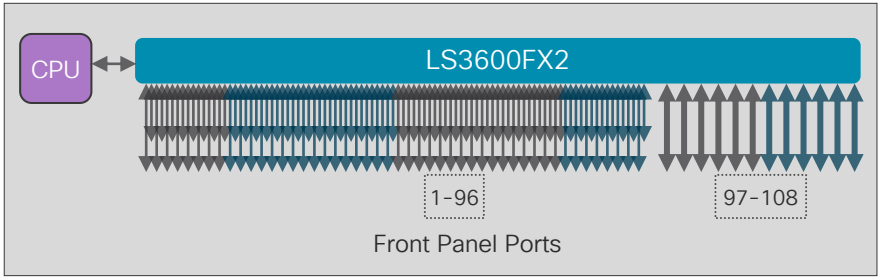
C9336C-FX2 (100G)



C93240YC-FX2 (10/25G + 100G)



N9K-C93360YC-FX2 (10/25G + 100G)



N9K-C93216TC-FX2 (1/10G + 100G)

Nexus 9300-GX Cloud Scale TOR Switches

16-port 400G QSFP-DD
N9K-C9316D-GX – LS6400GX-based
ACI: 4.2(2)
NX-OS: 9.3(3)

28-port 100G QSFP28 +
8-port 400G QSFP-DD
N9K-C93600CD-GX – LS6400GX-based
ACI: 4.2(2)
NX-OS: 9.3(3)

64-port 100G QSFP28
N9K-C9364C-GX – LS6400GX-based
ACI: 4.2(3l)
NX-OS: 9.3(3)

Key Features

Dual capability – ACI and NX-OS mode

First 400G-capable Cloud Scale platforms

400G ACI/standalone spine (9316D-GX)

100G leaf with 400G uplinks (93600CD)

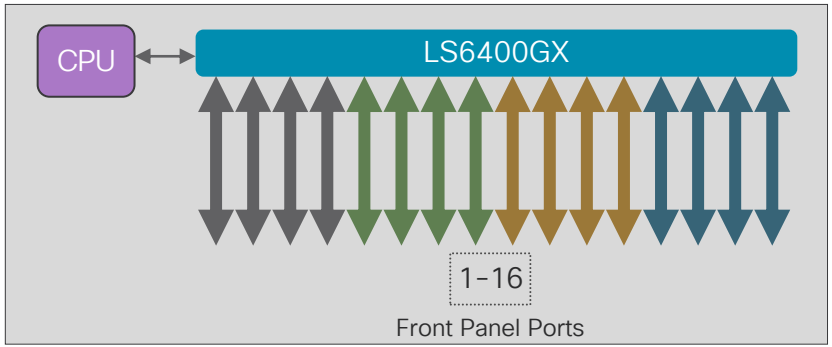
64-port 100G fixed TOR

400G/100G/50G/40G/10G with breakout capability

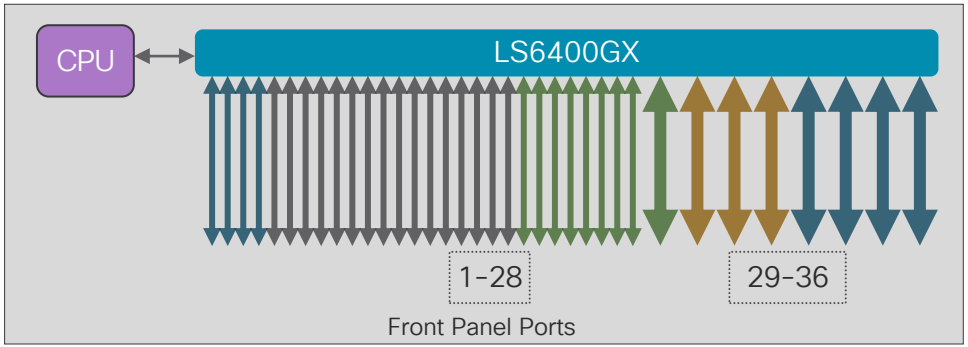
Flow Table for Network Insights, Netflow

Smart buffer capability (AFD / DPP)

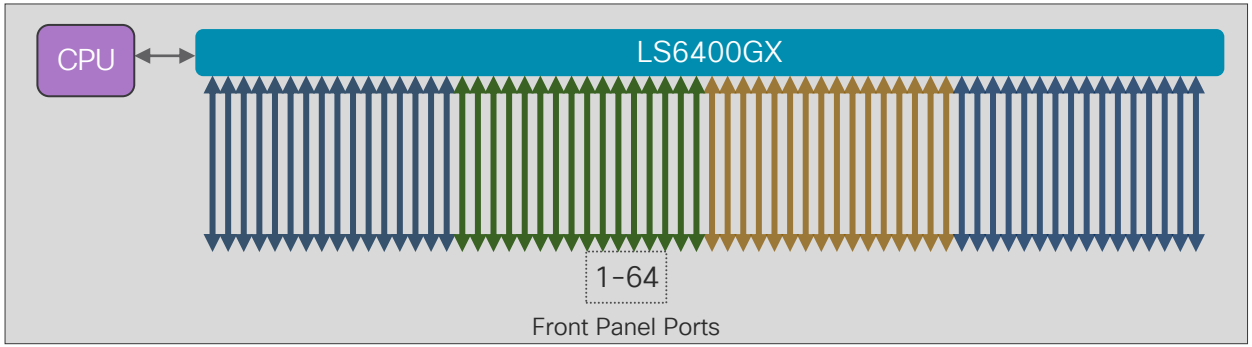
Nexus 9300-GX Switch Architecture



C9316D-GX (400G)



C93600CD-GX (100G + 400G)



C9364C-GX (100G)

Slice 0 Slice 2
Slice 1 Slice 3

Nexus 9300-GX2B Cloud Scale TOR Switches

32-port 400G QSFP-DD
N9K-C9332D-GX2B
LS12800GX2B-based
ACI: 5.2(3)
NX-OS: 10.2(1q)

Key Features

Dual capability – ACI and NX-OS mode

400G-capable Cloud Scale platforms

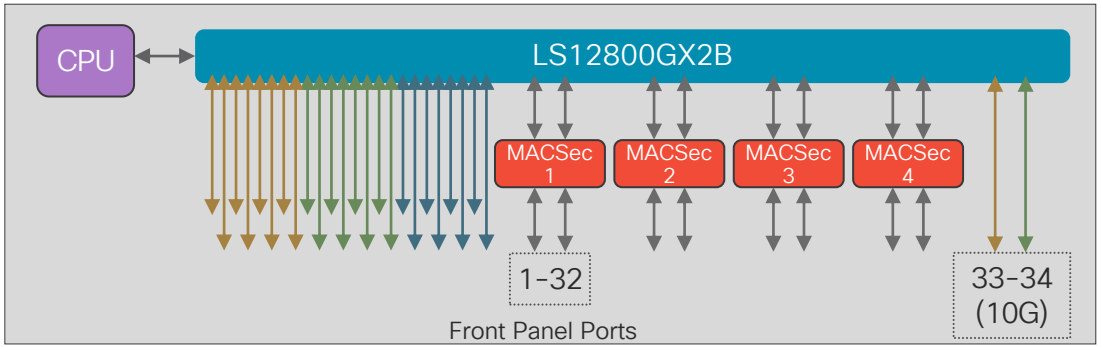
400G ACI/standalone spine

400G/100G/50G/40G/10G with breakout capability

Flow Table for Network Insights, Netflow

Smart buffer capability (AFD / DPP)

Nexus 9300-GX2B Switch Architecture



C9332D-GX2B (32x400G)

Slice 0 Slice 2
Slice 1 Slice 3

Nexus 9300-GX2A Cloud Scale TOR Switches

64-port 400G QSFP-DD

N9K-C9364D-GX2A – LS25600GX2A-based
ACI: 5.2(5)
NX-OS: 10.2(3)

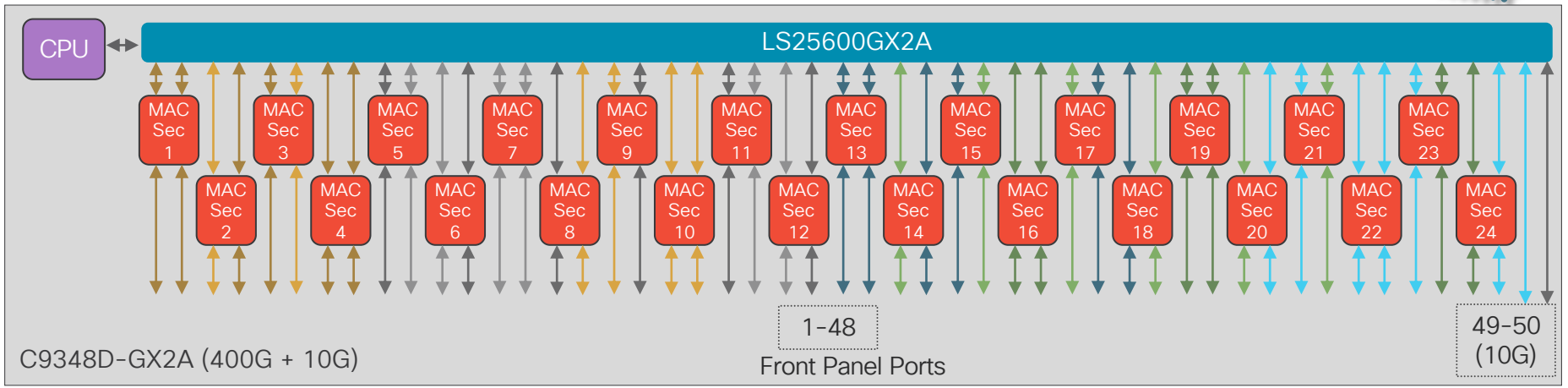
48-port 400G QSFP-DD

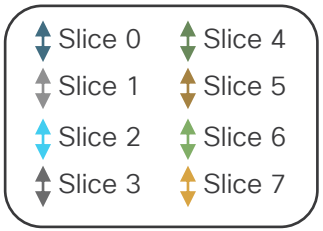
N9K-C9348D-GX2A – LS25600GX2A-based
ACI: 5.2(5)
NX-OS: 10.2(3)

