
TOMORROW
starts here.

Nexus 7000 / 7700 Architecture and Design Flexibility for Evolving Data Centres

BRKARC-3601

Tim Stevenson
Distinguished Engineer, Technical Marketing

#clmel

Cisco live!

Session Abstract

This session will discuss the foundations of the Nexus 7000 and 7700 series switches, including chassis, I/O modules, and NX-OS software. Examples will show common use-cases for different module types and considerations for module interoperability. The focus will then shift to key platform capabilities and features – including VPC, FabricPath, OTV, VDCs, and others – along with real-world designs and deployment models. The session concludes with a discussion of emerging architectures and designs, including the role of Nexus 7000 and 7700 in VXLAN and Application Centric Infrastructure (ACI) environments.

Session Goals

- To provide an understanding of the Nexus 7000 / Nexus 7700 switching architecture, which provides the foundation for flexible, scalable Data Centre designs
- To examine key Nexus 7000 / Nexus 7700 design building blocks and illustrate common design alternatives leveraging those features and functionalities
- To see how the Nexus 7000 / Nexus 7700 platform plays in emerging technologies and architectures

4

Other Relevant Sessions

- BRKARC-3470 – Advanced – Cisco Nexus 7000/7700 Switch Architecture (2014 San Francisco – www.ciscolive.com)
- BRKDCT-2048 - Deploying Virtual Port Channel (vPC) in NXOS (Cisco Live Melbourne 2015)
- BRKDCT-2081 - Cisco FabricPath Technology and Design (2014 San Francisco – www.ciscolive.com)
- BRKDCT-2121 – Virtual Device Context (VDC) Design and Implementation Considerations with Nexus 7000 (2014 San Francisco – www.ciscolive.com)
- BRKDCT-2049 – Data Centre Interconnect with Overlay Transport Virtualisation (Cisco Live Melbourne 2015)
- BRKDCT-3445 – Building scalable data centre networks with NX-OS and Nexus 7000 (2014 San Francisco – www.ciscolive.com)
- BRKDCT-2404 – VXLAN Deployment Models – A Practical Perspective (Cisco Live Melbourne 2015)

Agenda

- Introduction to Nexus 7000 / Nexus 7700
- Nexus 7000 / Nexus 7700 Architecture
  - Chassis
  - Supervisor engines and NX-OS software
  - I/O modules (M2/F2E/F3)
- I/O Module Interoperability
- Data Centre Designs with Nexus 7000 / Nexus 7700
  - STP/VPC
  - L4-7 services integration
  - VDCs
  - FabricPath
  - VRF/MPLS VPNs
  - OTV
- Next-Generation Data Centres with Nexus 7000 / Nexus 7700
  - Evolved FabricPath
  - ACI integration
  - VXLAN / VXLAN + EVPN

Introduction to Nexus 7000 / Nexus 7700

Introduction to Nexus 7000 / Nexus 7700 Platform

Data-centre class Ethernet switches designed to deliver high performance, high availability, system scale, and investment protection

Designed for wide range of Data Centre deployments, focused on feature-rich 10G/40G/100G density and performance

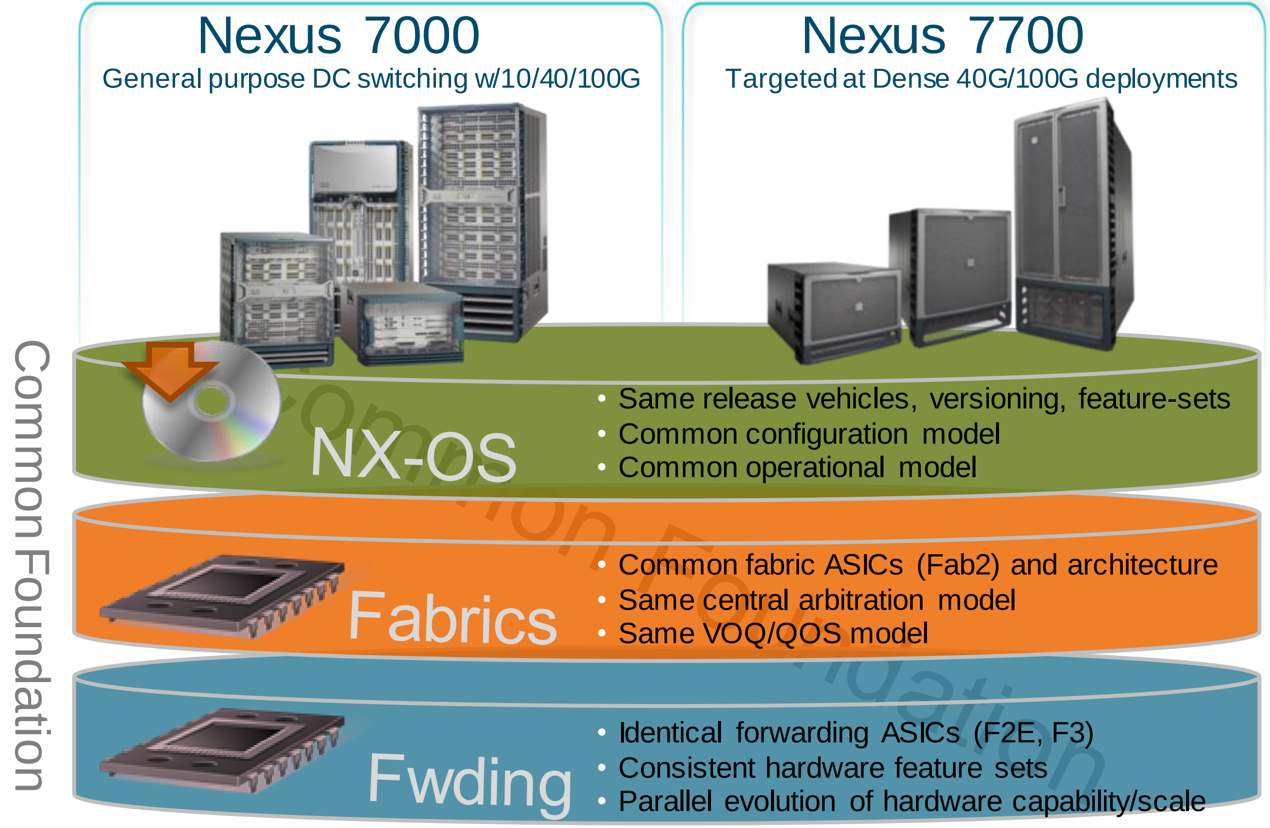
I/O Modules

Chassis

Supervisor Engines

Fabrics

Nexus 7000 / Nexus 7700 – Common Foundation



Agenda

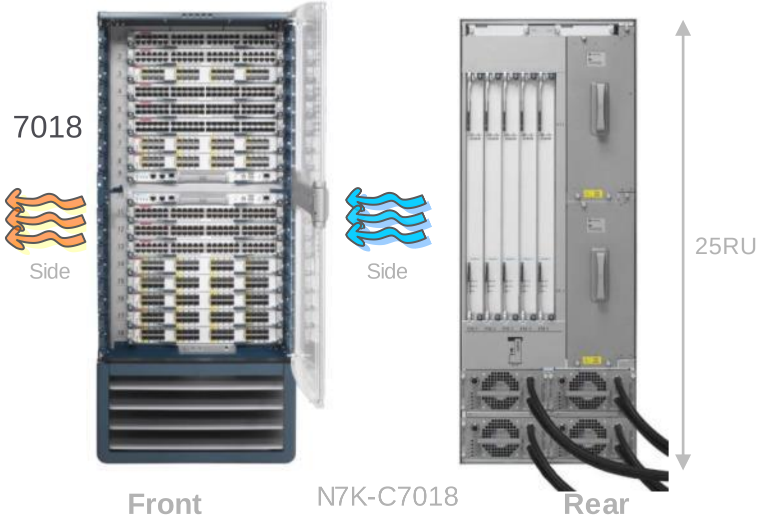
- Introduction to Nexus 7000 / Nexus 7700
- **Nexus 7000 / Nexus 7700 Architecture**
  - **Chassis**
  - **Supervisor engines and NX-OS software**
  - **I/O modules (M2/F2E/F3)**
- I/O Module Interoperability
- Data Centre Designs with Nexus 7000 / Nexus 7700
  - STP/VPC
  - L4-7 services integration
  - VDCs
  - FabricPath
  - VRF/MPLS VPNs
  - OTV
- Next-Generation Data Centres with Nexus 7000 / Nexus 7700
  - Evolved FabricPath
  - ACI integration
  - VXLAN / VXLAN + EVPN

Nexus 7000 / Nexus 7700 Architecture

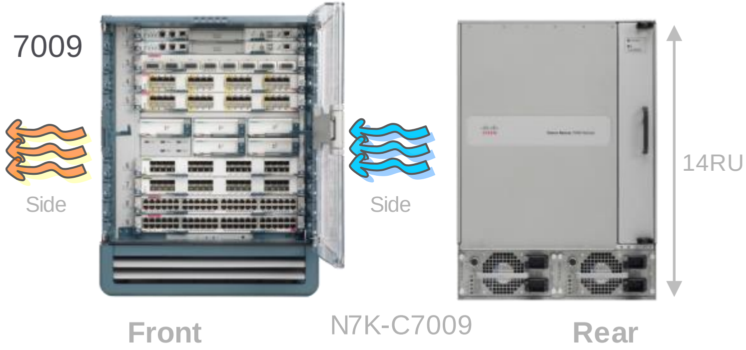
Nexus 7000 Chassis Family

Nexus 7010

Nexus 7018



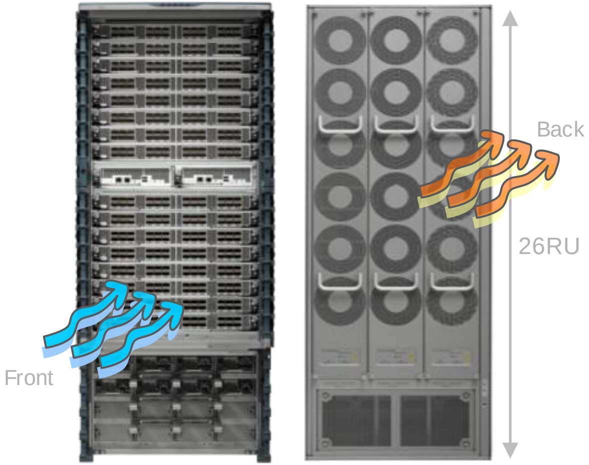
Nexus 7009



Nexus 7004

Nexus 7700 Chassis Family

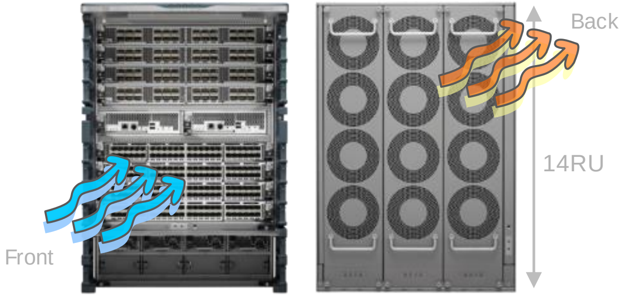
Nexus 7718



Front Rear

N77-C7718

Nexus 7710



Front Rear

N77-C7710

Nexus 7706

Front Rear

N77-C7706

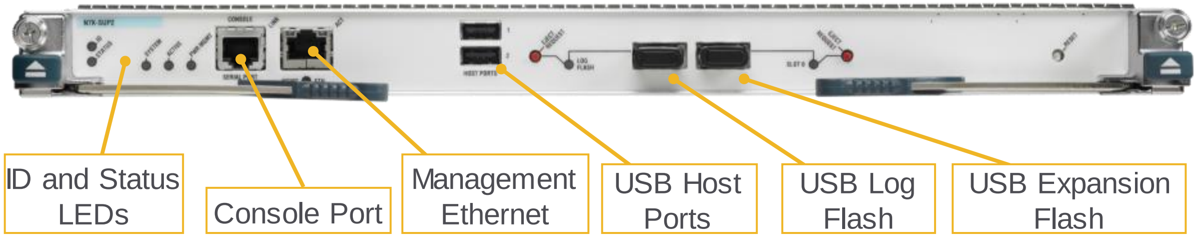
Supervisor Engine 2 / 2E

- System supervisor engines providing control plane and management functions

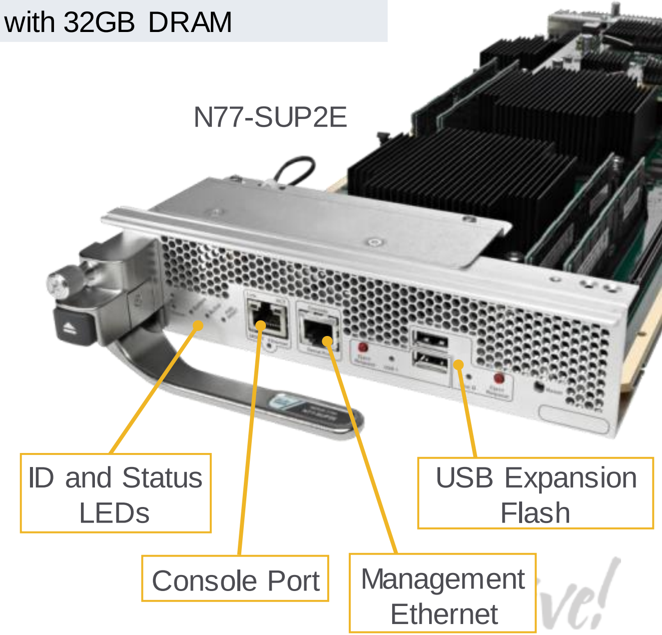
Supervisor Engine 2 (Nexus 7000)	Supervisor Engine 2E (Nexus 7000 / Nexus 7700)
Base performance	High performance
One quad-core 2.1GHz CPU with 12GB DRAM	Two quad-core 2.1GHz CPU with 32GB DRAM

- Connects to fabric via 1G inband interface

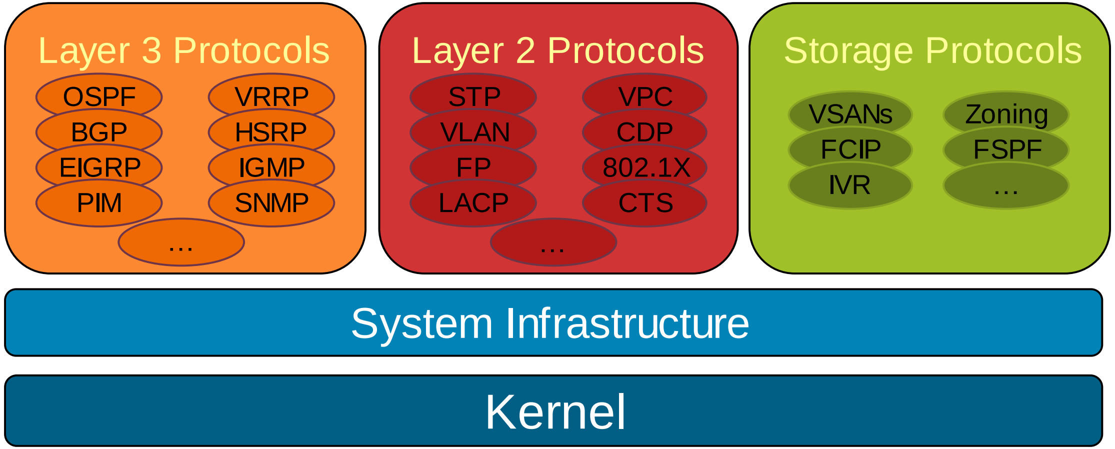
N7K-SUP2/N7K-SUP2E



N77-SUP2E



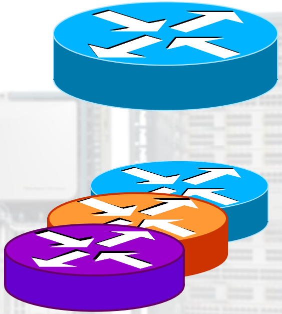
NX-OS



- Linux kernel provides preemptive multitasking, virtual memory, threading, etc.
- System infrastructure: Reliable messaging (IPC), state database, process management / monitoring
- Comprehensive Layer 3 protocol implementation
- Data-centre focused Layer 2 feature set
- Storage feature set from SAN-OS

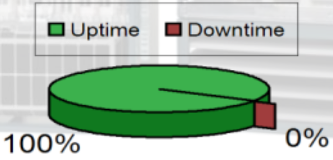
Nexus 7000 / Nexus 7700 and NX-OS – Comprehensive Data Centre Feature Set

Layer 3
- Distributed IPv4 and IPv6 unicast hardware forwarding
- OSPF, EIGRP, IS-IS, BGP, RIP, PBR
- PIM-SM, SSM, Bidir, MSDP, MP-BGP
- IGMP/MLD
- 32-way ECMP
- HSRP, GLBP, VRRP with object tracking



Virtualisation
- VLANs/VRF-lite
- MPLS VPNs
- Virtual Device Contexts (VDCs)
- Overlay Transport Virtualisation (OTV)
- Location/ID Separation Protocol (LISP)

High Availability
- In-Service Software Upgrade (ISSU)
- Stateful supervisor switchover
- Stateful process restarts
- Graceful restart for routing protocols
- Smart Call Home
- GOLD



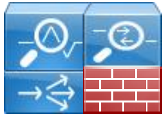
Security
- RACLs, VACLs, PACLs
- Cisco TrustSec: SGACLs, LinkSec (802.1AE)
- CoPP and rate limiters
- DHCP snooping, DAI, IP source guard
- Port security and 802.1x
- Storm control
- Unicast RPF check
- Roles-based management

Layer 2
- Distributed Layer 2 hardware switching
- Hardware MAC learning
- PVRST, MST
- Virtual Port Channels (VPC)
- FabricPath
- IGMP snooping
- BPDU Guard, Root Guard, BPDU Filter, Bridge Assurance
- Link Aggregation Control Protocol (LACP/802.1AD)
- Private VLANs

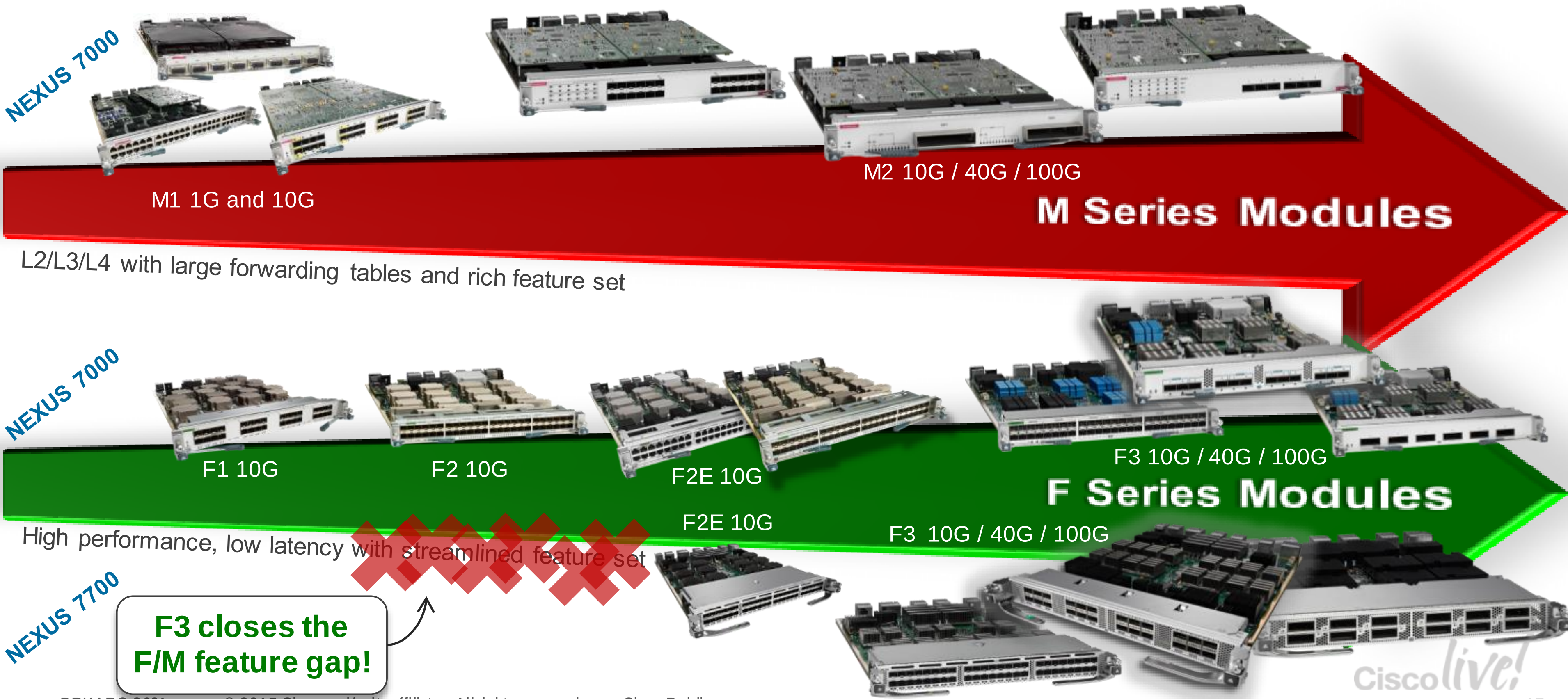
Operational Manageability & Programmability
- Python/TCL shell
- XMPP
- Netflow and NDE
- SPAN/ERSPAN
- SNMP/XML
- Configuration rollback
- EEM