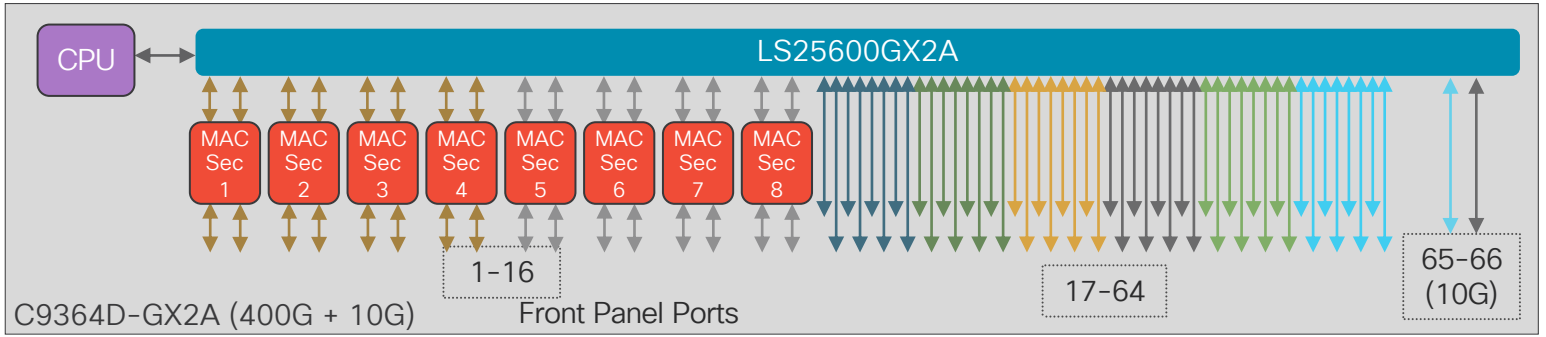
Key Features

- Dual capability – ACI and NX-OS mode
- 400G-capable Cloud Scale platforms
- 400G/100G/50G/40G/10G with breakout capability
- MACSEC support:
  - N9364D-GX2A: First 16 ports
  - N9348D-GX2A: All of the ports
- Flow Table for Network Insights, Netflow
- Smart buffer capability (AFD / DPP)

Nexus 9300-GX2A Switch Architecture







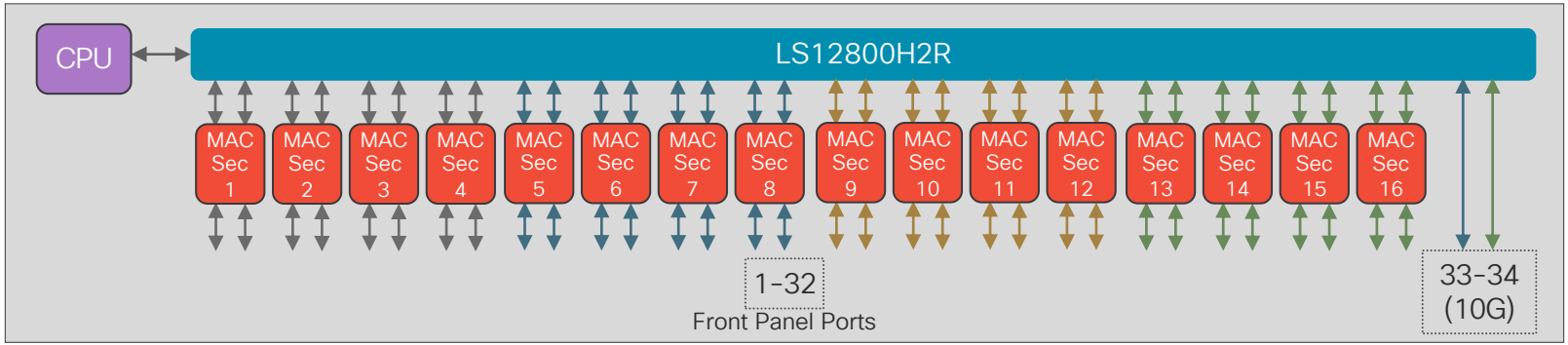
Nexus 9300-H2R Cloud Scale TOR Switches

32-port 400G QSFP-DD
N9K-C9332D-H2R
LS12800H2R-based
ACI: 6.1(1)
NX-OS: 10.4(1)

Key Features

- Dual capability – ACI and NX-OS mode
- 400G-capable Cloud Scale platforms
- 400G ACI/standalone spine
- 400G/100G/50G/40G/10G with breakout capability
- MACsec on all ports
- Flow Table for Network Insights, Netflow
- Smart buffer capability (AFD / DPP)
- Off-Chip 8GB HBM Buffer
- Telecom PTP and SyncE

Nexus 9300-H2R Switch Architecture



C9332D-H2R (32x400G)

Slice 0 Slice 2
Slice 1 Slice 3

Nexus 9300-H1 Cloud Scale TOR Switches

48-port 50G SFP56
4-port 400G QSFP-DD
N9K-C93400LD-H1
LS6400H1 -based
ACI: 6.1(1)
NX-OS: 10.4(2)

Key Features

Dual capability – ACI and NX-OS mode

Flexible port configurations – 100M/1/10/25/40/50/100G

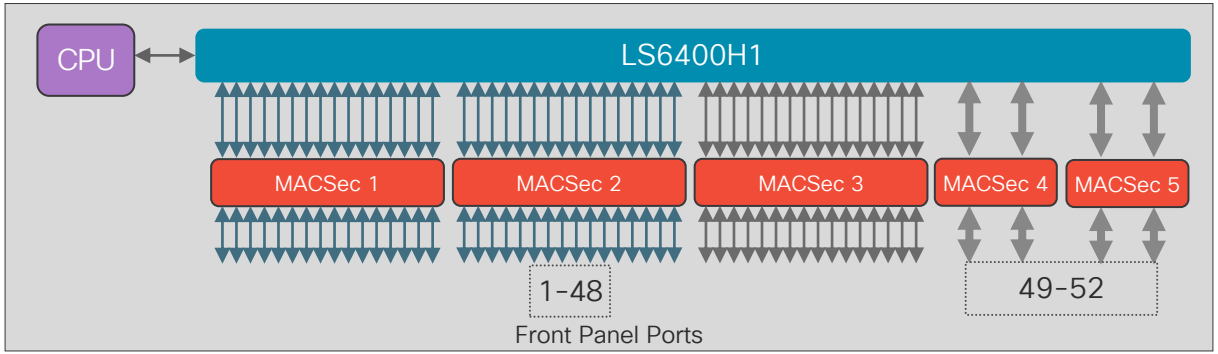
Flow Table for ND Insights, Netflow

MACSEC on all ports

Smart buffer capability (AFD / DPP)

Telecom PTP and SyncE

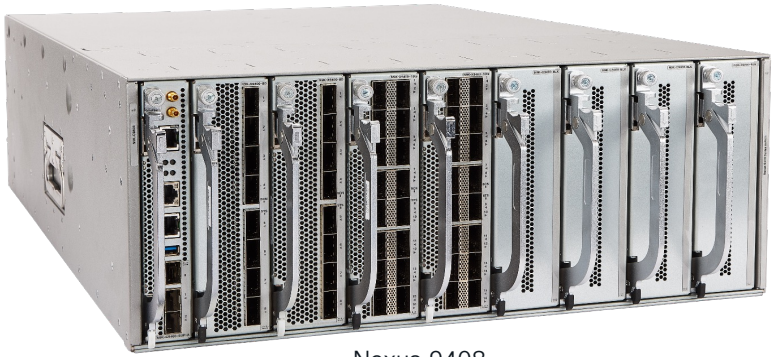
Nexus 9300-H1 Switch Architecture



C93400LD-H1 (10/25/50G + 400G)

Slice 0
Slice 1

Nexus 9400 Centralized Modular Switches



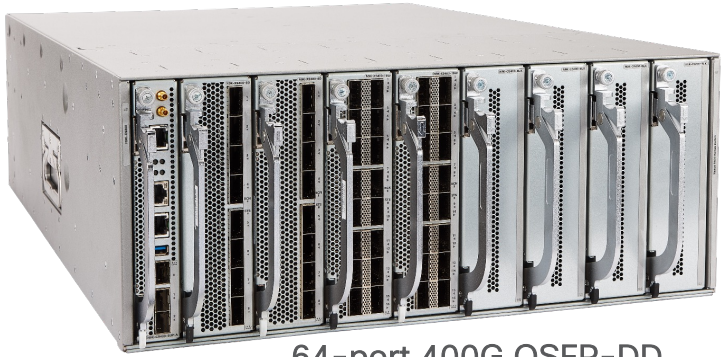
Nexus 9408

Chassis and Centralized Forwarding

+

Line-card Expansion-Module (LEM)

Nexus 9408 Cloud Scale Centralized Modular



64-port 400G QSFP-DD
N9K-C9408 – LS25600GX2A-based
ACI: 6.0(2)
NX-OS: 10.3(2)