Quality of Service
- Ingress/ egress queuing
- Marking policies and mutation
- Ingress and egress policing
- Colour-aware policing
- MQC CLI model

L4-7 Services
- NAM service module (Nexus 7000)
- Remote Integrated Services (RISE)
- Intelligent Traffic Director (ITD)
- WCCP



Nexus 7000 / 7700 I/O Module Families



Nexus 7000 M2 I/O Modules

N7K-M224XP-23L / N7K-M206FQ-23L / N7K-M202CF-22L

- 10G / 40G / 100G M2 I/O modules
- Share common hardware architecture
- Two integrated forwarding engines (120Mpps)
- Feature-rich L2/L3/L4 with large tables

N7K-M224XP-23L

N7K-M206FQ-23L

N7K-M202CF-22L

Module	Port Density	Optics	Bandwidth
M2 10G	24 x 10G (plus Nexus 2000 FEX support)	SFP+	240G
M2 40G	6 x 40G (or up to 24 x 10G via breakout)	QSFP+	240G
M2 100G	2 x 100G	CFP	200G

Nexus 7000 / Nexus 7700 F2E I/O Modules

N7K-F248XP-25E / N7K-F248XT-25E / N77-F248XP-23E

- 48-port 1G/10G with SFP/SFP+ transceivers
- 480G full-duplex fabric connectivity
- System-on-chip (SOC) forwarding engine design
  - 12 independent SOC ASICs
- Layer 2/Layer 3 forwarding with L3/L4 services (ACL/QOS)
- Interoperability with M1/M2, in Layer 2 mode on Nexus 7000
  - Proxy routing for inter-VLAN/L3 traffic

Nexus 7000
N7K-F248XP-25E

Nexus 7000
N7K-F248XT-25E

Nexus 7700
N77-F248XP-23E

Nexus 7000 F3 I/O Modules

N7K-F348XP-25 / N7K-F312FQ-25 / N7K-F306CK-25

- 10G / 40G / 100G F3 I/O modules
- Share common hardware architecture
- SOC-based forwarding engine design
  - 6 independent SOC ASICs per module
- Layer 2/Layer 3 forwarding with L3/L4 services (ACL/QOS) and advanced features
- **Require Supervisor Engine 2 / 2E**

N7K-F348XP-25

N7K-F312FQ-25

N7K-F306CK-25

Module	Port Density	Optics	Bandwidth
F3 10G	48 x 1/10G (plus Nexus 2000 FEX support)	SFP+	480G
F3 40G	12 x 40G (or up to 48 x 10G via breakout)	QSFP+	480G
F3 100G	6 x 100G	CPAK	550G

* Roadmap item

Nexus 7700 F3 I/O Modules

N7K-F348XP-25 / N7K-F312FQ-25 / N7K-F306CK-25

- 10G / 40G / 100G F3 I/O modules
- Share common hardware architecture
- SOC-based forwarding engine design
  - 6 independent SOC ASICs per 10G module
  - 12 independent SOC ASICs per 40G/100G module
- Layer 2/Layer 3 forwarding with L3/L4 services (ACL/QOS) and advanced features


N77-F348XP-23


N77-F324FQ-25

Module	Port Density	Optics	Bandwidth
F3 10G	48 x 1/10G (plus Nexus 2000 FEX support)	SFP+	480G
F3 40G	24 x 40G (or up to 76 x 10G + 5 x 40G via breakout)	QSFP+	960G
F3 100G	12 x 100G	CPAK	1.2T


N77-F312CK-26

* Roadmap item

Agenda

- Introduction to Nexus 7000 / Nexus 7700
- Nexus 7000 / Nexus 7700 Architecture
  - Chassis
  - Supervisor engines and NX-OS software
  - I/O modules (M2/F2E/F3)
- **I/O Module Interoperability**
- Data Centre Designs with Nexus 7000 / Nexus 7700
  - STP/VPC
  - L4-7 services integration
  - VDCs
  - FabricPath
  - VRF/MPLS VPNs
  - OTV
- Next-Generation Data Centres with Nexus 7000 / Nexus 7700
  - Evolved FabricPath
  - ACI integration
  - VXLAN / VXLAN + EVPN

I/O Module Interoperability

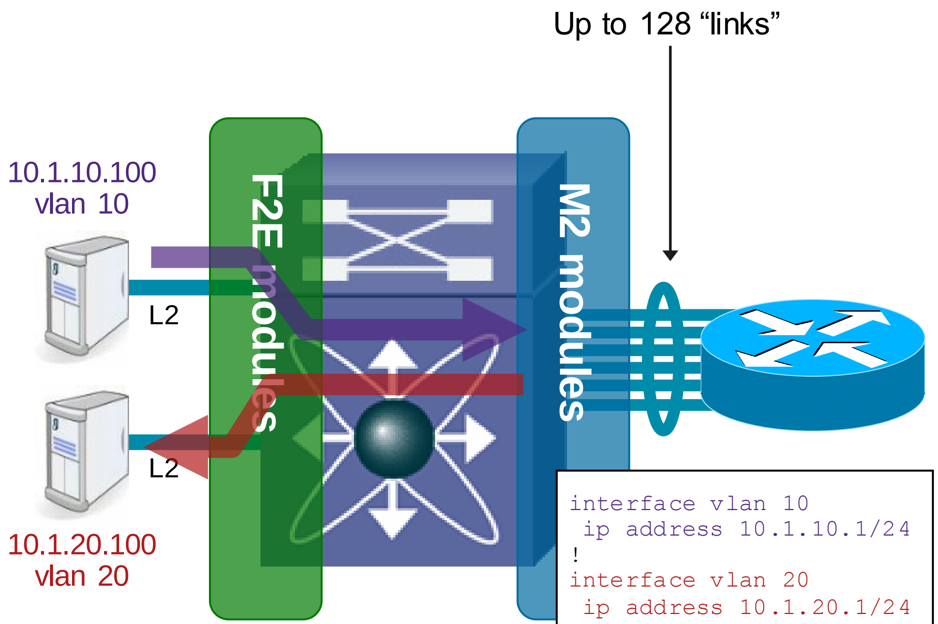
I/O Module Interoperability

- General module interoperability rule is: “+/-1 generation” in same Virtual Device Context (VDC)
- System-level coexistence based on chassis support matrix
  - E.g., cannot run F1 modules in Nexus 7004
- Layer 3 forwarding behaviour in VDC is key difference between interop models:
  - “Proxy Forwarding”
  - “Ingress Forwarding” with Lowest Common Denominator

Proxy Forwarding Model

M2 + F2E VDC

- F2E modules run in pure Layer 2 mode – all L3 functions disabled
- From F2E perspective, Router MAC reachable through port-channel with all ports on M2 modules
- All packets destined to Router MAC forwarded through fabric toward one “member port” in that channel
- M2 modules(s) perform all L3 forwarding and policy, pass packets back over fabric to output port
- Key consideration: M-series L3 routing capacity versus F-series front-panel port count




```
interface vlan 10
 ip address 10.1.10.1/24
!
interface vlan 20
 ip address 10.1.20.1/24
```


Ingress Forwarding with Lowest Common Denominator Model

M2 + F3 VDC *-or-* F2E + F3 VDC

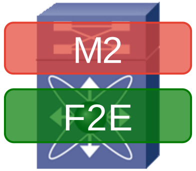
- F3 module interoperability always “Ingress Forwarding” – NO proxy forwarding
  - Ingress module makes all forwarding decisions
- Supported feature set and scale based on Lowest Common Denominator
  - Feature available if **all** modules support the feature

Not all features supported by software today…

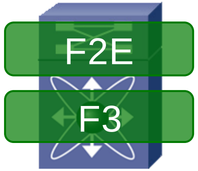
Module Types in VDC	Layer 2	Layer 3	VPC	Fabric Path	VXLAN	FEX	MPLS	OTV	LISP	FCOE	Table Sizes
F3	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	F3 size
M2 + F3	✓	✓	✓	✗	✗	✓	✓	✓	✗	✗	F3 size
F2E + F3	✓	✓	✓	✓	✗	✓	✗	✗	✗	✓	F2E size
M2 + F2E + F3	Not supported										

Module Interoperability Use Cases

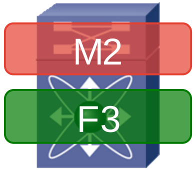
- M2 + F2E VDC
  - Provide higher-density 10G while supporting M2 features and L3 functions
  - Full internet routes, MPLS VPNs
  - FabricPath with increased MAC address scale (proxy L2 learning)



- F2E + F3 VDC
  - Introduction of 40G/100G into existing 10G environments
  - Migration to larger table sizes
  - Transition to additional features/functionality (OTV, MPLS, VXLAN, etc.)

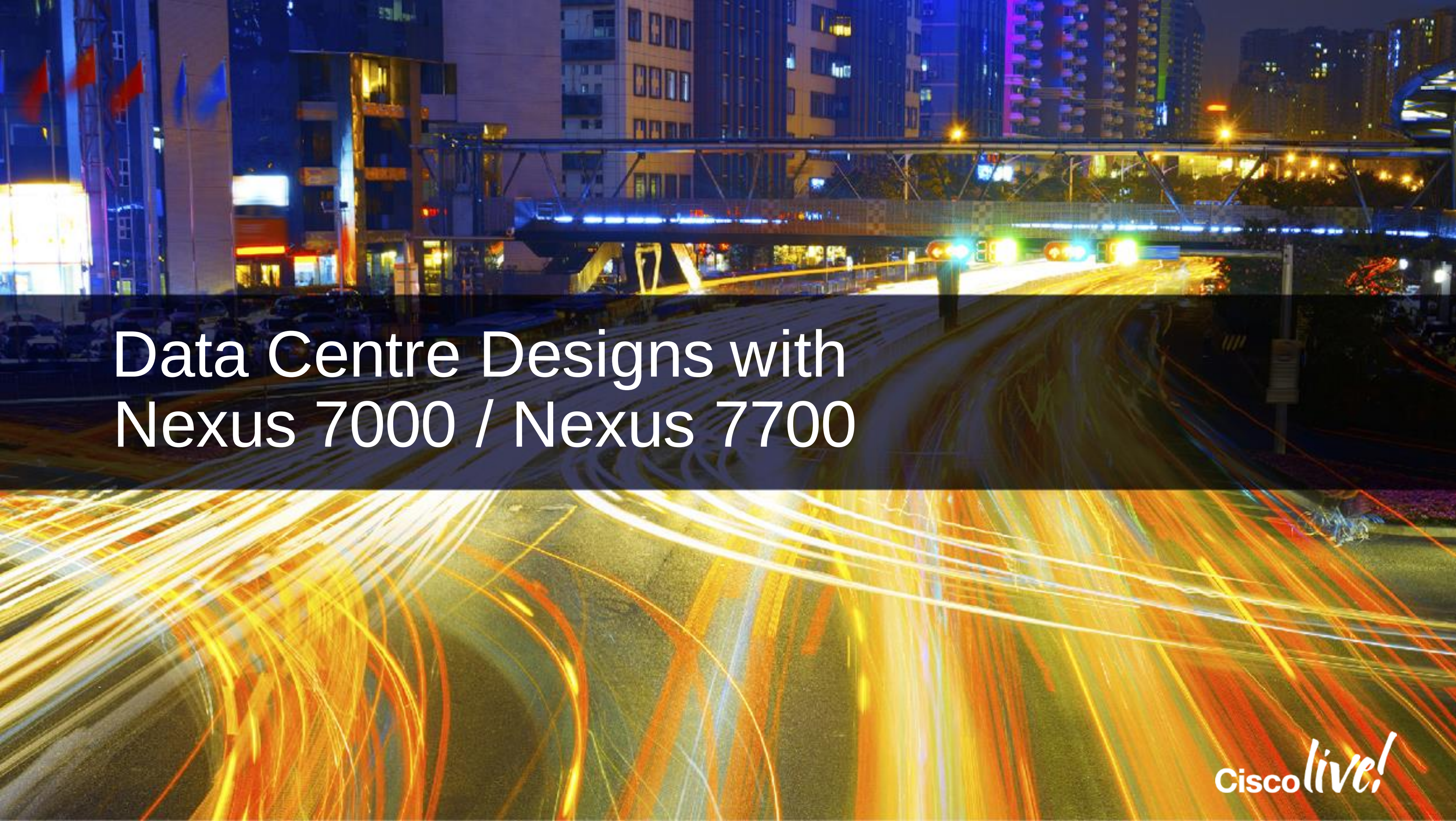


- M2 + F3 VDC
  - Introduce higher 10G/40G/100G port-density while maintaining feature-set
  - Avoid proxy-forwarding model for module interoperability
  - Migrate to 40G/100G interfaces with full-rate flow capability



Agenda

- Introduction to Nexus 7000 / Nexus 7700
- Nexus 7000 / Nexus 7700 Architecture
  - Chassis
  - Supervisor engines and NX-OS software
  - I/O modules (M2/F2E/F3)
- I/O Module Interoperability
- **Data Centre Designs with Nexus 7000 / Nexus 7700**
  - **STP/VPC**
  - **L4-7 services integration**
  - **VDCs**
  - **FabricPath**
  - **VRF/MPLS VPNs**
  - **OTV**
- Next-Generation Data Centres with Nexus 7000 / Nexus 7700
  - Evolved FabricPath
  - ACI integration
  - VXLAN / VXLAN + EVPN



Data Centre Designs with Nexus 7000 / Nexus 7700

Cisco live!

Nexus 7000 / Nexus 7700 Design Building Blocks

Foundational:

- Spanning Tree (RSTP+/MST)
- Virtual Port Channel (VPC)
- FabricPath
- Virtual Routing and Forwarding (VRF) and MPLS VPNs

Innovative:

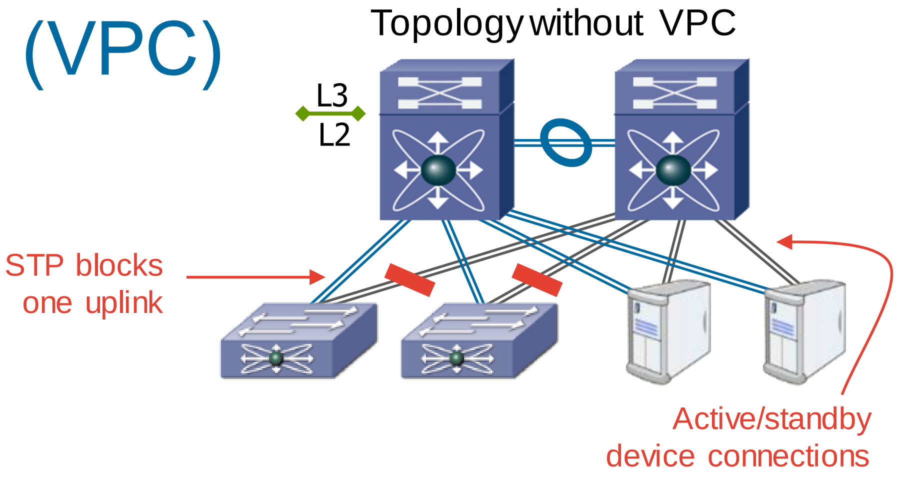
- Remote Integrated Service Engine (RISE)
- Intelligent Traffic Director (ITD)
- Virtual Device Context (VDC)
- Overlay Transport Virtualisation (OTV)

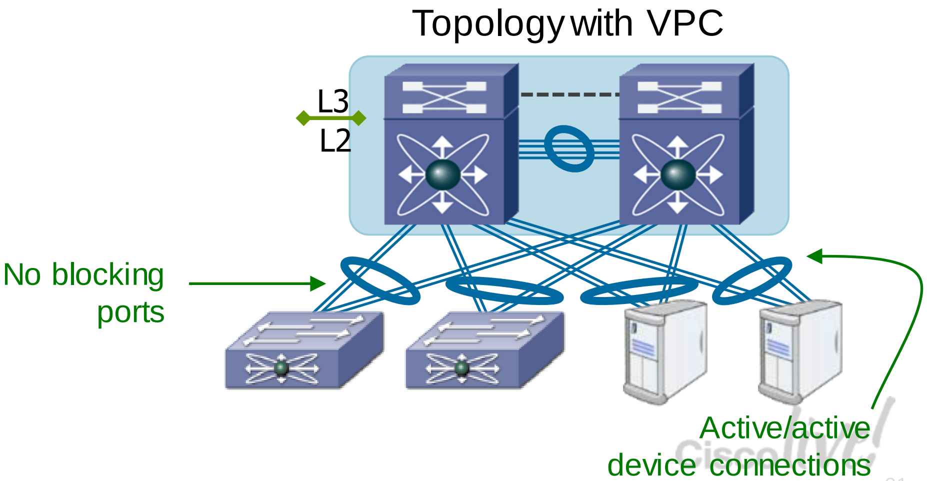
Emerging:

- FabricPath evolution
- ACI integration
- VXLAN flood and learn / VXLAN + EVPN

STP → Virtual Port Channel (VPC)

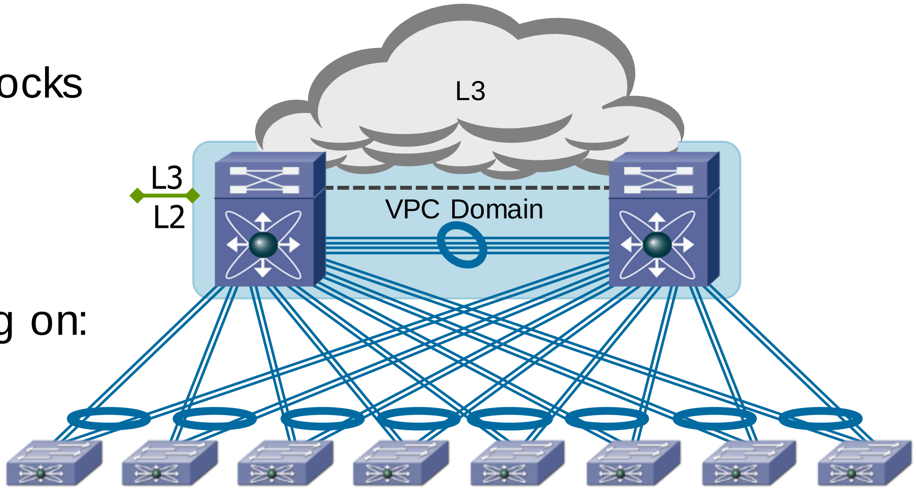
- Most customers have taken this step
- Leverages all available uplink bandwidth by eliminating STP blocked ports
- Eliminates active-standby mode on dual-homed servers
- Provides active-active HSRP
- Works seamlessly with current network designs/topologies
- Simple L4-7 services integration
- Works with any module type (M2/F2E/F3)
- Follow documented best practices for VPC to avoid issues





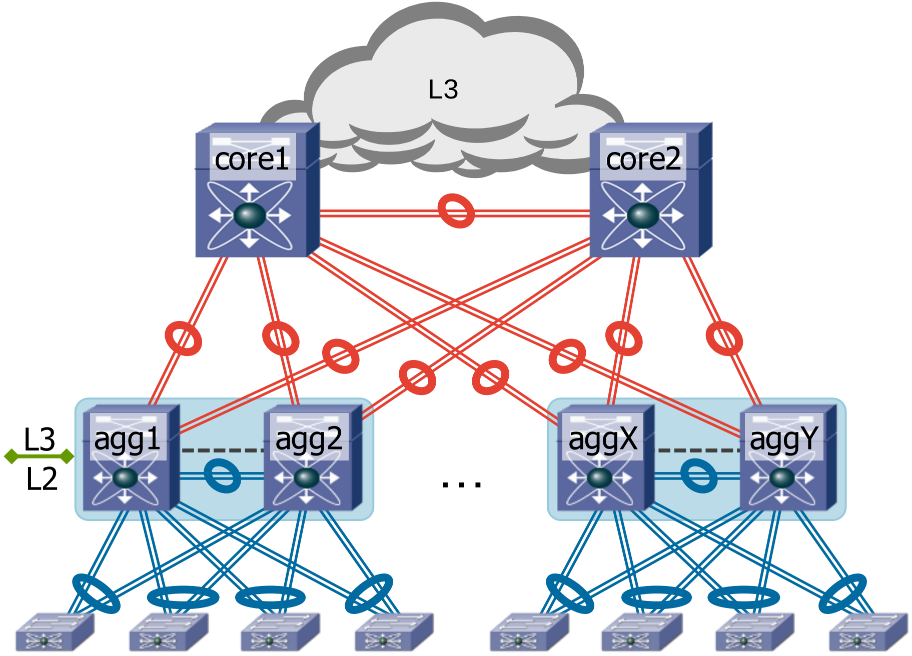
Collapsed Core/Aggregation

- Nexus 7000 / Nexus 7700 as Data Centre collapsed core/aggregation
- Consolidate multiple aggregation building blocks into single switch pair
- Reduce number of managed devices
- Simplify East-West communication path
- M-series or F-series I/O modules, depending on:
  - Port density and feature-set requirements
  - Desired level of oversubscription



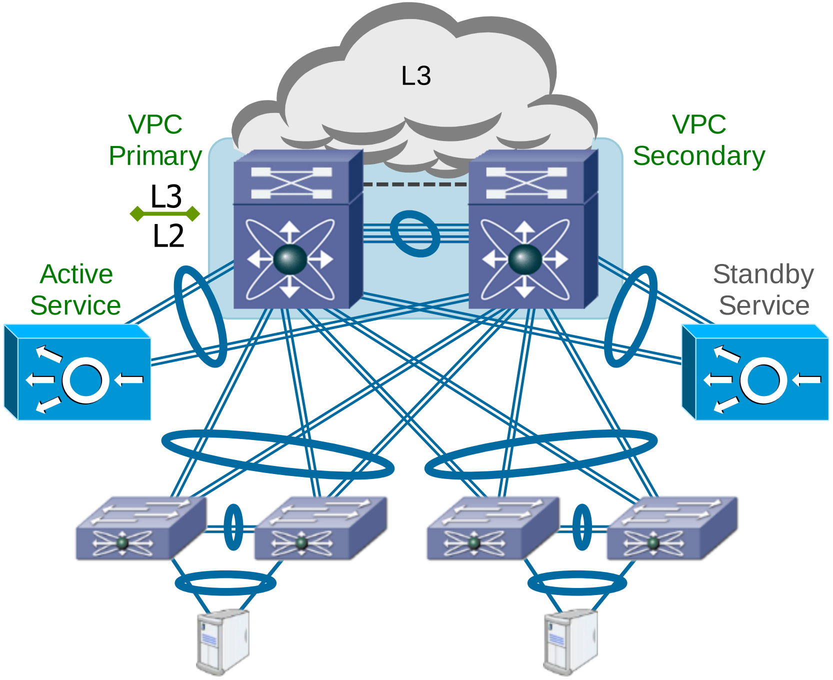
Traditional 3-Tier Hierarchical Design

- Extremely wide customer-deployment footprint
- Nexus 7000 / Nexus 7700 in both Data Centre aggregation and core
  - Provides high-density, high-performance 10G / 40G / 100G
  - Same module-type considerations as collapsed core
- Scales well, but scoping of failure domains imposes some restrictions
  - VLAN extension / workload mobility options limited



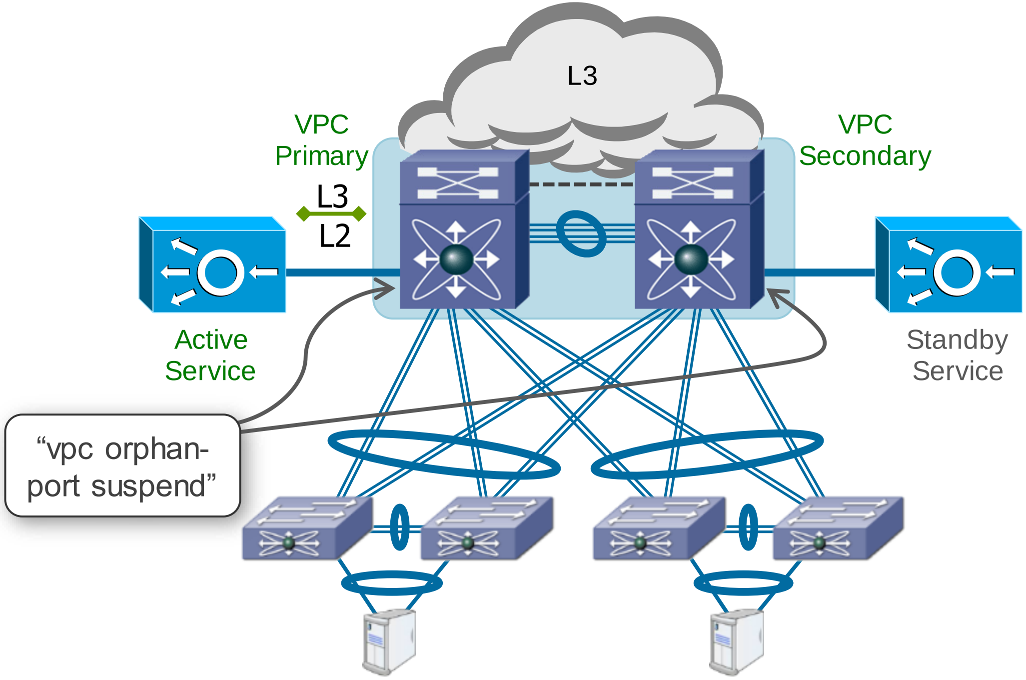
L4-7 Services Integration – VPC Connected

- VPC designs well-suited for L4-7 services integration – pair of aggregation devices makes service appliance connections simple
- Multiple service types possible – transparent services, appliance as gateway, active-standby or active-active models
- VPC-connected appliances preferred:
  - Ensures that all traffic – data plane, fault-tolerance, and management – sent direct via VPC port-channels
  - Minimises VPC peer link utilisation in steady state



L4-7 Services Integration – Orphan Ports

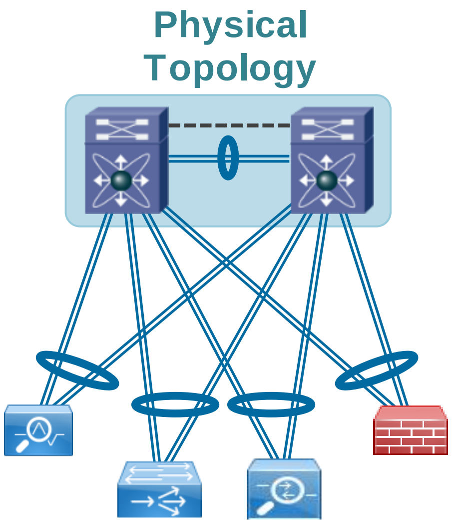
- Sometimes services appliance does not support port-channels
- Use orphan ports with "vpc orphan-port suspend" on appliance-connected interfaces
  - Ensures if VPC peer-link fails, services on VPC secondary taken down as well
- Orphan-port connected appliances means data plane, fault-tolerance, and management traffic may traverse VPC peer-link
  - Be sure to provision peer-link accordingly



L4-7 Services Integration – RISE

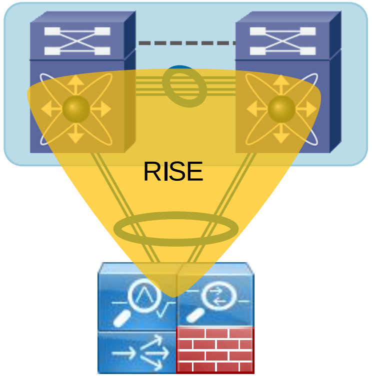
Remote Integrated Service Engine (RISE)

- Logical integration of external services appliance with Nexus 7000 / Nexus 7700
  - Citrix NetScaler and Cisco Prime NAM appliance supported today
- Enables feature integration and data-path acceleration between services appliance and Nexus 7000 / Nexus 7700 switches, including:
  - Discovery and bootstrap
  - Automated Policy Based Routing (APBR)
  - Route Health Injection (RHI) (future)

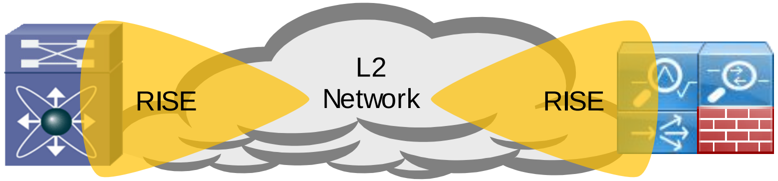


RISE Attach Modes

- Direct Mode – Services appliance directly connected to Nexus 7000 / Nexus 7700

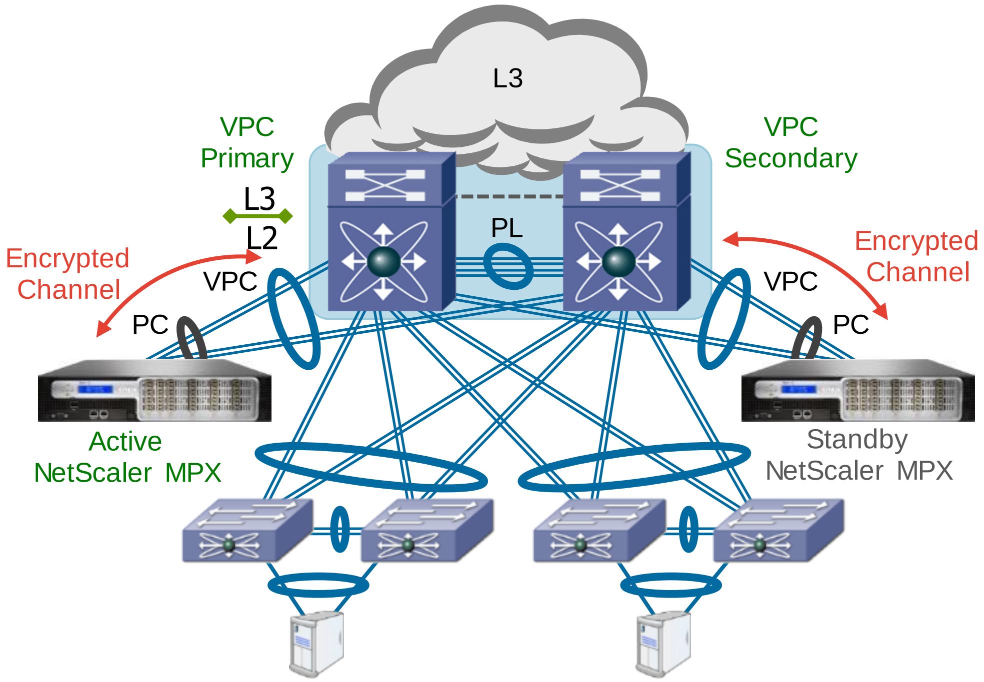


- Indirect Mode – Virtual service appliance (e.g., NetScaler VPX / 1000V) connected via L2 network



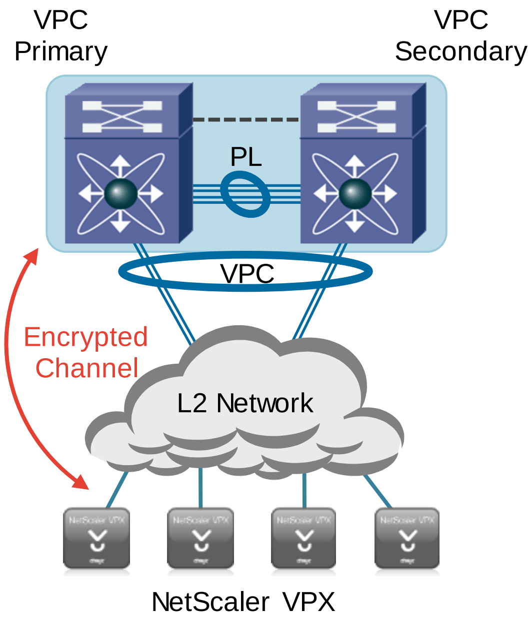
RISE Direct Attached Mode

- Services appliance directly connected to Nexus 7000 / Nexus 7700
- NetScaler appliance can connect to single switch or VPC peer switches
- Encrypted channel for switch ↔ appliance communications
- Supports Auto-Discovery and Bootstrap of NetScaler services appliance
  - Appliance's management IP configured via RISE during auto-discovery process



RISE Indirect Mode

- Services appliance indirectly connected to Nexus 7000 / Nexus 7700 via an L2 network
  - Typical design for virtual services
- Encrypted channel for switch ↔ appliance communications
- Auto-Discovery and Bootstrap not supported in indirect mode
  - User must manually connect to NetScaler to perform initial configuration



RISE – Virtual Services Modules

- **show module service** displays all RISE-attached services appliances, including type, status, version, and serial number

Physical MPX
Virtual VPX

```
n7k1-dclagg1# show module service
Mod  Ports  Module-Type                          Model              Name         Status
---  -----  ------------------------------------ ------------------ ------------ ----------
9    4     Network Analysis Module NAM-NX1      N7K-SM1-NAM-a21    NAM9        ok
332  2      NSMPX-11500 12*CPU+2*E1K+8*E1K+4*IX NetScaler          mpx205a      ok
333  2      NSMPX-11500 12*CPU+2*E1K+8*E1K+4*IX NetScaler          mpx205b      ok
334  0      NetScaler Virtual Appliance          NetScaler          vpx2         ok
335  1      NAM2220                              NAM2220            nam2220      ok

Mod  Sw                                Serial-Num
---  --------------------------------  --------------------------------
9   6.1(1)                            JAF1647AAGL
332  NetScaler NS10.1: Build 124.1308  MH8C02AM50
333  NetScaler NS10.1: Build 124.1308  JT7A22AM9E
334  NetScaler NS10.1: Build 124.1308  HE2H81UJ47
335  6.0(2-patch1)"                    FTX1504563W
```

Physical MPX
Virtual VPX

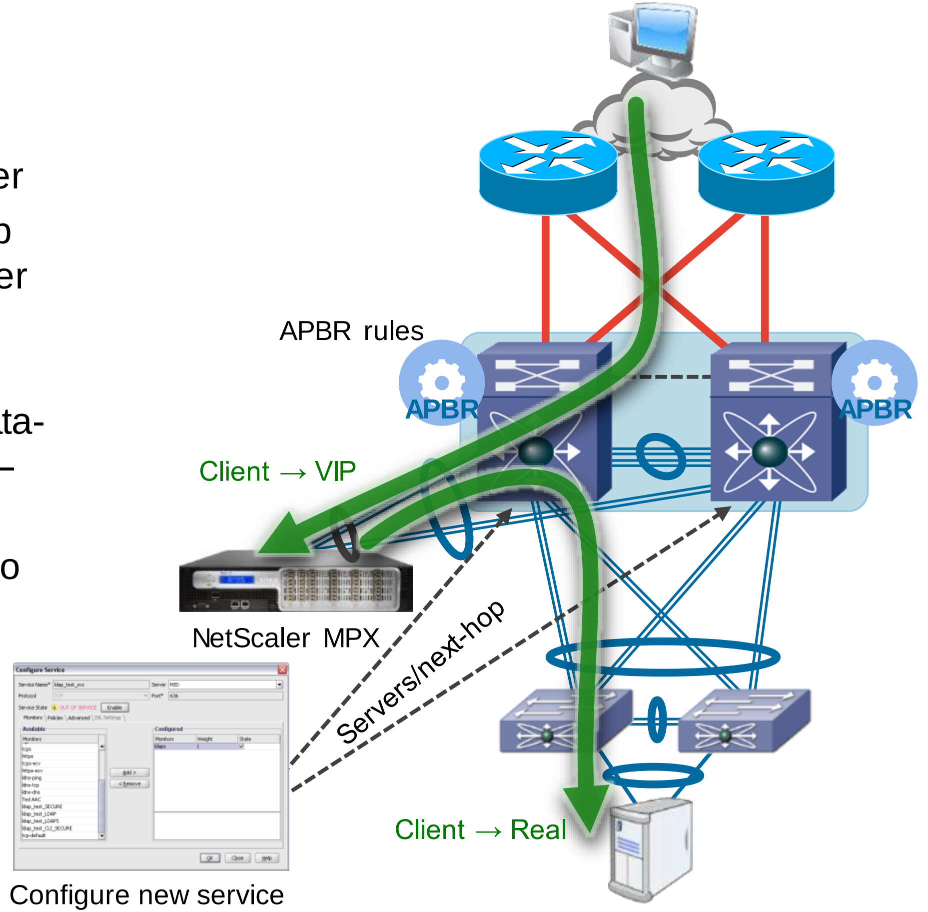
```
n7k1-dcl-agg1(config-rise)# attach rise slot 332
Attaching to RISE 332 ...
Username:nsroot
Warning: Permanently added '10.90.14.216' (RSA) to the list of known hosts.
Password:
Last login: Tue Mar 11 09:15:31 2014 from 10.90.14.138
Copyright (c) 1980, 1983, 1986, 1988, 1990, 1991, 1993, 1994
        The Regents of the University of California.  All rights reserved.

 Done
> show rise profile
1)      Service Name     : mpx205a
        Status           : Active
        Mode             : vPC-Direct
        Device Id        : TBM14257214
        Slot Number      : 332
        VDC Id           : 2
        vPC Id           : 2051
        SUP IP           : 10.90.14.138
        VLAN             : 99
        VLAN Group       : 20
        Interface        : LA/1
        ISSU             : None
```

- Can attach to RISE services appliance directly from switch command line

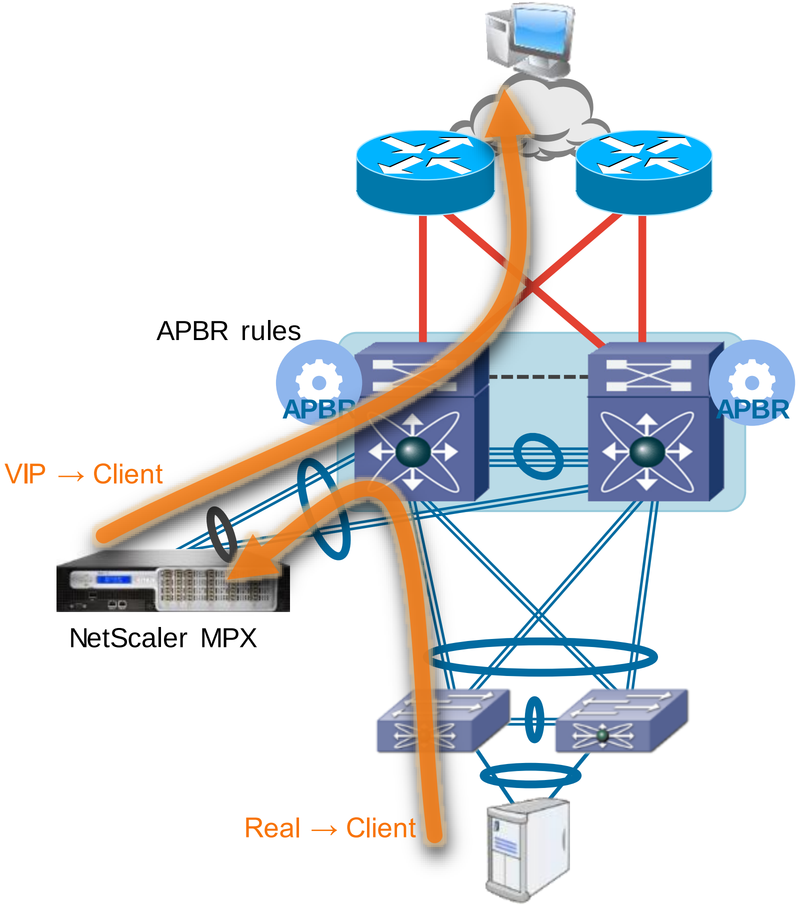
RISE Auto-PBR

- User configures new service in NetScaler
- NetScaler sends server list and next-hop interface to Nexus 7000/7700 switch over RISE control channel
- Switch automatically generates PBR route-maps and applies PBR rules in data-plane hardware to redirect target traffic – no manual configuration on switch
- Client traffic destined to VIP redirected to NetScaler for processing, destination rewritten to Real server IP