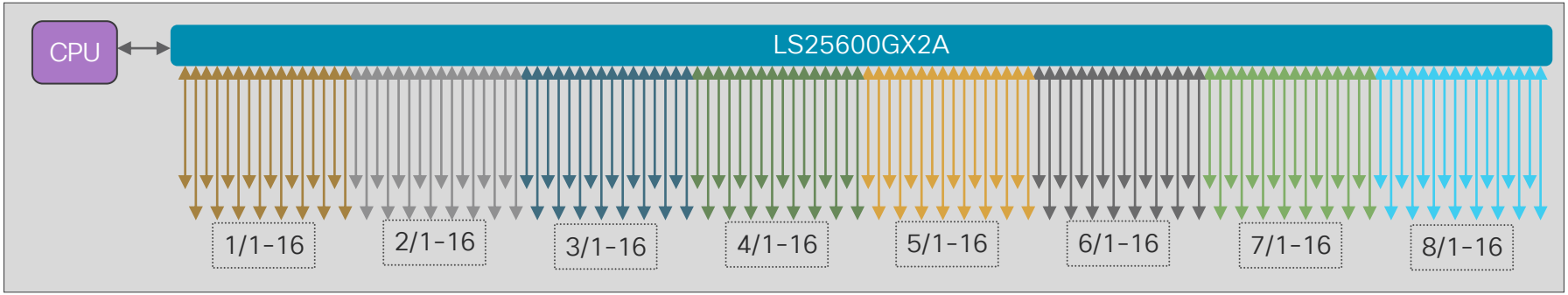
8-port 400G QSFP-DD
X9400-8D LEM

16-port 200G QSFP56
X9400-16W LEM

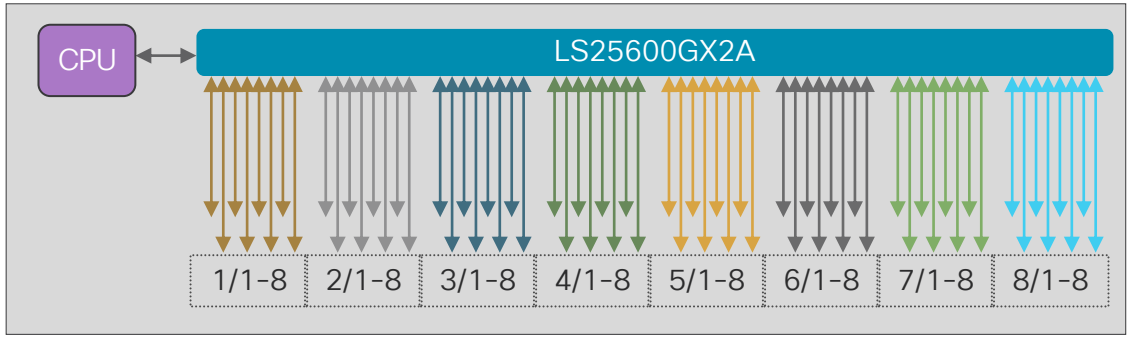
Key Features

- Dual capability – ACI and NX-OS mode
- 400G-capable Cloud Scale platforms
- Based on LS25600GX2A ASIC
- 400G/100G/50G/40G/10G with breakout capability
- MACSec support
- Flow Table for Network Insights, Netflow
- Smart buffer capability (AFD / DPP)
- Field replicable switch card
- Telecom PTP and SyncE

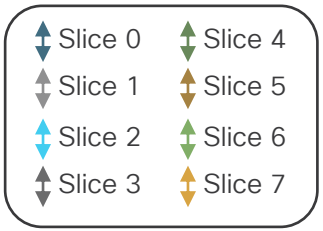
Nexus 9408 Switch Architecture



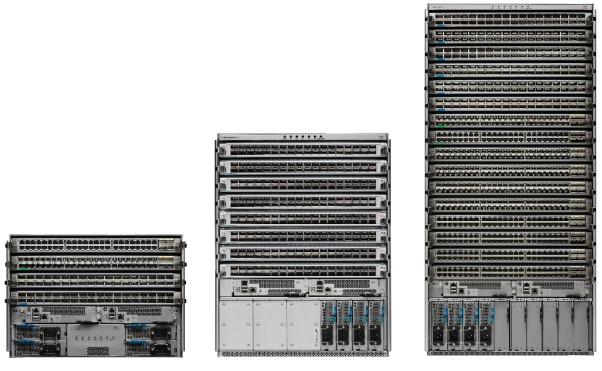
C9408 (128 x 200G)



C9408 (64 x 400G)



Nexus 9500 Modular Cloud Scale Switches



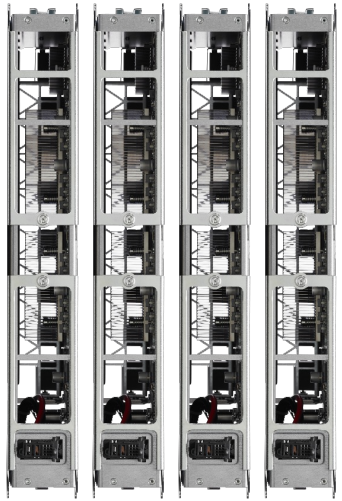
Nexus 9504 Nexus 9508 Nexus 9516

Common Equipment

+

FX / GX Series
Line Cards

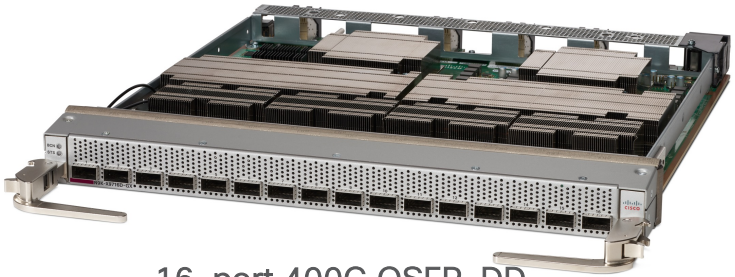
+



E2 / G-Series
Fabric Modules

X9700-GX 400G Cloud Scale Modules

N9K-X9716D-GX



16-port 400G QSFP-DD
X9716D-GX – LS6400GX-based
ACI: 5.1
NX-OS: 10.1(1)

Key Features

9716D-DX – Dual capability ACI and NX-OS

6.4Tbps capacity per module

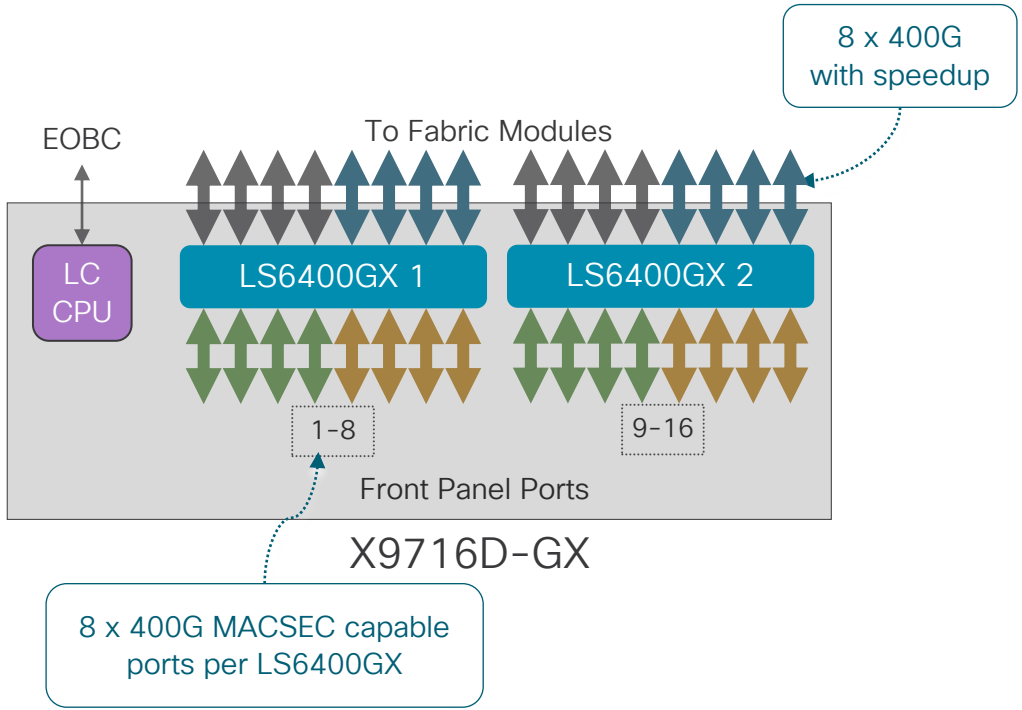
Flexible port configurations – 10/25/40/50/100/400G with breakout

Line-rate MACSEC on all ports

Flow Table for Network Insights, NetFlow

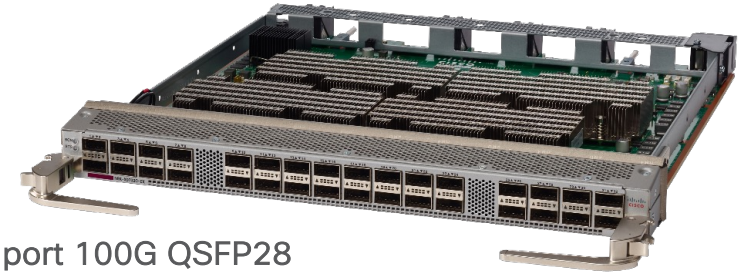
Smart buffer capability (AFD / DPP)

N9K-X9716D-GX Architecture



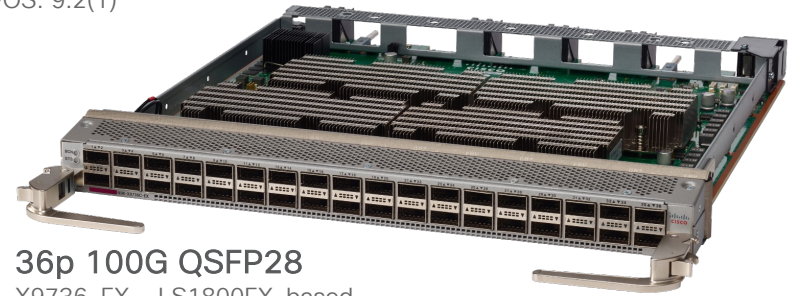
X9700-FX 100G Cloud Scale Modules

N9K-X9732C-FX / N9K-X9736C-FX



32-port 100G QSFP28

X9732C-FX – LS1800FX-based
ACI: Not supported
NX-OS: 9.2(1)