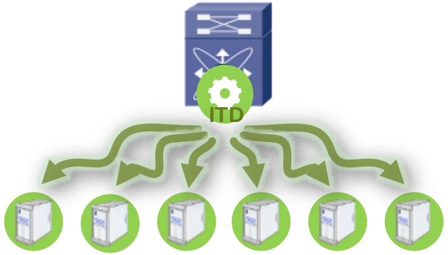
RISE Auto-PBR

- User configures new service in NetScaler
- NetScaler sends server list and next-hop interface to Nexus 7000/7700 switch over RISE control channel
- Switch automatically generates PBR route-maps and applies PBR rules in data-plane hardware to redirect target traffic – no manual configuration on switch
- Client traffic destined to VIP redirected to NetScaler for processing, destination rewritten to Real server IP
- Return traffic redirected to rewrite Real IP to VIP



ITD on Nexus 7000 / Nexus 7700

Intelligent Traffic Director (ITD)

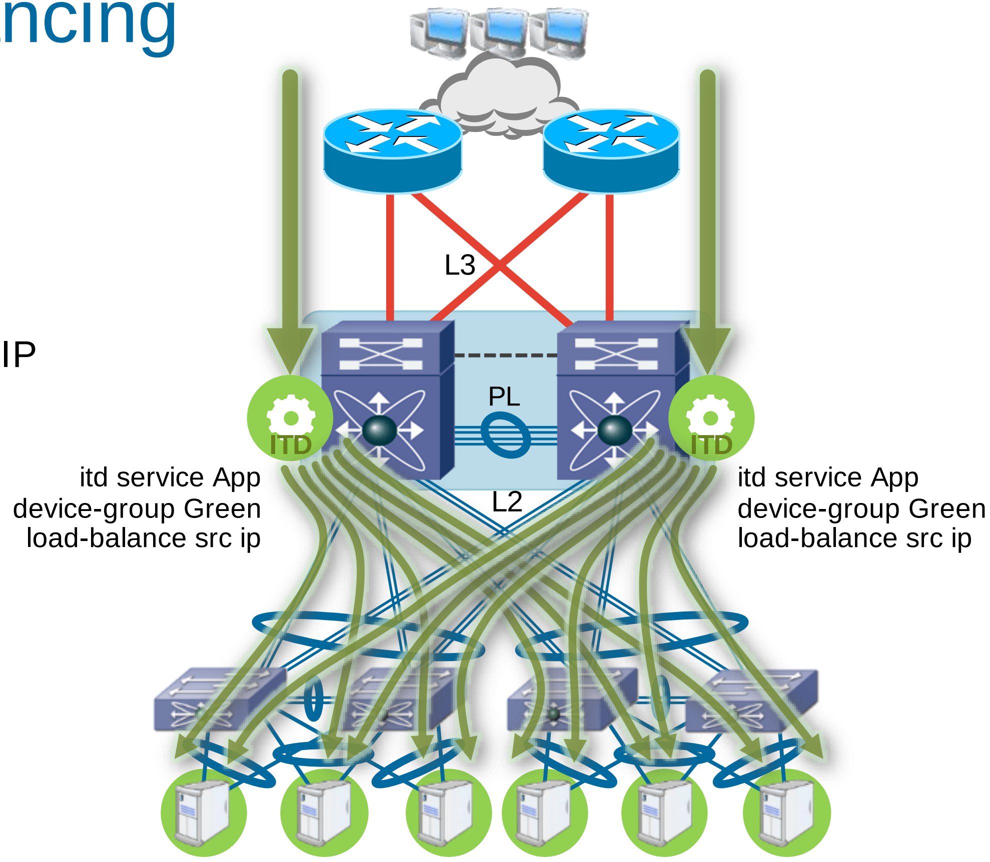


- Hardware-based L3/L4 redirection and weighted load-balancing
- Any Nexus 7000 / Nexus 7700 port can be used for load-balancing
  - No service module or external load-balancer required
  - Available on M2/F2E/F3
- Redirect line-rate traffic to any devices, including firewalls, web caches, WAAS, etc.
  - Servers/appliances do not have to be directly connected
- Supports both IPv4 and IPv6
- Bidirectional flow coherency – Same device receives forward and reverse traffic
- Performs health monitoring/probes and automatic failure handling

<u>Note: ITD is **not** a replacement for L7 load-balancers (URL-based, cookie-based, SSL, etc.)</u>

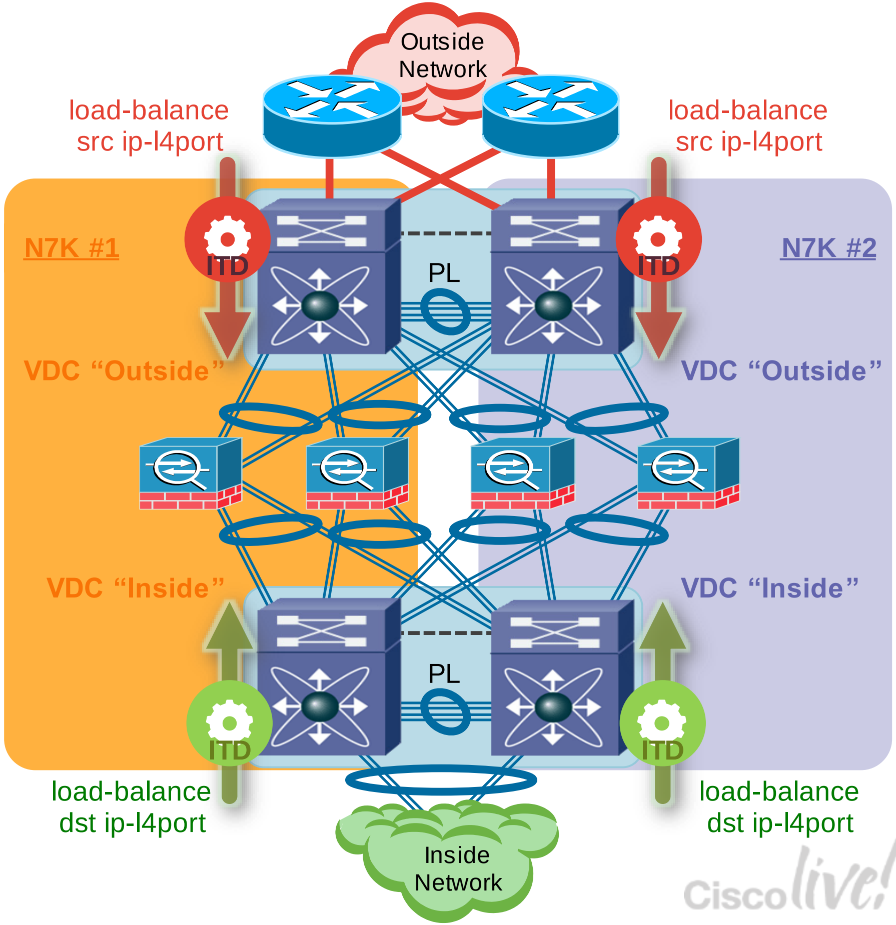
ITD for Server Load Balancing

- Use ITD to load-balance inbound traffic toward cluster of servers
- Typical configuration uses source-based load-balancing to direct flows to target servers
- Option to load-balance based on IP + L4 or just IP addresses
  - L4 option can match on subset of TCP/UDP ports
- Specify VIP, weighting, bucket count (granularity of load-balancing), hot-standby nodes, health probes, etc.



ITD + VDCs for Firewall Load Balancing and Security Domains

- Use VDCs to consolidate hardware while maintaining security-domain separation
- Use ITD to load-balance inbound/outbound traffic through multiple firewall appliances
- Pin bidirectional flows to same firewall device by using source-based load-balancing inbound and destination-based load-balancing outbound
- Option to load-balance based on IP + L4 or just IP addresses
    - L4 option can match on subset of TCP/UDP ports

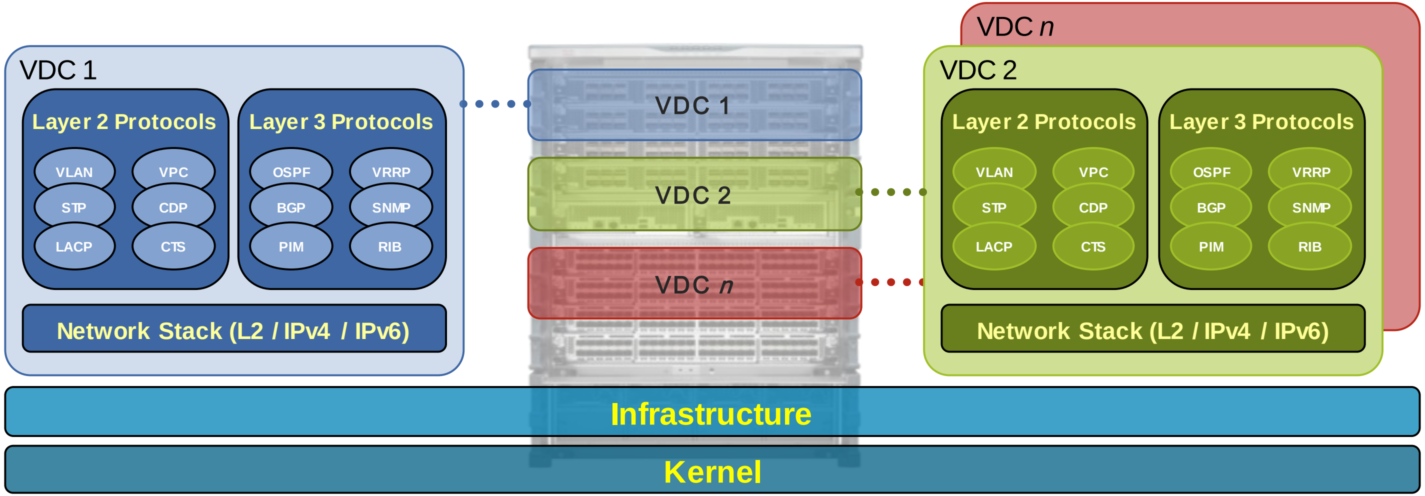


What Are VDCs?

Note: VDCs do *not* provide a hypervisor capability, or ability to run different OS versions in each VDC

Virtual Device Contexts

- Create multiple virtual devices out of one physical device
- Provide data-plane, control-plane, and management-plane separation
- Fault isolation and reduced fate sharing
- Flexible separation / allocation of hardware resources and software components

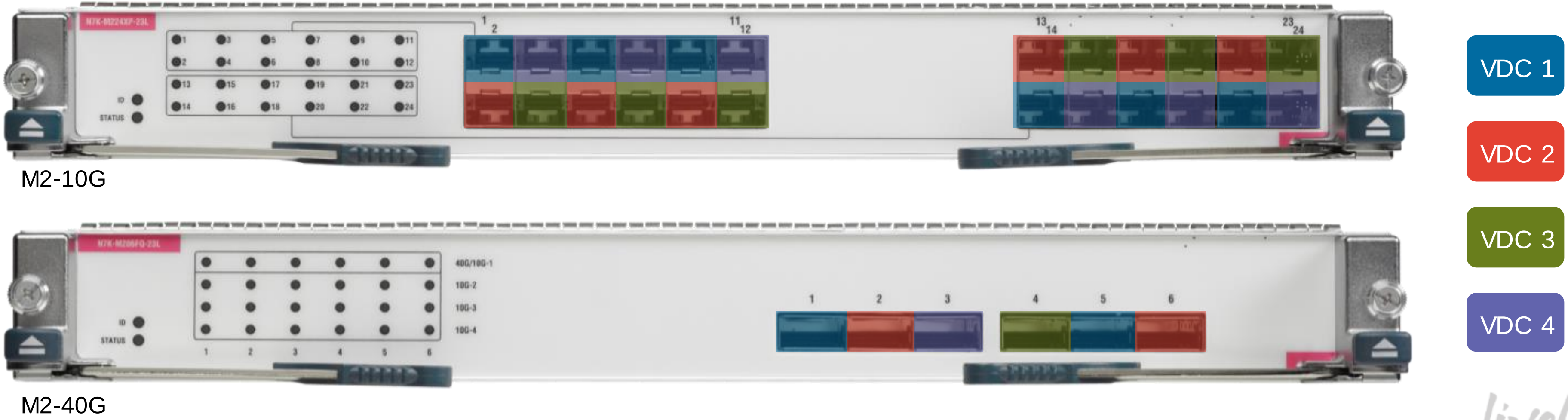


VDC Interface Allocation

- Physical interfaces assigned on per VDC basis, from default/admin VDC
- A single interface cannot be shared across multiple VDCs
- All subsequent interface configuration performed within the assigned VDC
- VDC type (“limit-resource module-type”) determines types of interfaces allowed in VDC
- VDC type driven by operational goals and/or hardware restrictions, e.g.:
  - Mix M2 and F2E in same VDC to increase MAC scale in FabricPath
  - Restrict VDC to F3 only to avoid lowest common denominator
  - Cannot mix M1 and F3 in same VDC

VDC Interface Allocation – M2

- Allocate any interface to any VDC
- But, be aware of shared hardware resources – backend ASICs may be shared by several VDCs
- Best practice: allocate entire module to one VDC to minimise shared hardware resources



VDC Interface Allocation – F2E / F3 Modules

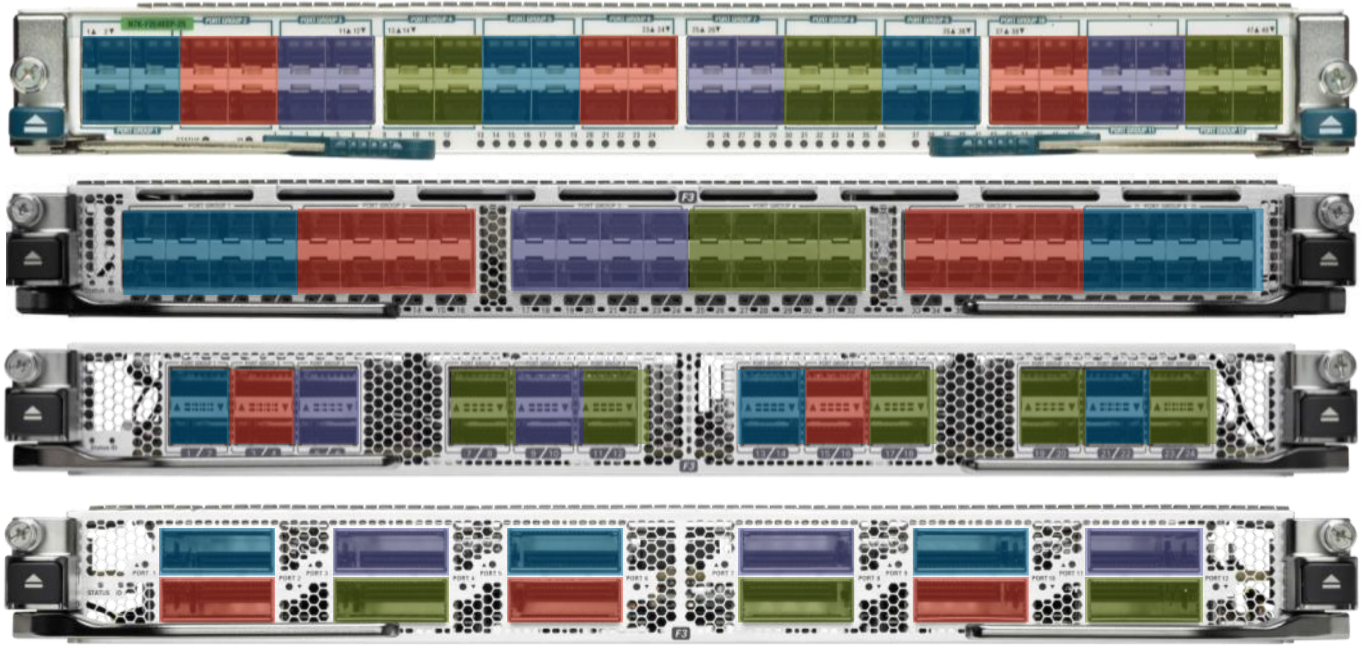
- Allocation on port-group boundaries – aligns ASIC resources to VDCs
- Port-group size varies depending on module type

F2E
4-port
port-group

F3-10G
8-port
port-group

F3-40G
2-port
port-group

F3-100G
1-port
port-group



VDC 1

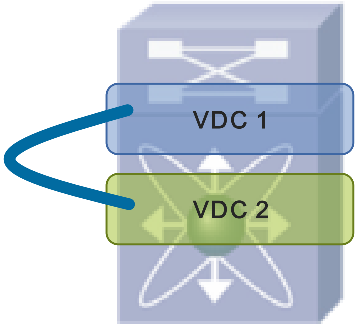
VDC 2

VDC 3

VDC 4

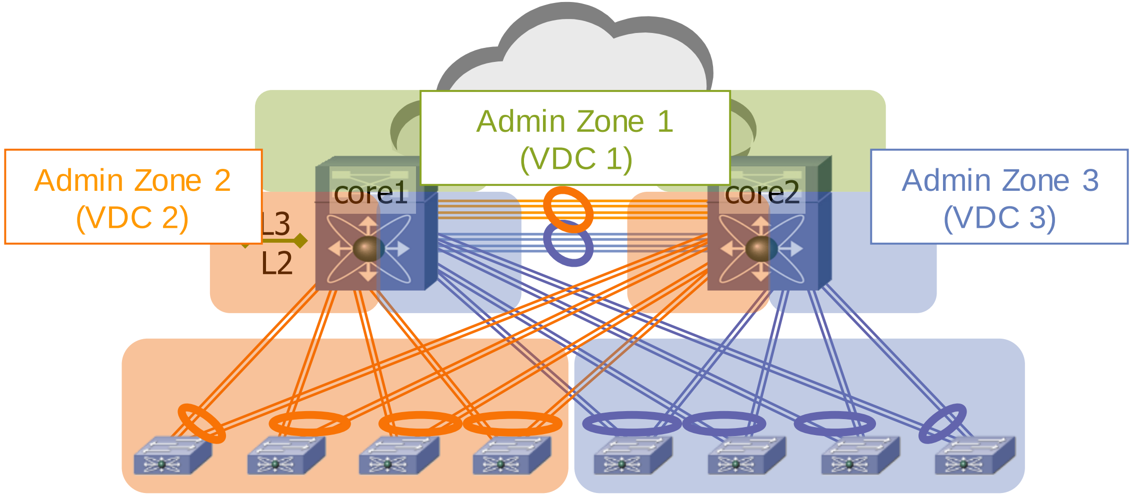
Communicating Between VDCs

- **Must** use front-panel ports to communicate between VDCs
  - No backplane inter-VDC communication
- No restrictions on L2/L3 configuration, module types, or physical media type – just like interconnecting two physical switches
  - Copper Twinax cables (CX-1) or 40G bidi optics provide low-cost interconnect options