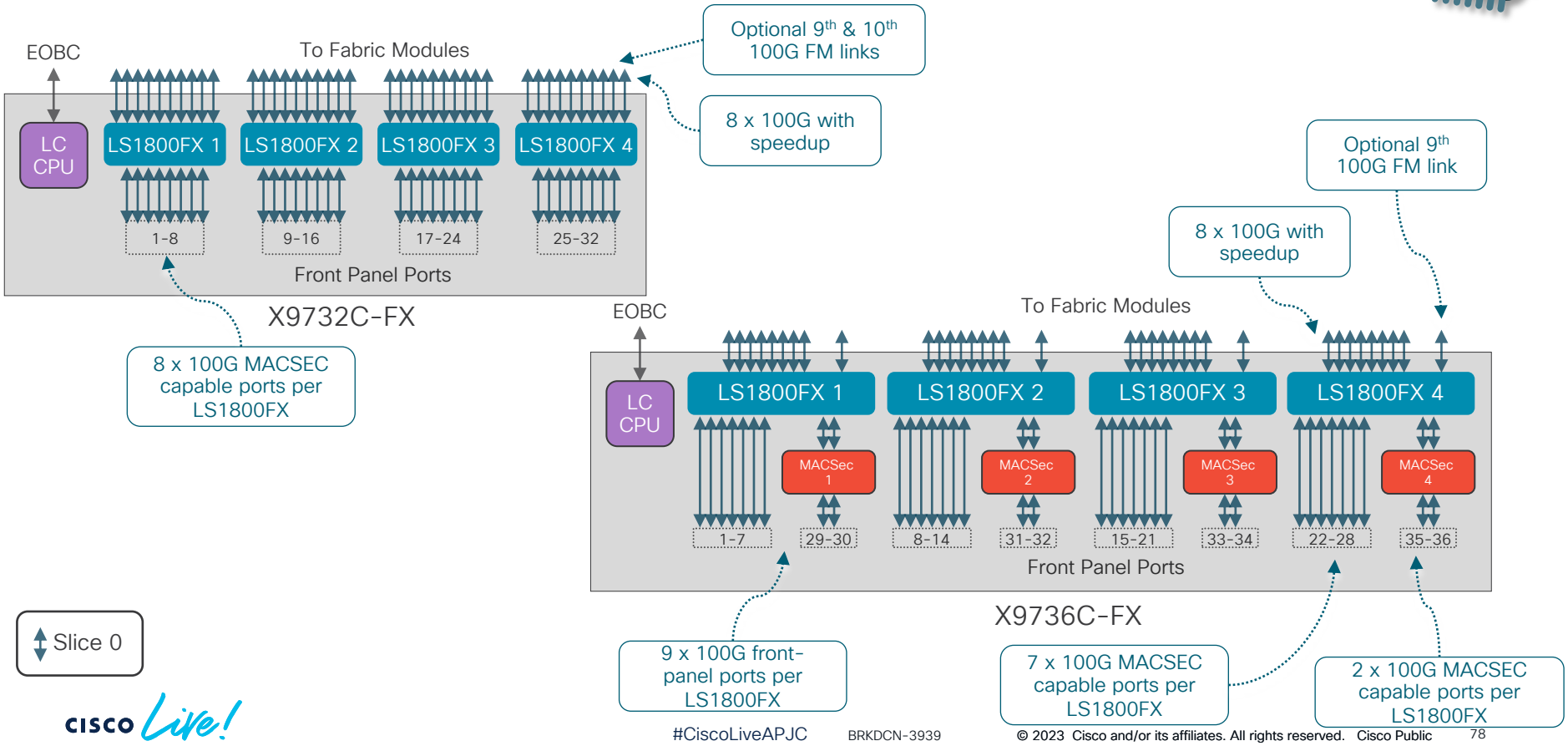
36p 100G QSFP28

X9736-FX – LS1800FX-based
ACI: 13.0(1)
NX-OS: NX-OS: 7.0(3)I7(3)

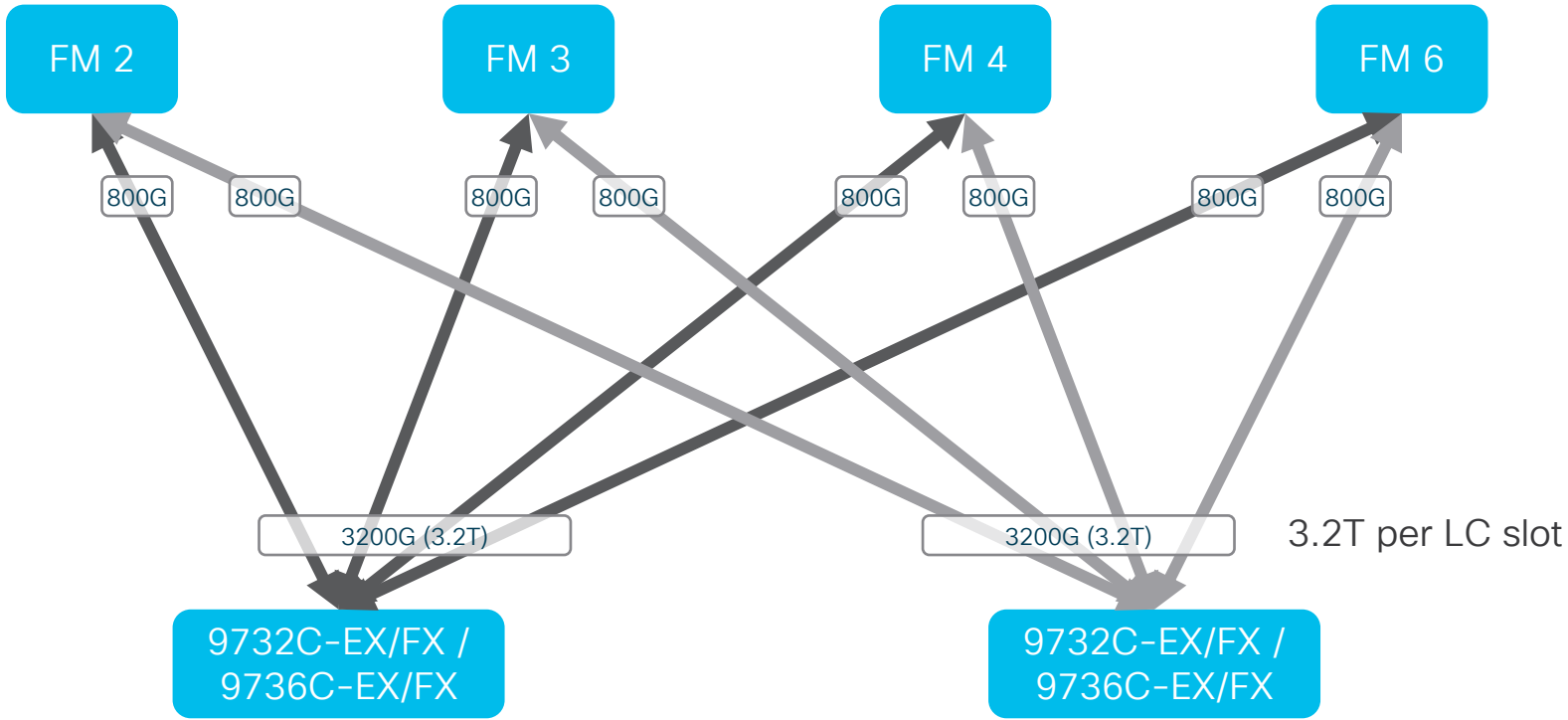
Key Features

- 9732C-FX – NX-OS only
- 9736C-FX – Dual capability ACI and NX-OS
- 3.2Tbps capacity per module
- 3.6Tbps capacity with optional 5th fabric module on 9736C-FX
- N+1 fabric redundancy option on 9732C-FX
- Flexible port configurations – 10/25/40/50/100G with breakout
- **Line-rate MACSEC on all ports**
- Flow Table for Network Insights, NetFlow
- Smart buffer capability (AFD / DPP)

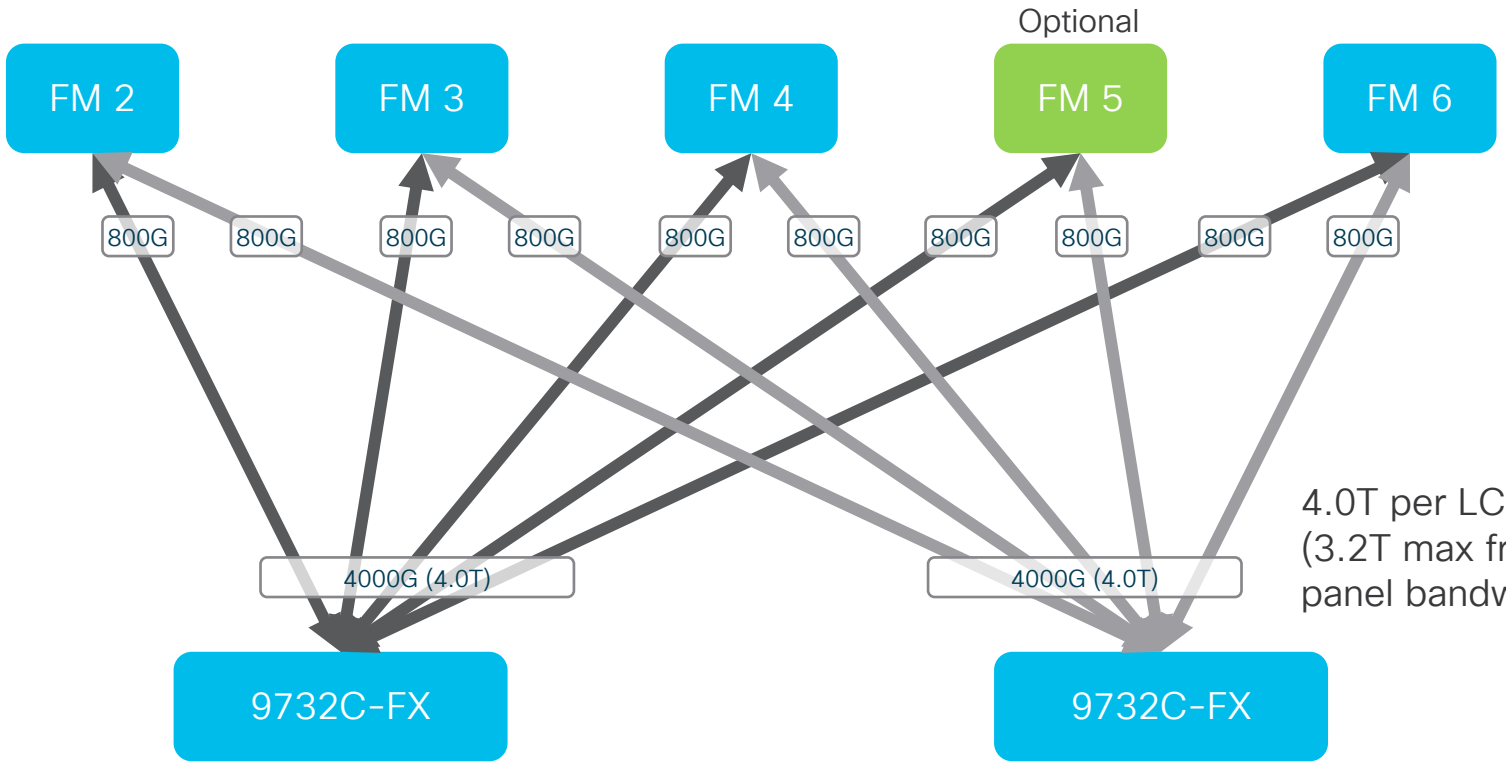
N9K-X9732C-FX / N9K-X9736C-FX Architecture



9732C-FX / 9736C-FX Fabric Connectivity

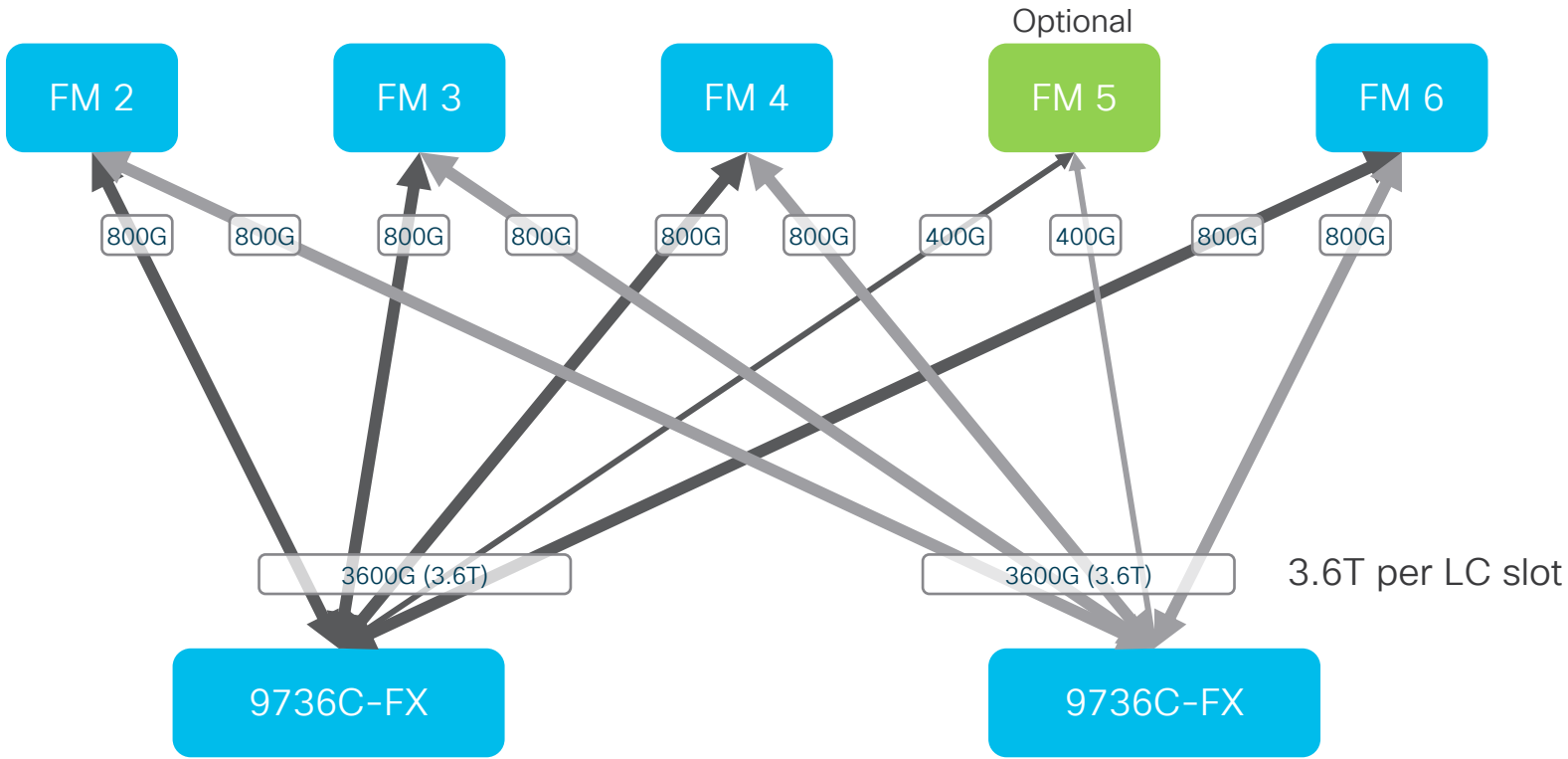


9732C-FX Fabric Connectivity – 5 FMs



4.0T per LC slot (3.2T max front panel bandwidth)

9736C-FX Fabric Connectivity – 5 FMs



Using 5 Fabric Modules

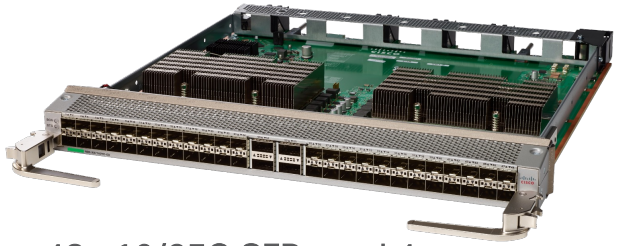
Limitations and Notes

- **All** modules installed in chassis must be either 9732C-FX or 9736C-FX to use 5 FMs
  - If other module type installed, 5th FM powered off automatically
- **9732C-FX:**
  - 5 FMs required for N+1 fabric module redundancy
- **9736C-FX:**
  - 5 FMs required for full bandwidth
  - Bandwidth reduction on FM failure varies depending on which FM failed

Note: 5 x FMs supported on all chassis types in standalone from 7.0(3)I7(2). 5 x FMs with 9736C-FX supported from in ACI 13.2(2).

X9700-EX/FX EOR/MOR Cloud Scale Modules

N9K-X97160YC-EX / N9K-X9788TC-FX

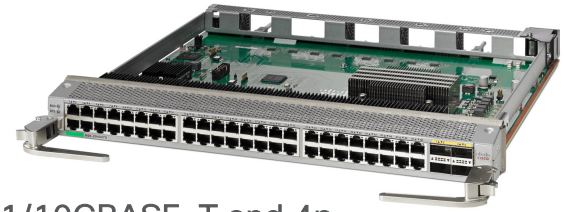


48p 10/25G SFP+ and 4p 100G QSFP28

X97160YC-EX – LSE-based

ACI: Not supported

NX-OS: 7.0(3)I5(2)



48p 1/10GBASE-T and 4p 100G QSFP28

X9788TC-FX – LS1800FX-based

ACI: Not supported

NX-OS: 7.0(3)I7(3)

Key Features

NX-OS mode only

Flow Table for Network Insights, NetFlow

Smart buffer capability (AFD / DPP)

97160-EX:

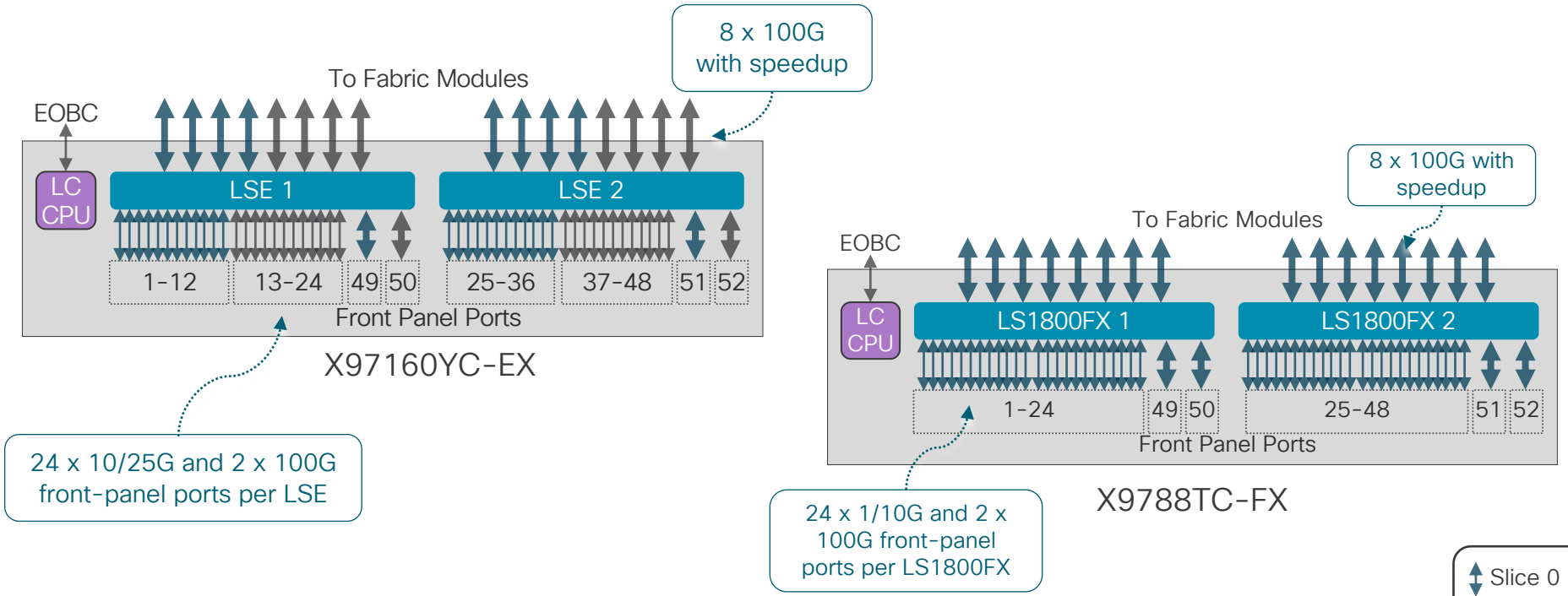
- 1.6Tbps capacity with line-rate performance
- Flexible port configurations –1/10/25G SFP28 ports, 1/10/25/40/50/100G QSFP28 ports