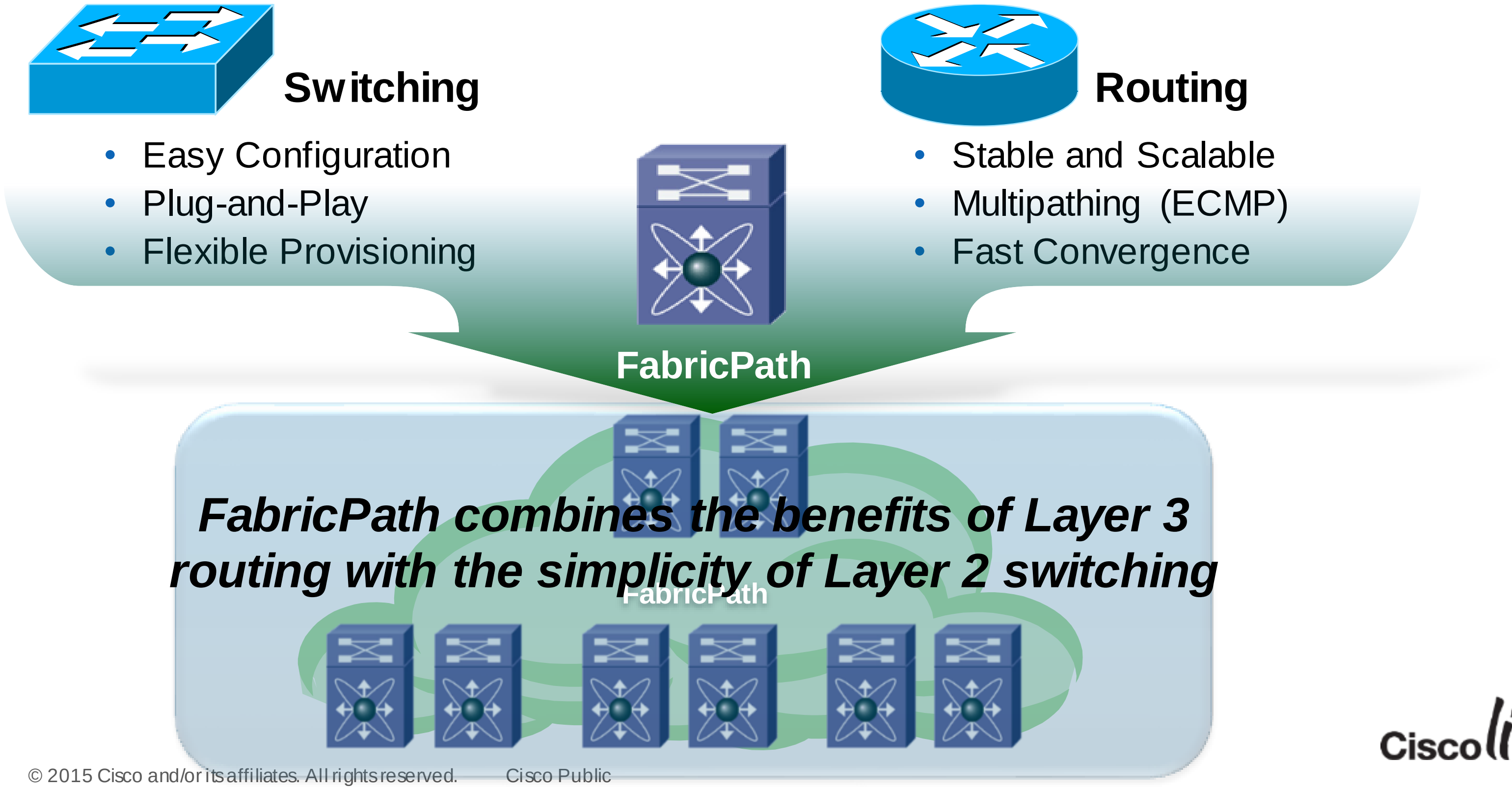
Collapsed Core Design with VDCs

- Maintain administrative segmentation while consolidating network infrastructure
- Maintain fault isolation between zones (independent L2, routing processes per zone)
- Firewalling between zones facilitated by VDC port membership model

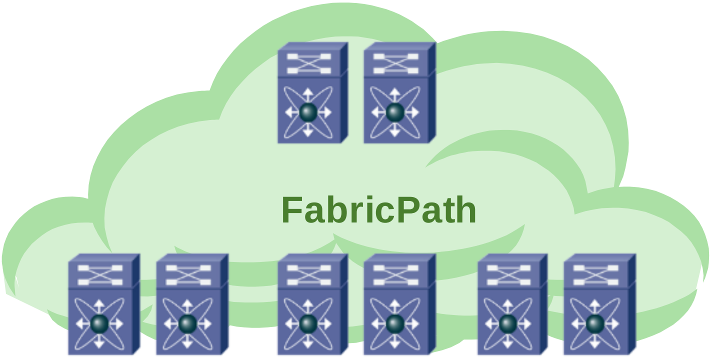


Scalable Layer 2 Fabrics – Cisco FabricPath

Migration to Data Centre Fabric Designs



Why FabricPath?

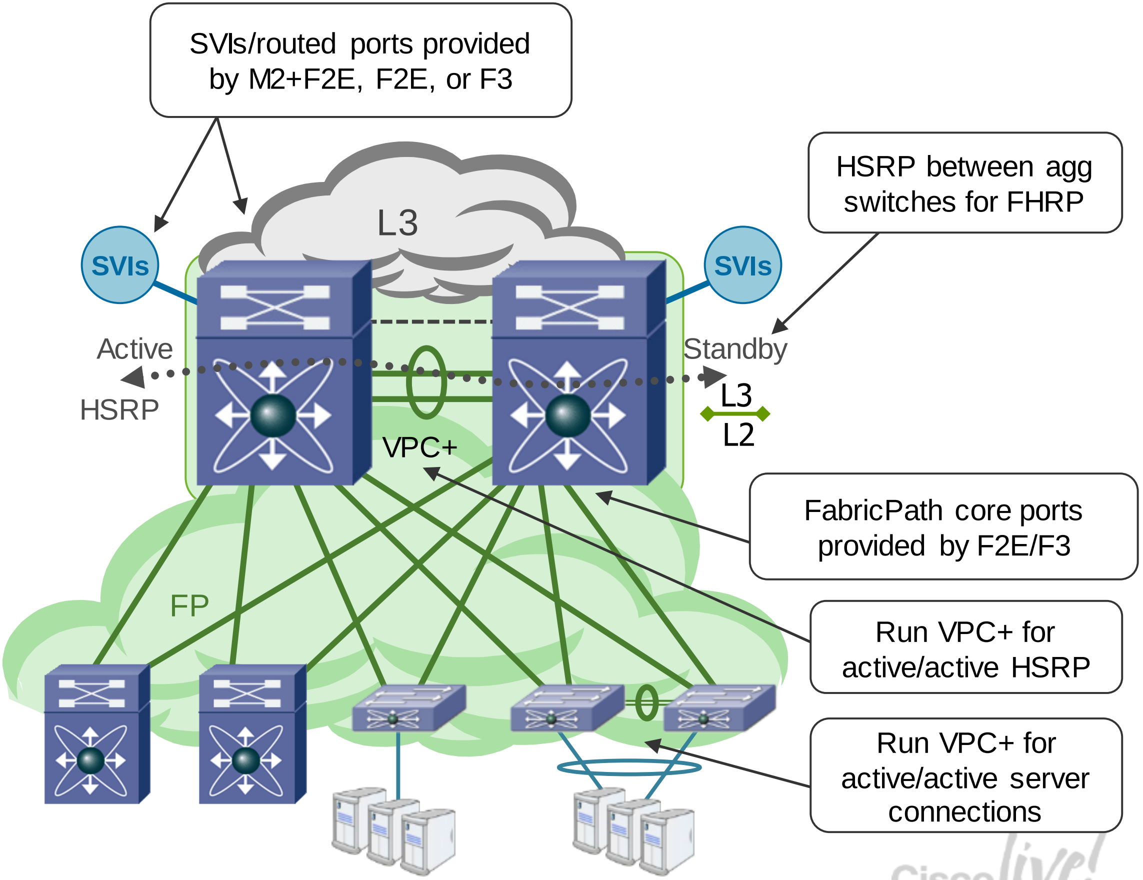


- Reduction / elimination of Spanning-Tree Protocol (STP)
- Better stability and convergence characteristics
- Simplified configuration
- Leverage parallel paths at Layer 2 (ECMP)
- Deterministic throughput and latency using typical designs
- “VLAN anywhere” – flexibility, L2 adjacency, and VM mobility
- Supports legacy/non-IP applications and protocols
- Wide customer-deployment footprint

FabricPath and VPC+

Two-Spine L2 Fabric Design

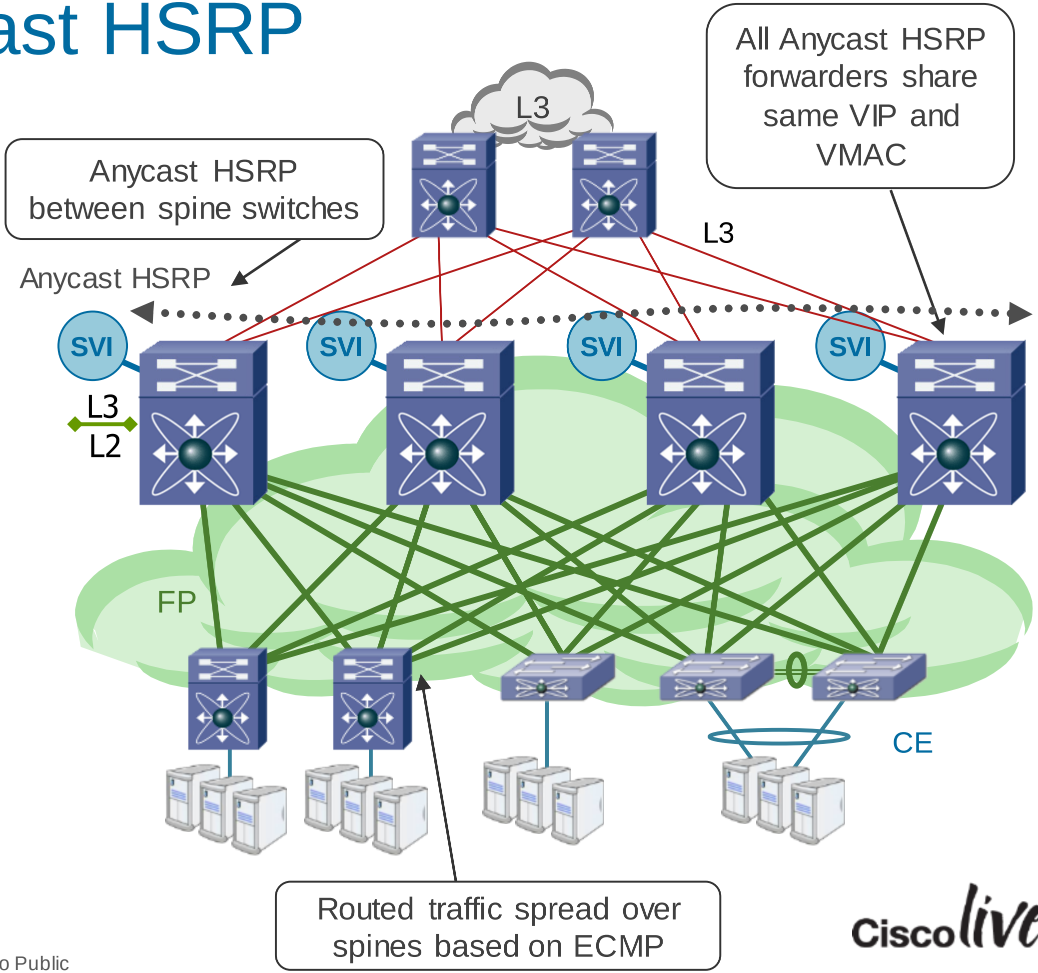
- Simplest FabricPath design option – Extension of traditional aggregation/access designs
- Provides immediate benefits:
  - Removal of STP
  - Active/active gateways
  - "VLAN anywhere" at access layer
  - Topological flexibility
- Positions network for emerging technologies and topologies



FabricPath with Anycast HSRP

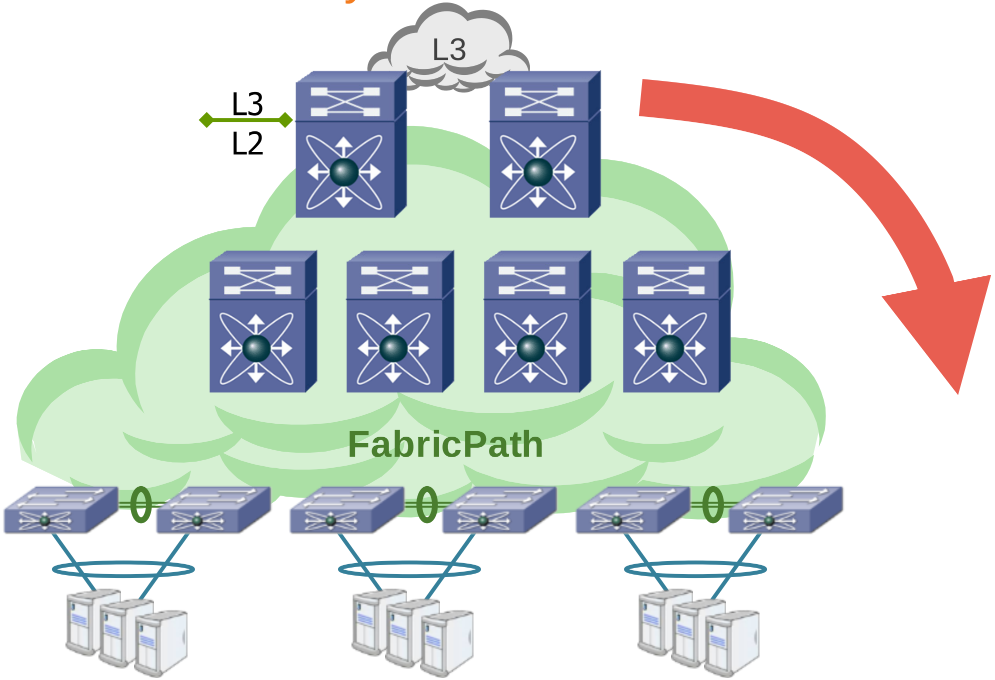
Four-Spine L2 Fabric Design

- Extends existing L2 fabric design – expands spine layer
- Moves topology toward modern DC-fabric trends
- Increases bisectional bandwidth within fabric
- Decreases failure impact
- Increases deployment flexibility



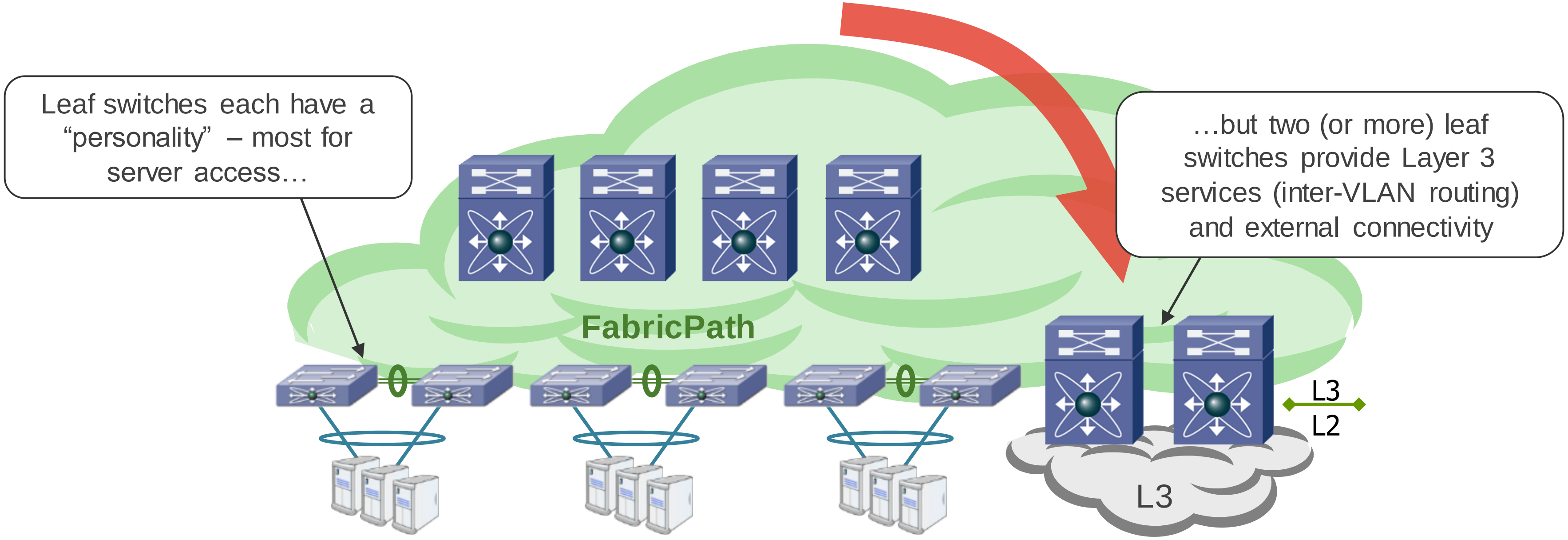
FabricPath Layer 3 Functions

Where does the L2/L3 boundary sit in a FabricPath network?