9788-FX:

- 880Gbps capacity with line-rate performance
- Flexible port configurations –1/10GBASE-T ports, 1/10/25/40/50/100G QSFP28 ports

Line-rate MACSEC on all ports

N9K-X97160YC-EX / N9K-X9788TC-FX Architecture

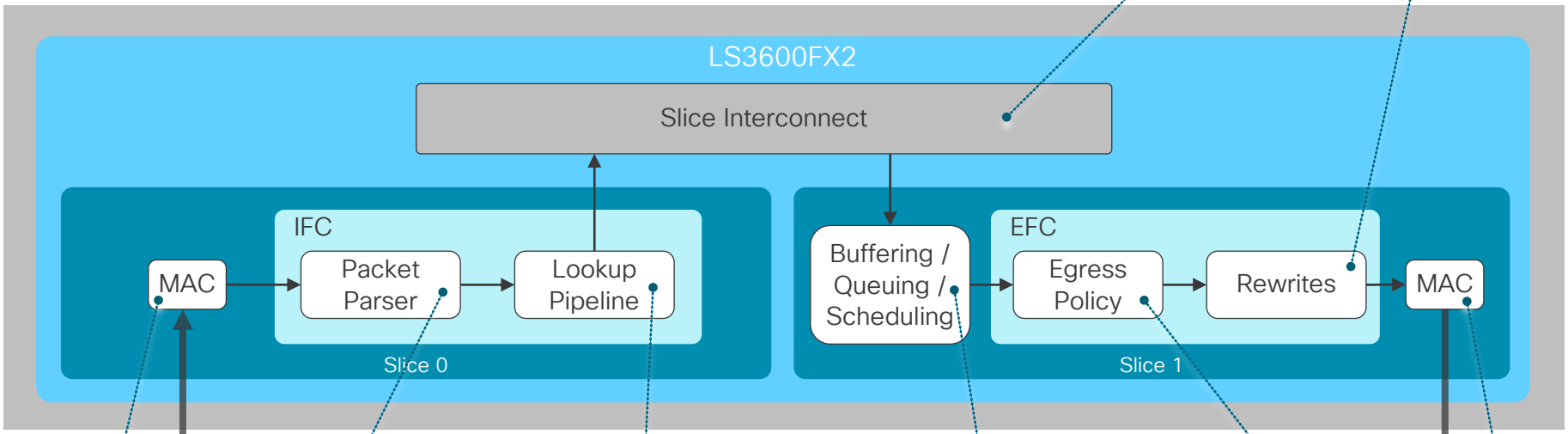


Agenda

- Data Center and Silicon Strategy
- Cloud Scale ASIC Architecture
- Cloud Scale Switching Platforms
- Packet Walks
- Key Takeaways

Packet Walk (TOR) – IP Unicast

9336C-FX2



Receive from Slice 0, transmit to Slice 1

Rewrite SMAC/DMAC
TTL decrement
Write QOS fields

Receive frame
FCS checking
VLAN checking

MTU checks
Extract header fields
Generate lookup key

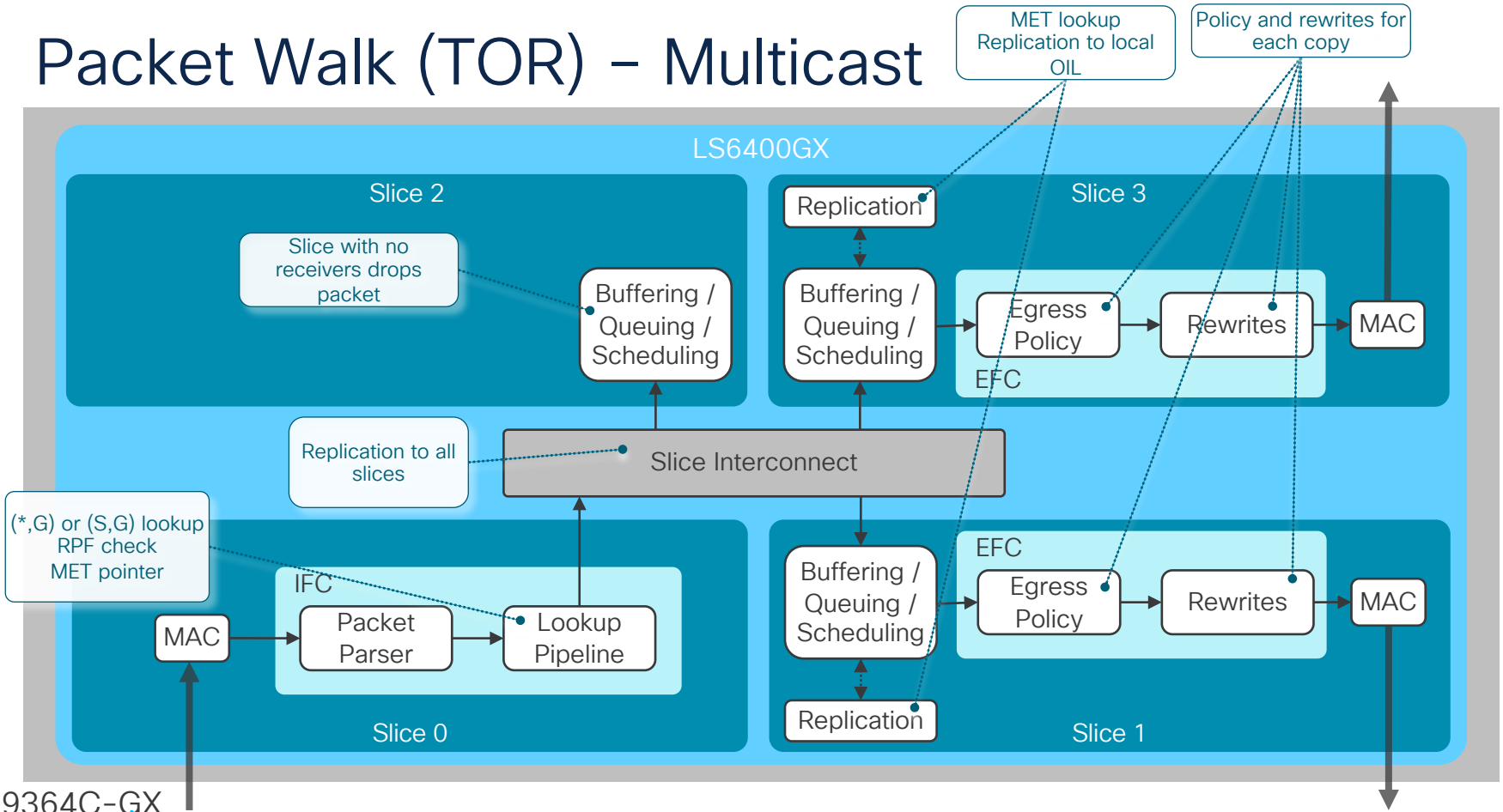
Longest-match prefix lookup
Adjacency/ECMP pointer
ECMP decision
Ingress policy enforcement
AFD/DPP
Flow table

Receive from slice interconnect
Buffer packet in queue (DBP)
AFD drops
Scheduling