FabricPath Layer 3 Functions

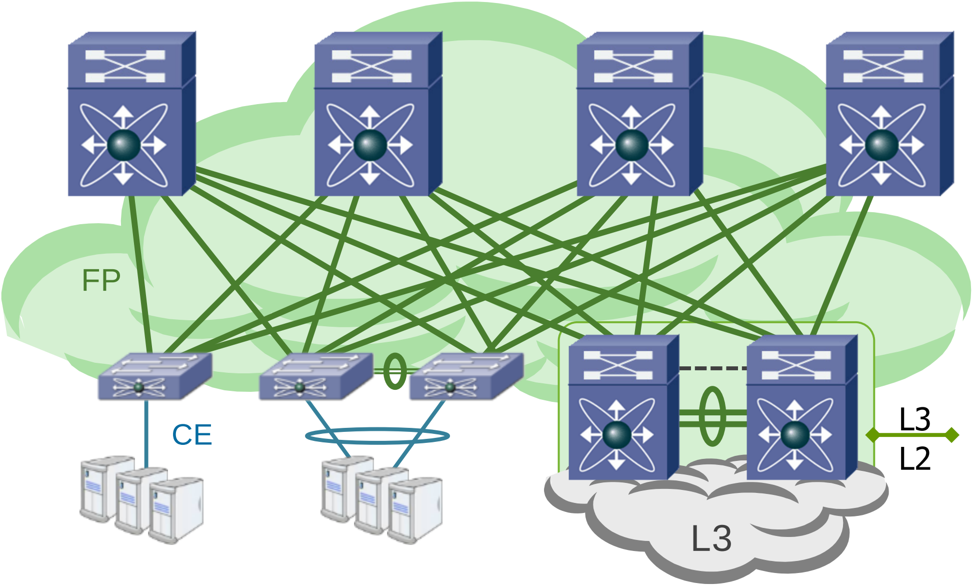
Alternative View – "Border Leaf"



FabricPath with Border Leaf

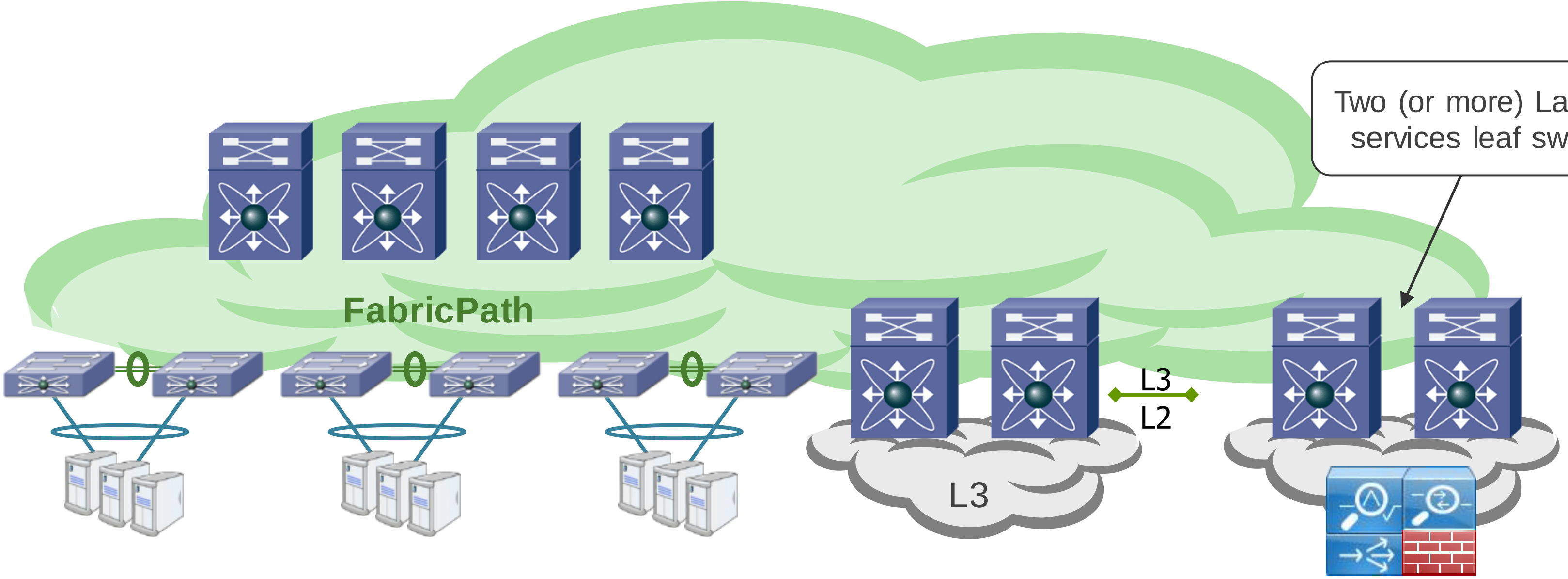
Pure Spine/Leaf Fabric

- Paradigm shift with respect to typical designs – Traditional “aggregation” layer becomes pure FabricPath spine
- Provides uniform any-to-any connectivity between leaf switches
- Two or more leaf switches provide L2/L3 boundary, inter-VLAN routing and North ↔ South routing
- Separates interconnection function from routing function



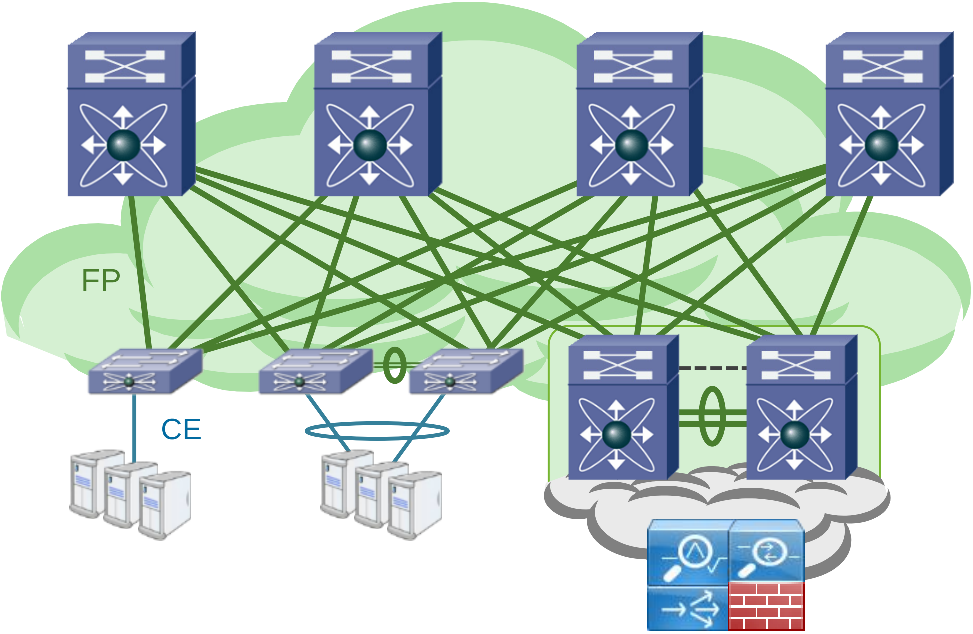
FabricPath L4-7 Services

"L4-7 Services Leaf"



FabricPath with Services Leaf

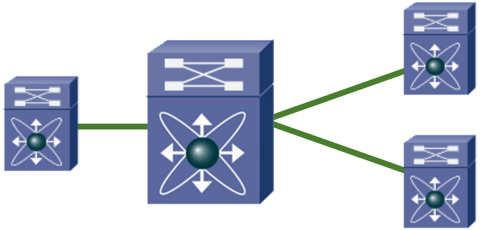
- Services leaf pair simplifies L4-7 services attachment to fabric
- Can leverage same designs and technologies as for traditional networks – VPC+, RISE, ITD, etc.
- Deploy as many services leaf pairs as necessary/desired
- Could be co-located with L3 services leaf switches



FabricPath Multi-Topology

- Traffic engineering for FabricPath VLANs
- Extend some VLANs DC-wide, limit others to subset of physical topology
- One VLAN belongs to one and only one FabricPath topology
- FabricPath core ports always belong to default topology, may belong to as many other topologies as desired
- SPF run for each topology – routing occurs per topology

Configure FabricPath core ports

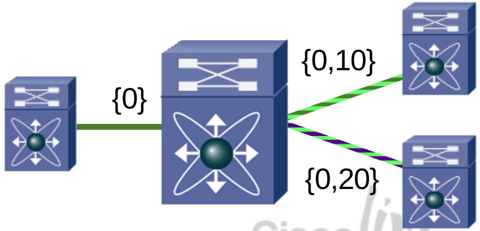


Create additional FP topologies

Map VLANs to topologies

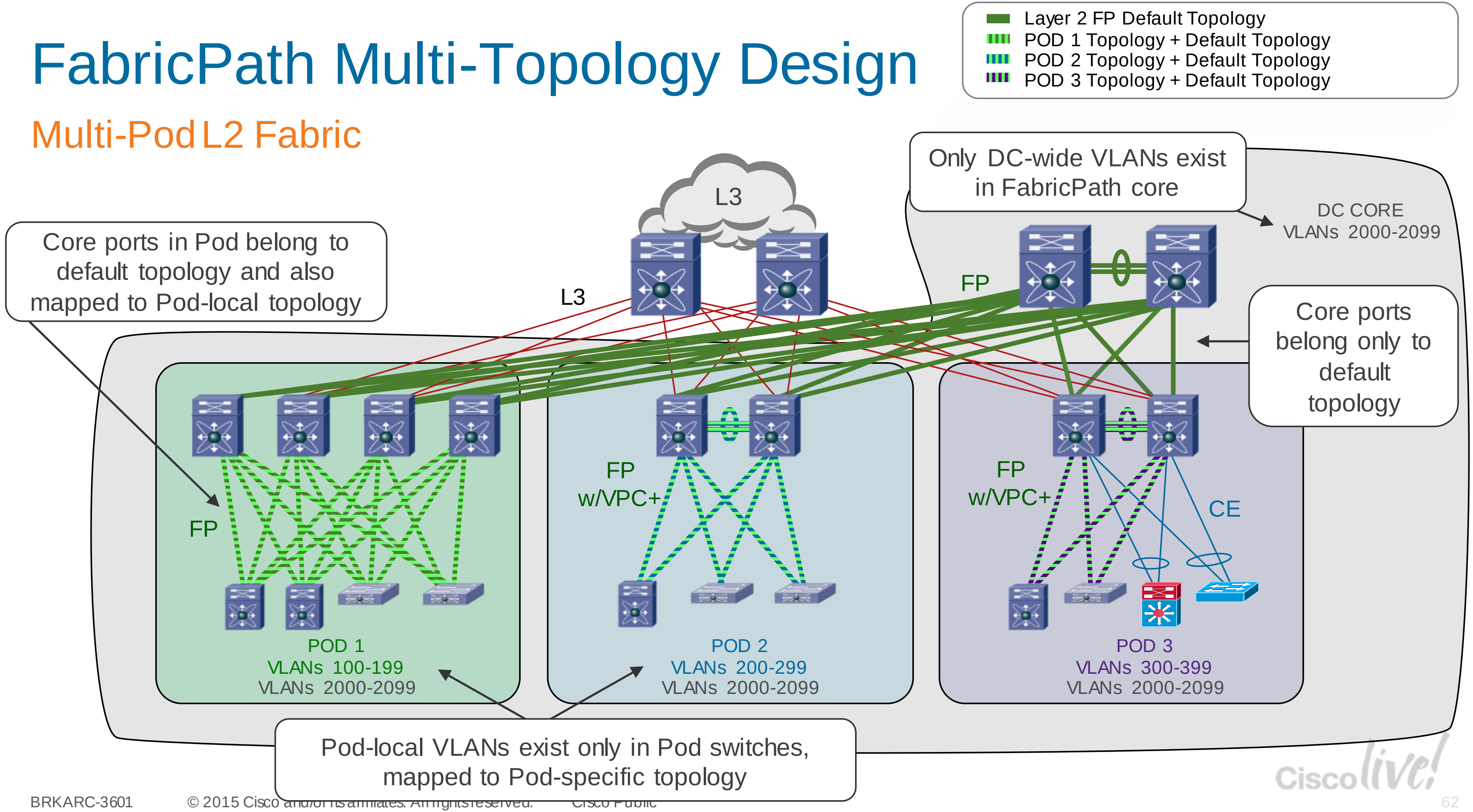
Topology	VLANs
0 (default)	1-100, 301-4095
10	101-200
20	201-300

Map topologies to interface(s)



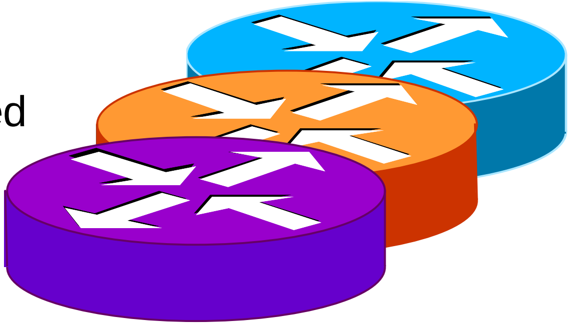
FabricPath Multi-Topology Design

Multi-Pod L2 Fabric



VRF / MPLS VPNs

- Provides network virtualisation – One physical network supporting multiple virtual networks
  - While maintaining security/segmentation and access to shared services
- VRF-lite segmentation for simple/limited virtualisation environments
- MPLS L3VPN for larger-scale, more flexible deployments



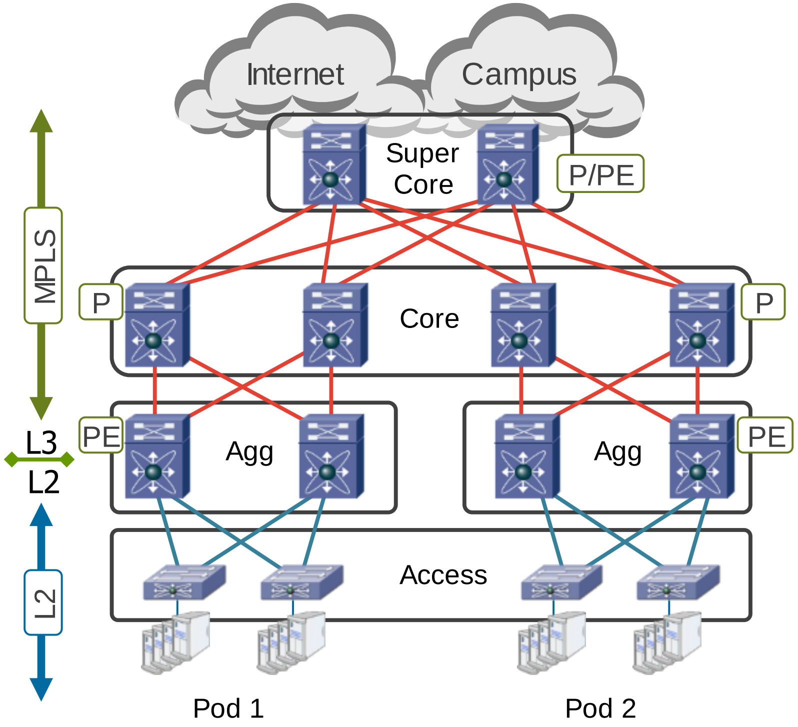
MPLS Layer 3 VPN – Secure Multi-Tenant Data Centre

Requirement:

- Secure segmentation for hosted / enterprise data centre

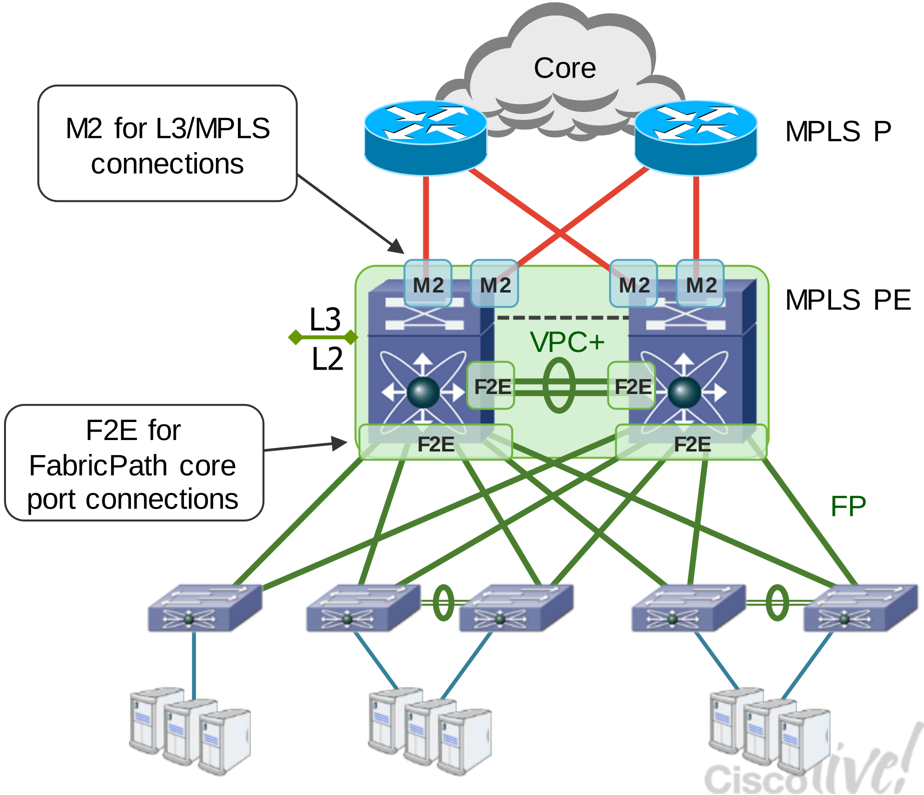
Solution:

- MPLS network infrastructure for all services
- MPLS PE boundary in Pod aggregation layer
- Direct PE-PE or PE-P-PE networks
- L2 or L3 below MPLS boundary (VRF-lite with PE-CE)



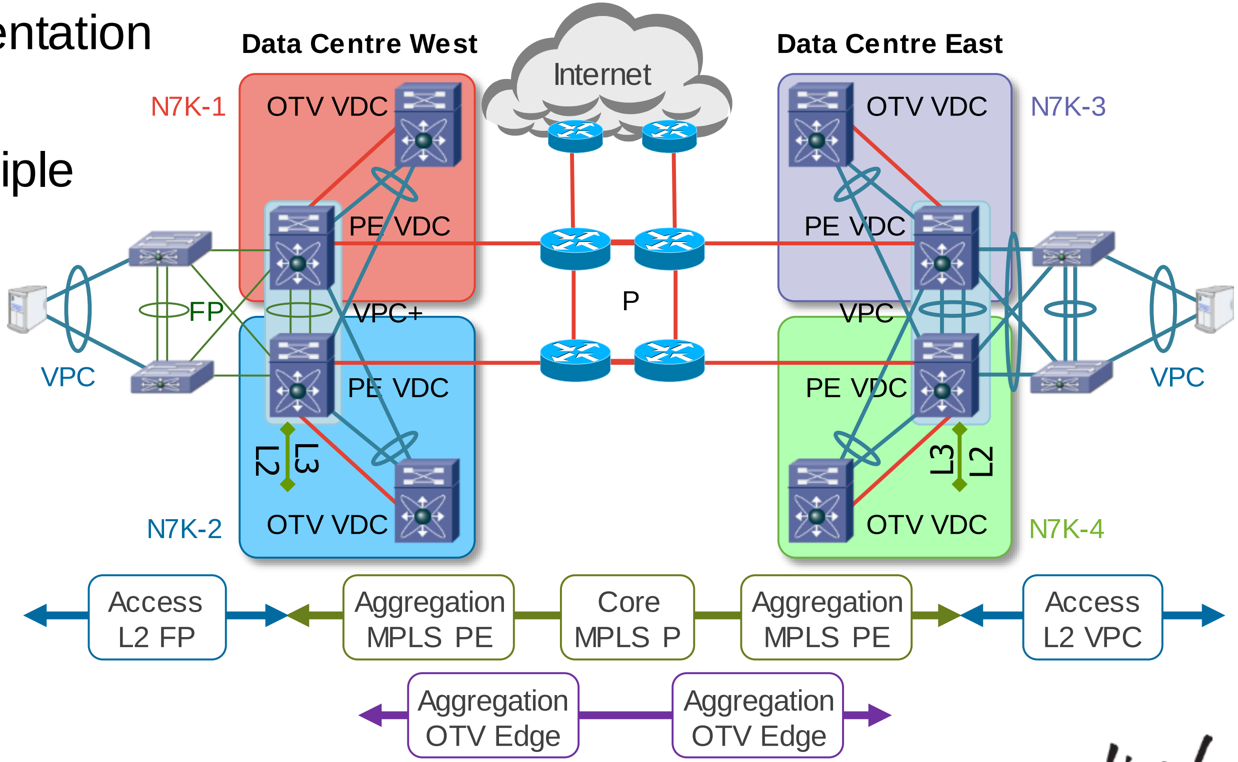
MPLS + FabricPath for Multi-tenancy and VLAN Anywhere

- Mixed M2+F2E system provides FabricPath "south" and MPLS L3VPNs "north"
- F2E modules provide FabricPath for STP-free Layer 2 with flexible VLAN provisioning
- VPC+ provides active-active HSRP into L2 network
- M2 modules provide VRF membership on SVIs and MPLS L3VPNs for multi-tenancy on shared infrastructure
- MPLS functions at L2/L3 spine or on border leaf



Multi-Tenant Data Centre with MPLS + OTV + VDC

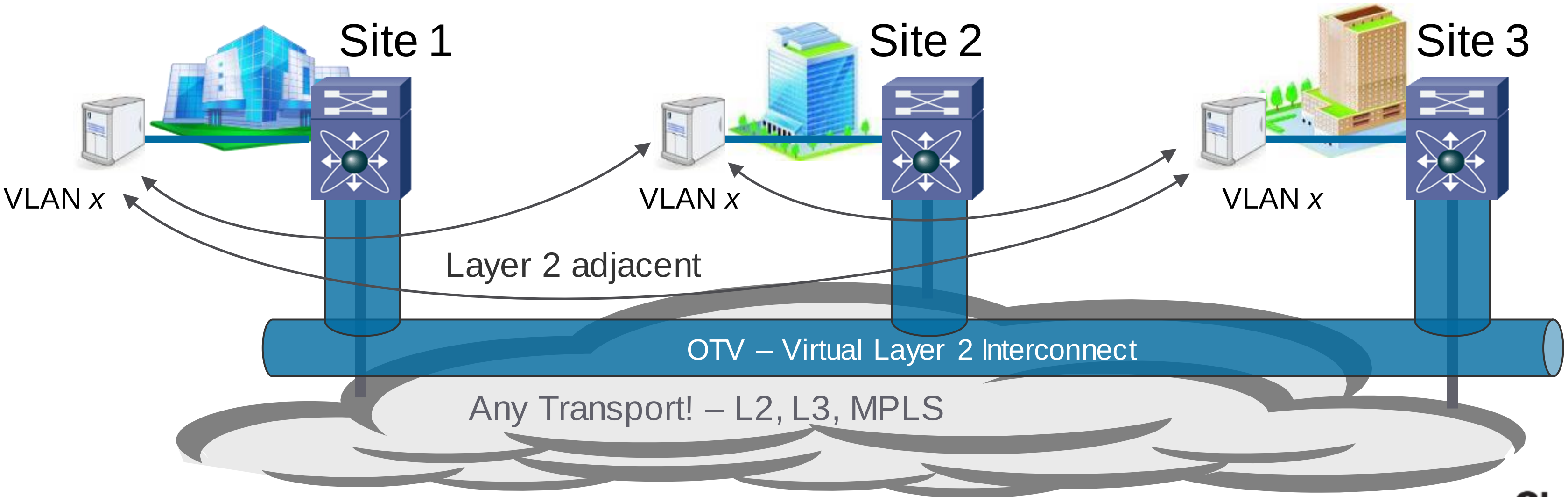
- MPLS for L3VPNs/segmentation
- OTV for multi-site DCI
- VDCs to consolidate multiple network functions (L2/L3 gateways, MPLS PE, and OTV edges) on same physical infrastructure



What Is OTV?