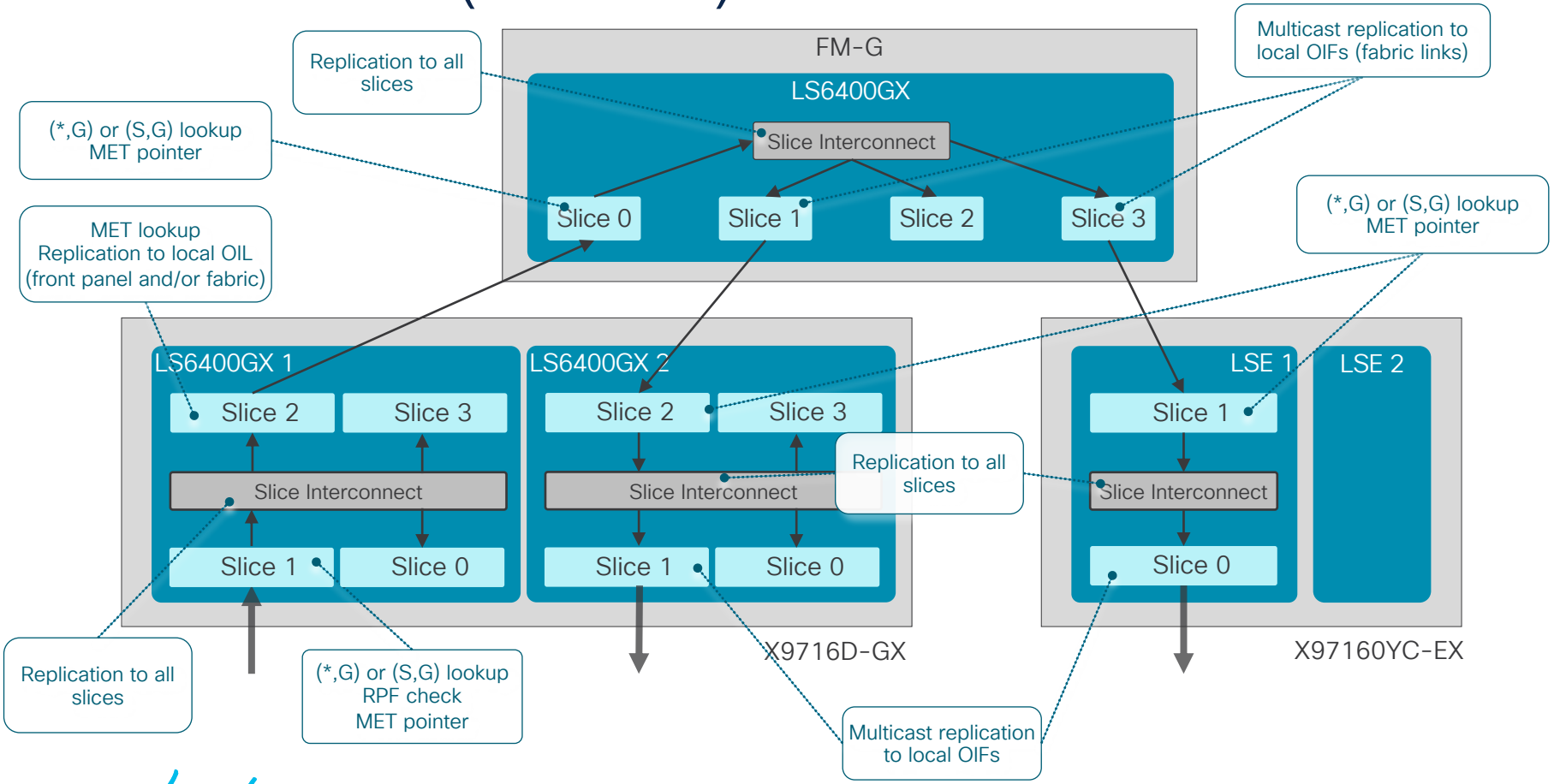
Egress policy enforcement

FCS generation
Transmit frame

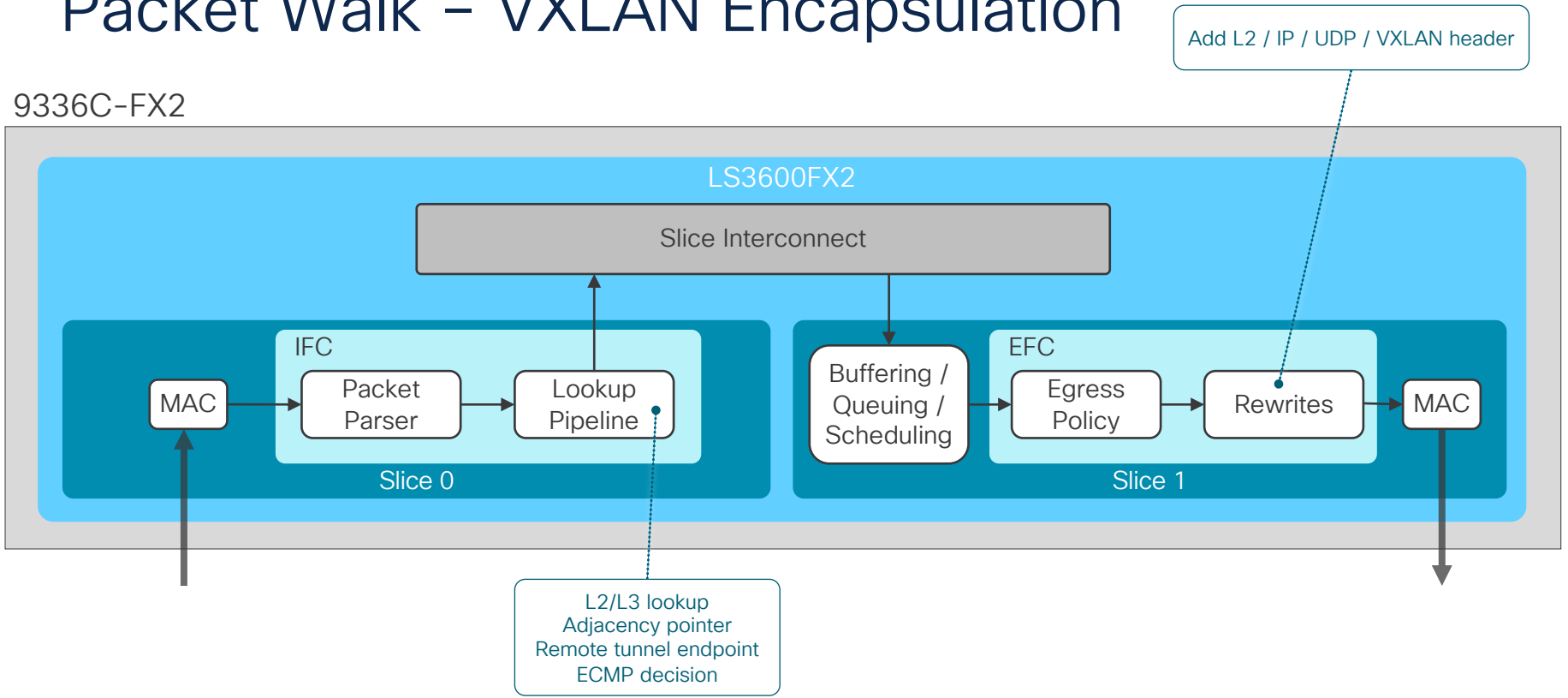
Packet Walk (TOR) – Multicast



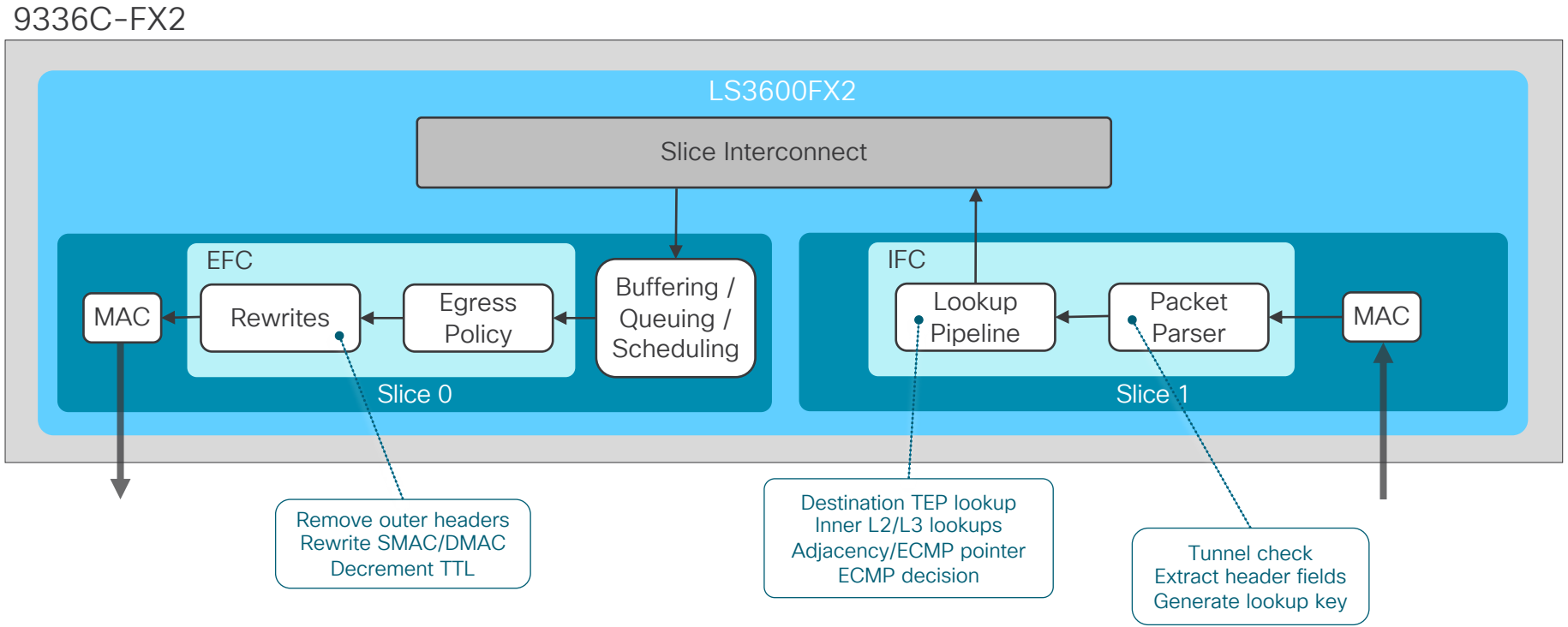
Packet Walk (Modular) – Multicast



Packet Walk – VXLAN Encapsulation



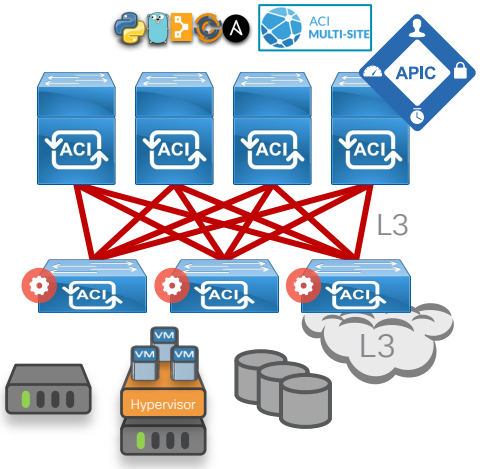
Packet Walk – VXLAN Decapsulation



Agenda

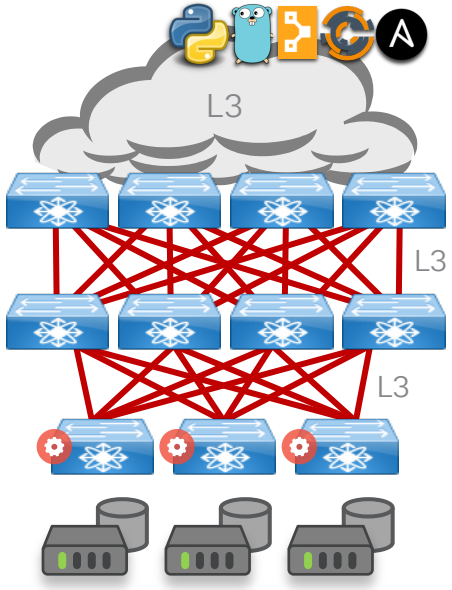
- Data Center and Silicon Strategy
- Cloud Scale ASIC Architecture
- Cloud Scale Switching Platforms
- Packet Walks
- Key Takeaways