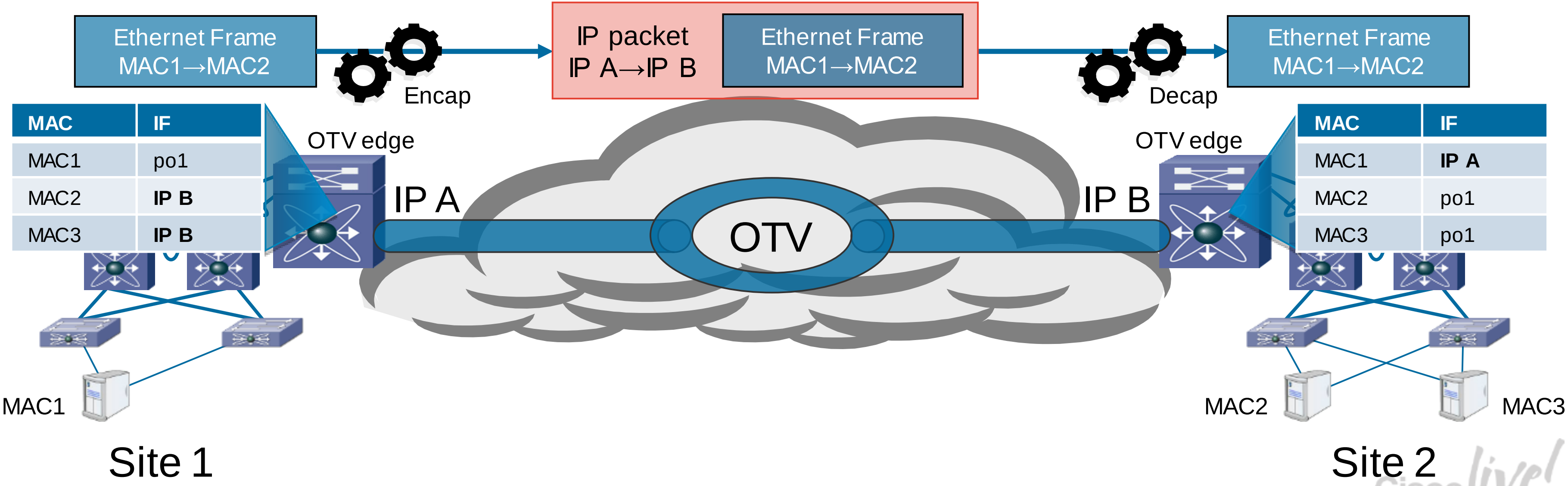
Overlay Transport Virtualisation (OTV)

- Provides multi-site Layer 2 Data Centre Interconnect (DCI)
- Dynamic "MAC in IP" encapsulation with forwarding based on MAC "routing" table
- No pseudo-wire or tunnel state maintained



OTV at a Glance

- MAC addresses advertised in routing protocol (control plane learning) between Data Centre sites
- Ethernet traffic between sites encapsulated in IP: “MAC in IP”

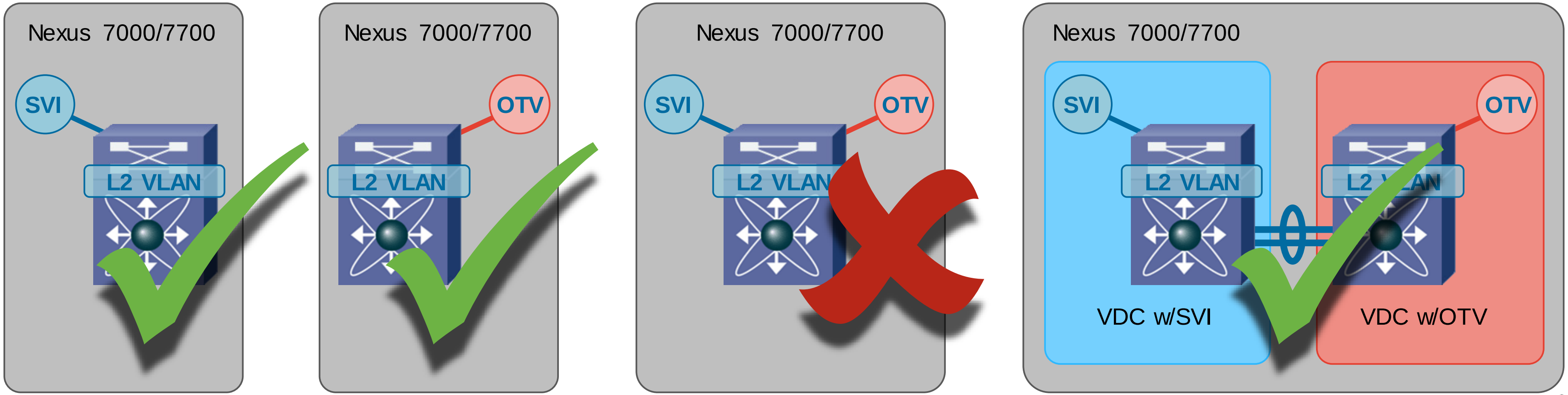


MAC	IF
MAC1	po1
MAC2	**IP B**
MAC3	**IP B**

MAC	IF
MAC1	**IP A**
MAC2	po1
MAC3	po1

OTV VDC Requirement

- Current limitation – SVI (for VLAN termination at L3) and OTV overlay interface (for VLAN extension over OTV) cannot exist in same VDC
- Typical designs move OTV to separate VDC, or separate switch (e.g. Nexus 7004)



Agenda

- Introduction to Nexus 7000 / Nexus 7700
- Nexus 7000 / Nexus 7700 Architecture
  - Chassis
  - Supervisor engines and NX-OS software
  - I/O modules (M2/F2E/F3)
- I/O Module Interoperability
- Data Centre Designs with Nexus 7000 / Nexus 7700
  - STP/VPC
  - L4-7 services integration
  - VDCs
  - FabricPath
  - VRF/MPLS VPNs
  - OTV
- **Next-Generation Data Centres with Nexus 7000 / Nexus 7700**
  - **Evolved FabricPath**
  - **ACI integration**
  - **VXLAN / VXLAN + EVPN**

Next-Generation Data Centres with Nexus 7000 / Nexus 7700

Next-Generation Data Centre Building Blocks

- Evolved FabricPath-based networks
  - Introduce BGP-based host- and subnet-route learning
  - Distribute L3 gateway function to leaf layer
  - Central point of fabric management

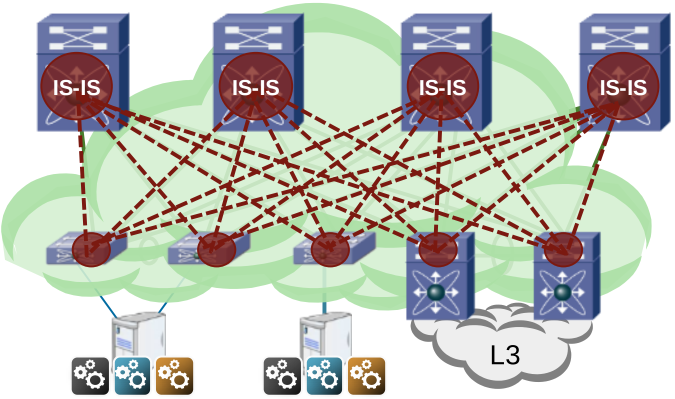
- Application-Centric Infrastructure (ACI)
  - Policy-based fabric management
  - Holistic application deployment and management model – application / compute / network / services

- VXLAN and VXLAN + EVPN
  - “Standards-based” DC fabrics with flexible overlay
  - Multi-tenancy, workload mobility, integration of physical and virtual

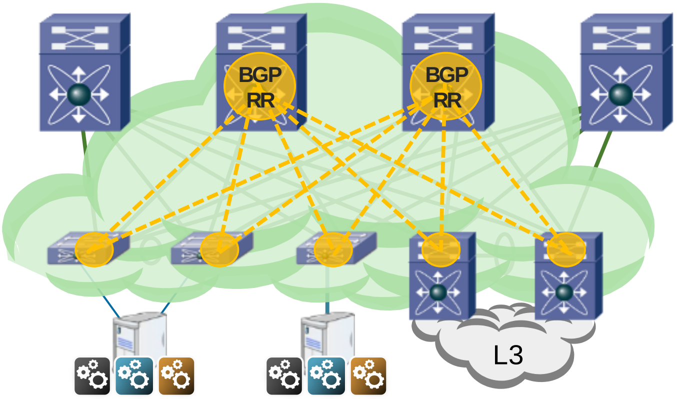
Evolved FabricPath-Based Networks

Fabric and Host / Route Reachability

- FabricPath IS-IS for fabric-node reachability and multidestination tree construction
- FabricPath encapsulation at the data plane



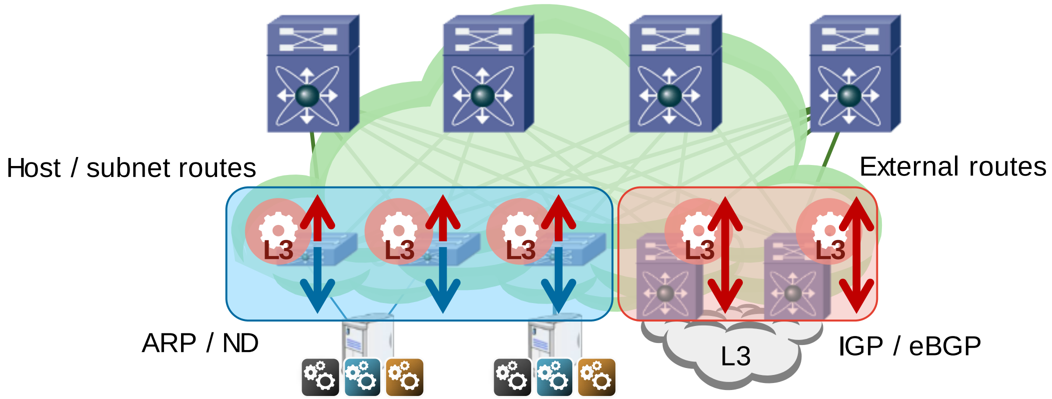
- MP-BGP for host- and subnet-route distribution (VPNv4/v6 address family)
- Route reflectors (RR) for scalability
  - Reduces number of iBGP peering sessions



Evolved FabricPath-Based Networks

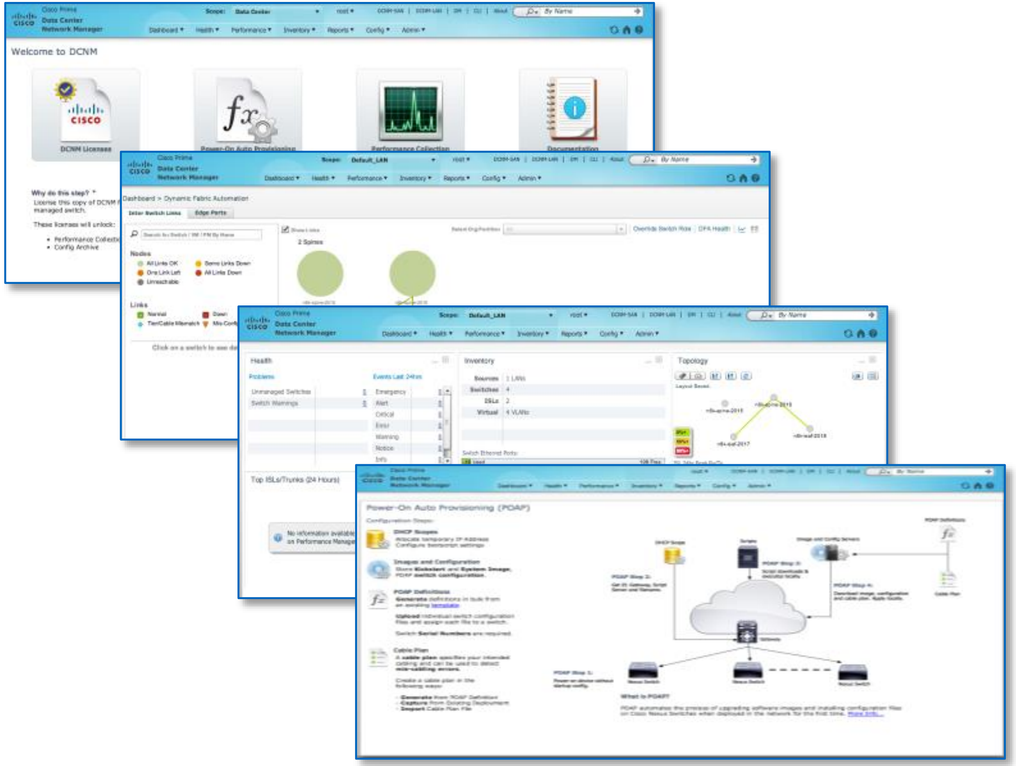
Distributed Gateway

- All leaf switches share gateway IP and MAC in each subnet
  - Any subnet anywhere – Any leaf can instantiate any subnet
  - No HSRP
- ARP/ND terminated on leaf switch, no flooding of IP traffic



Evolved FabricPath-Based Networks

Central Point of Fabric Management

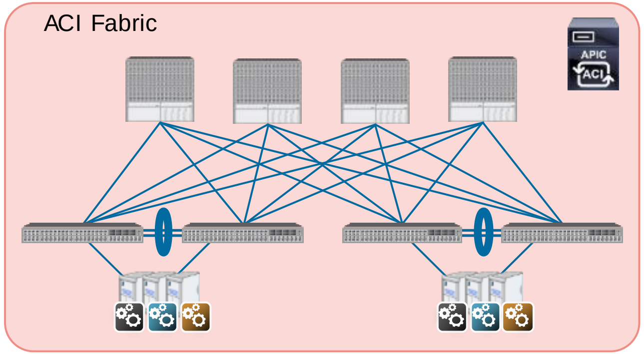


DCNM 7.0 provides central point of fabric management

- Power-on Auto-Provisioning (PoAP)
- Cable-plan consistency checks
- Fabric health-monitoring
- Performance monitoring
- Topology map
- Image and configuration repository

What Is ACI?

- Application Centric Infrastructure (ACI) – A policy-based fabric
- Holistic application deployment and management model – application / compute / network / services
- Native support only on Nexus 9500/9300 today

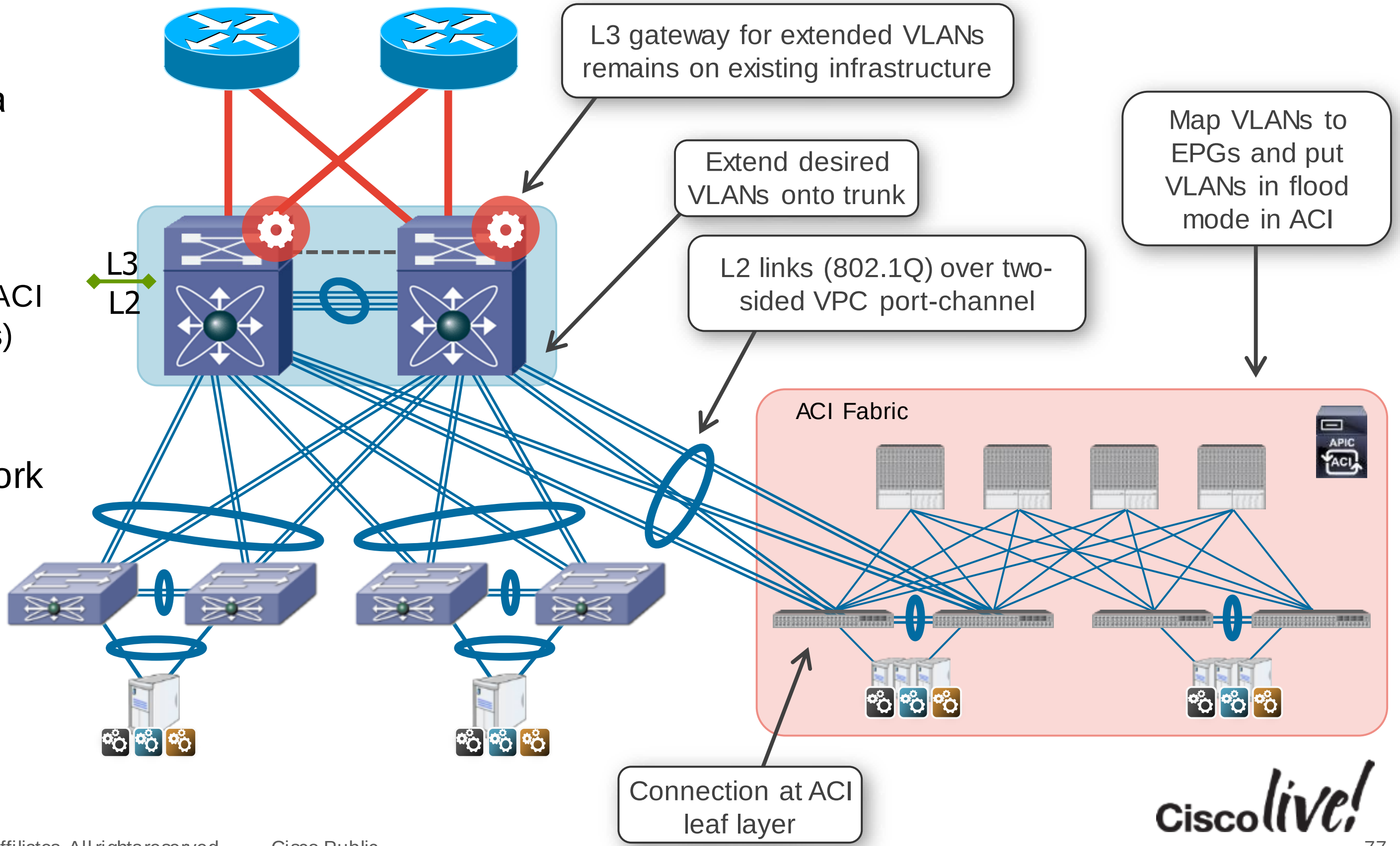


Adding ACI to existing Data Centre network:

- Extend existing workloads / applications into ACI
- Integration via L2
- Integration as “Services Fabric”
- Integration via L3 + overlay