Building Data Center Fabrics with Nexus 9000

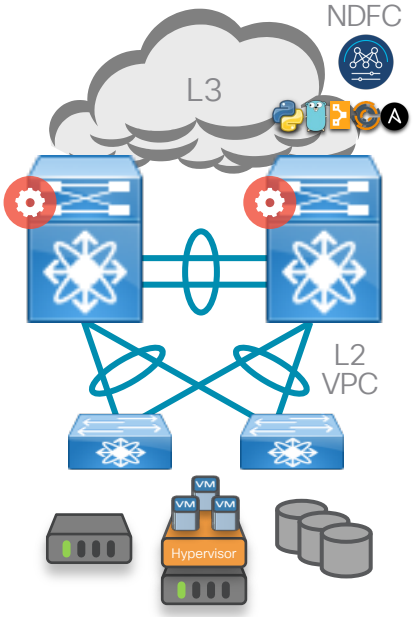


ACI – Turnkey Fabric

Standalone – Programmable Fabric with VXLAN+EVPN



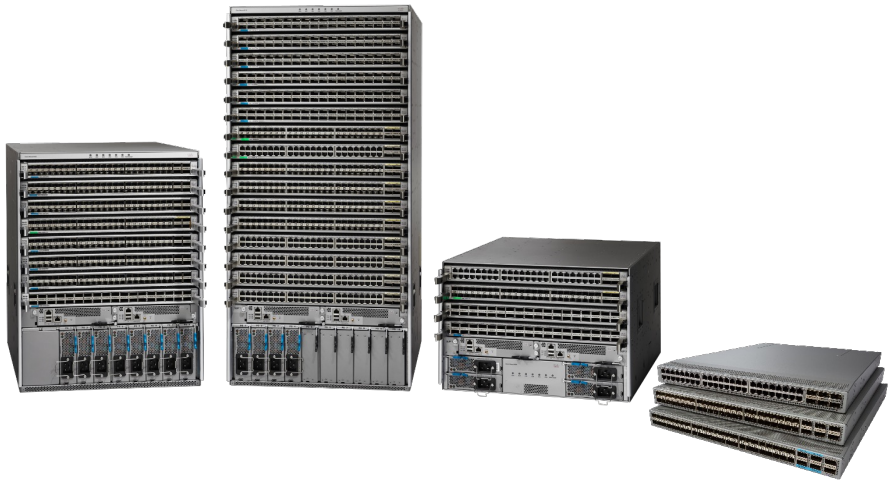
Standalone – Programmable IP Network



Standalone – Traditional Data Center Network

Key Takeaways

- You should now have a thorough understanding of the Nexus 9000 Cloud Scale switching platform architecture
- Feature-rich, innovative switching platform addresses virtually every deployment scenario
- Nexus 9000 Cloud Scale platform forms foundation of Cisco Data Center strategy



Recommended Sessions

- Cisco Live On-Demand Library:
  - BRKDCN-2112 - Network Best Practices for Artificial Intelligence Data Centre
  - BRKAPP-2698 – Building a High-Performance AI Data Center using Nexus, UCS, and NVIDIA GPUs
  - BRKDCN-3966 – VXLAN EVPN: A Deep Dive into Packet Forwarding
  - BRKSPG-3003 - Cisco 8000 Technical Update: Powered by Silicon One, IOS XR7 & SONiC

Recommended Literature

- Flexible Forwarding Table on Nexus 9000 White Paper
- Classification TCAM with Cisco CloudScale ASICs for Nexus 9000 Series Switches White Paper
- Intelligent Buffer Management on Cisco Nexus 9000 Series Switches White Paper
- Cisco Nexus 9800 Series Switches White Paper

cisco Live!

Did you know?

You can have a
one-on-one session with
a technical expert!

Visit Meet the Expert in The HUB
to meet, greet, whiteboard & gain
insights about your unique questions
with the best of the best.

Meet the Expert Opening Hours:

Tuesday	**3:00pm – 7:00pm**
Wednesday	**11:15am – 7:00pm**
Thursday	**9:30am – 4:00pm**
Friday	**10:30am – 1:30pm**

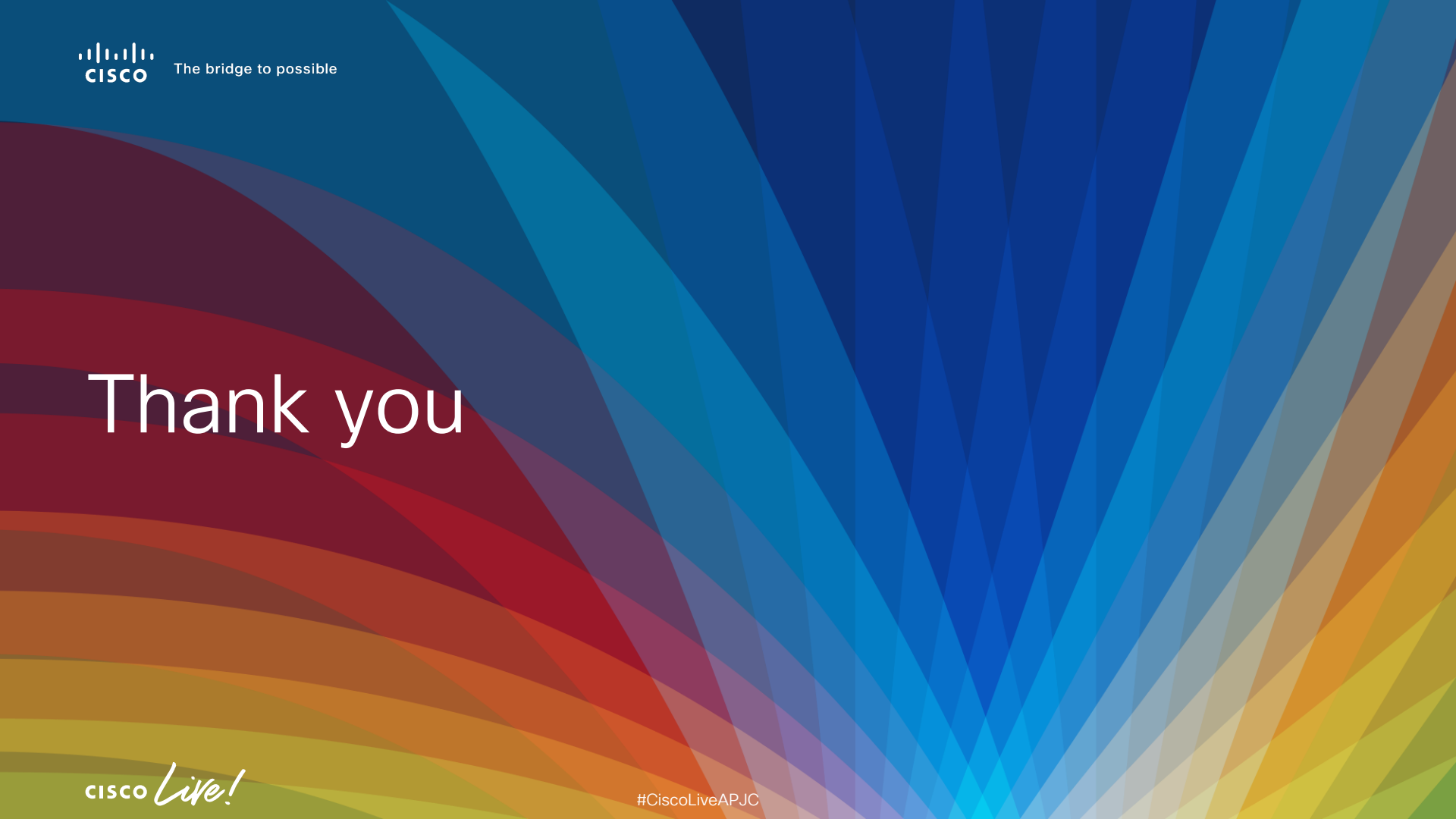
Session Surveys

We would love to know your feedback on this session!

- Complete a minimum of four session surveys and the overall event surveys to claim a Cisco Live T-Shirt

Continue your education

- Visit the Cisco Showcase for related demos
- Book your one-on-one Meet the Expert meeting
- Attend the interactive education with DevNet, Capture the Flag, and Walk-in Labs
- Visit the On-Demand Library for more sessions at www.CiscoLive.com/on-demand


cisco
The bridge to possible
Thank you
cisco Live!
#CiscoLiveAPJC

cisco Live!

Let's go

#CiscoLiveAPJC