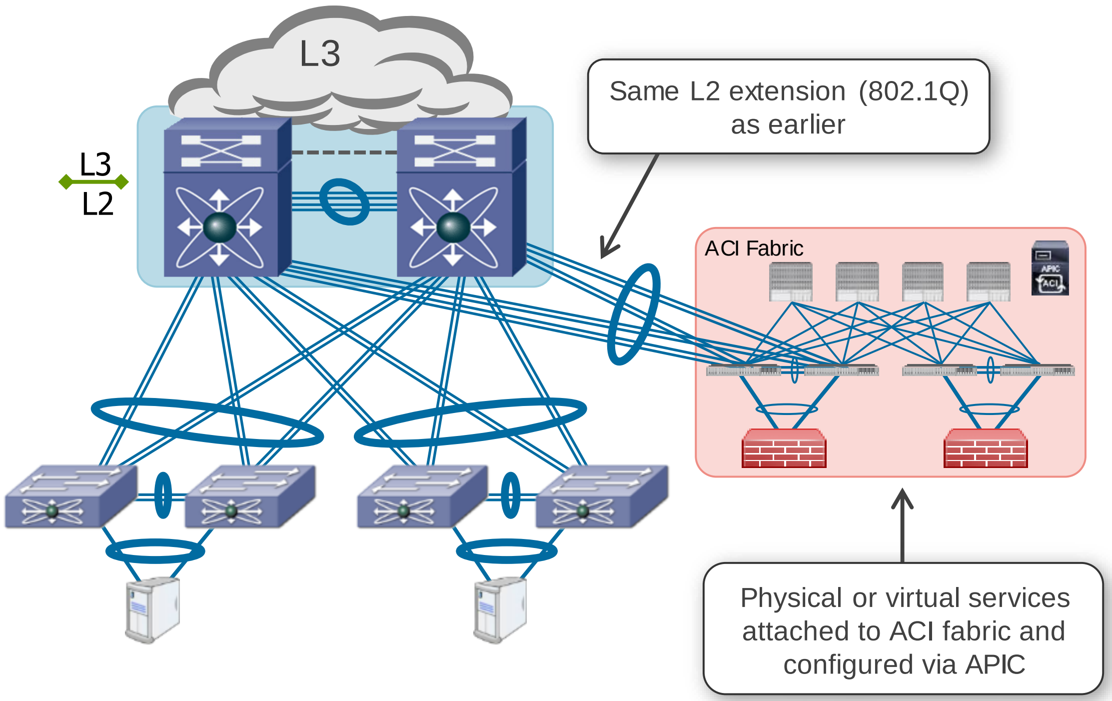
Adding ACI to the Data Centre – Layer 2 Option

- ACI fabric as "leaf switch" in existing Data Centre network
- Extend L2 VLANs into ACI fabric
  - Map 802.1Q VLANs to ACI end-point groups (EPGs)
- Applications and workloads extend between existing network and ACI fabric



Adding ACI to the Data Centre – ACI as Services Fabric

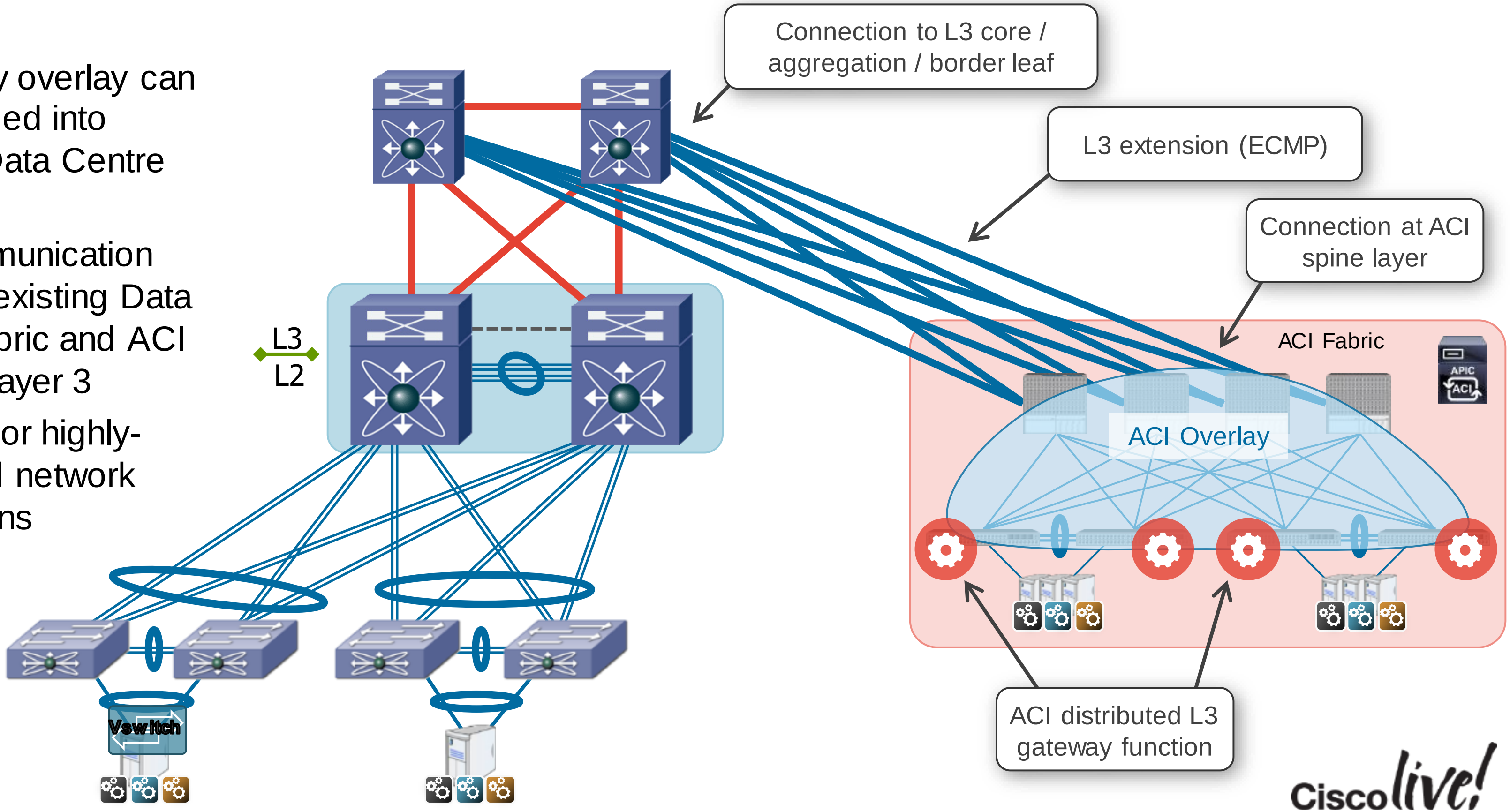
- Hosts in ACI fabric provide Layer 4-7 services – physical and/or virtual
- Easily scale up/down services based on changing requirements
- Security and other L4-7 services policy managed via APIC controller



Extending an ACI Overlay into the Data Centre

Roadmap

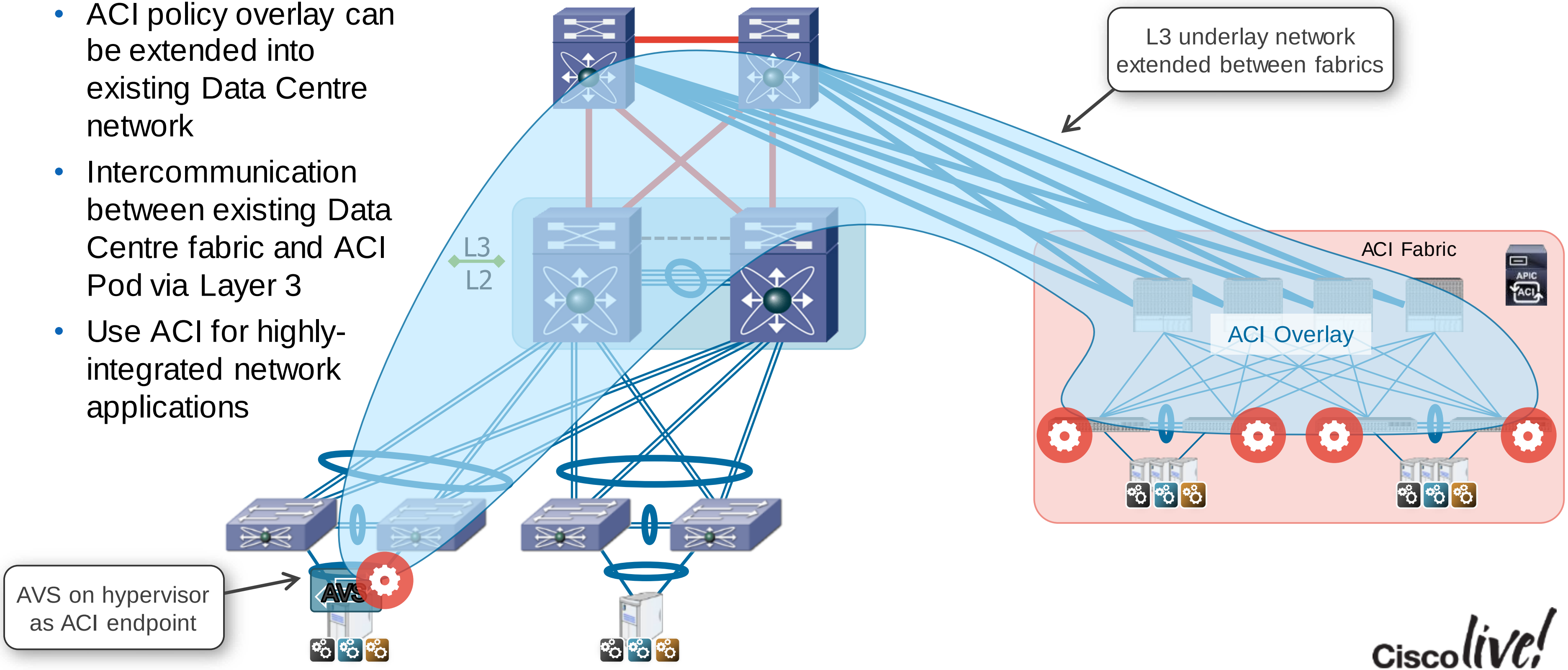
- ACI policy overlay can be extended into existing Data Centre network
- Intercommunication between existing Data Centre fabric and ACI Pod via Layer 3
- Use ACI for highly-integrated network applications



Cisco live!

Extending an ACI Overlay into the Data Centre

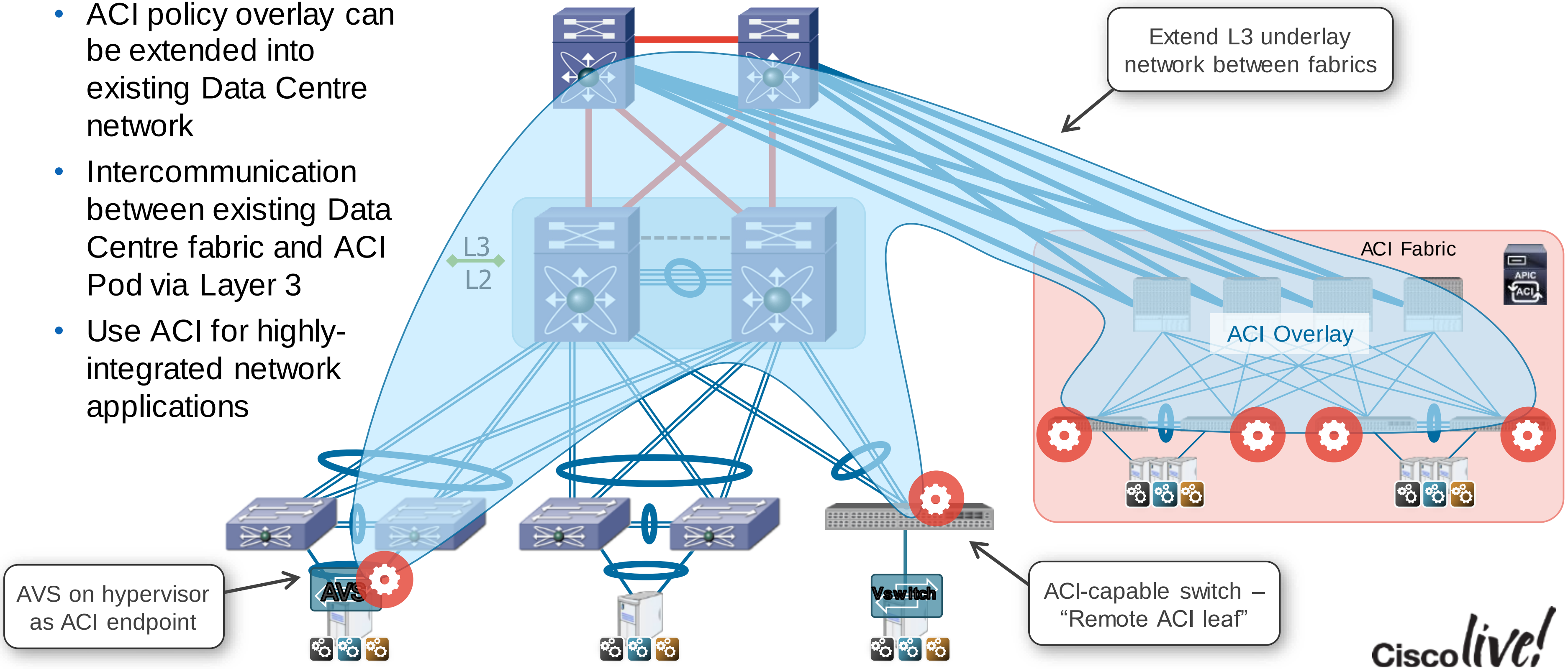
- ACI policy overlay can be extended into existing Data Centre network
- Intercommunication between existing Data Centre fabric and ACI Pod via Layer 3
- Use ACI for highly-integrated network applications



Extending an ACI Overlay into the Data Centre

Roadmap

- ACI policy overlay can be extended into existing Data Centre network
- Intercommunication between existing Data Centre fabric and ACI Pod via Layer 3
- Use ACI for highly-integrated network applications



   Cisco live!

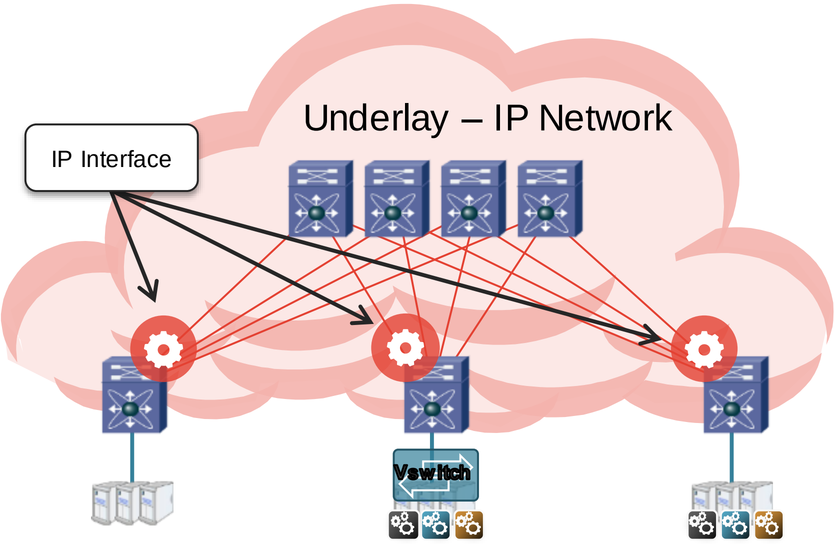
Emergence of VXLAN

- “Standards-based” overlay technology (RFC 7348) – New encapsulation for data-centre fabric
- Provides segmentation, IP mobility, and scale to Data Centre networks
- Leverages IP-based underlay with L3 ECMP

VXLAN Basics

Roadmap

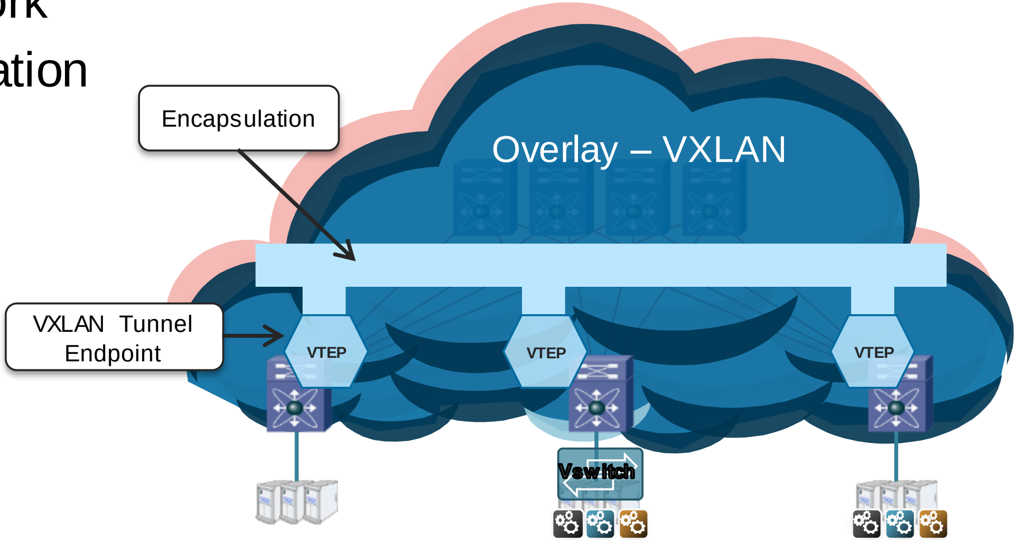
- Underlay – Layer 3 IP network



VXLAN Basics

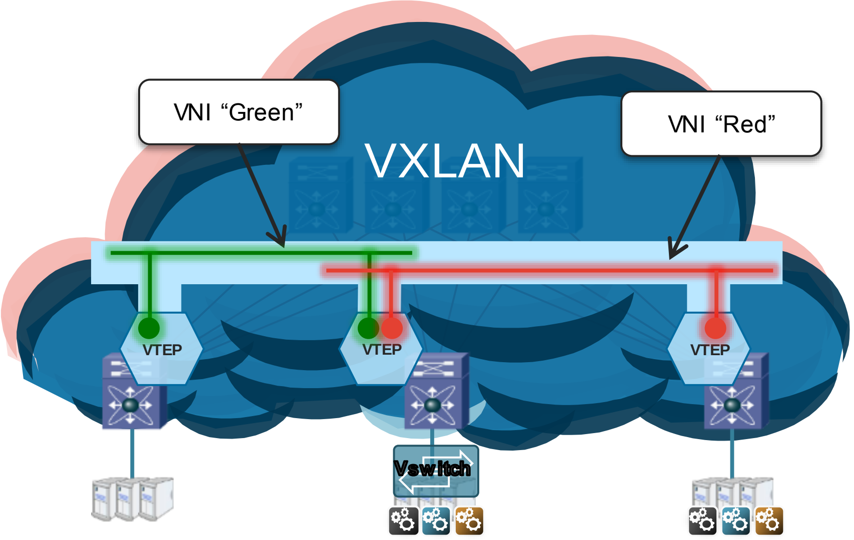
Roadmap

- Underlay – Layer 3 IP network
- Overlay – VXLAN encapsulation



VXLAN Basics

- Underlay – Layer 3 IP network
- Overlay – VXLAN encapsulation
- VNI – VLXAN Network Identifier

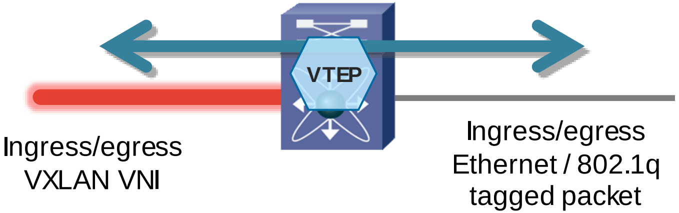


Roadmap

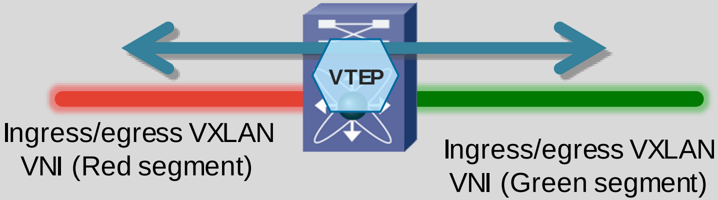
VXLAN on Nexus 7000 / Nexus 7700

- Near-term roadmap item for Nexus 7000 / Nexus 7700 **F3 I/O modules**
- Comprehensive VXLAN VTEP functionality in F3:
  - VXLAN bridging
  - VXLAN-to-VLAN bridging
  - VXLAN-to-VXLAN routing
  - VXLAN-to-VLAN routing

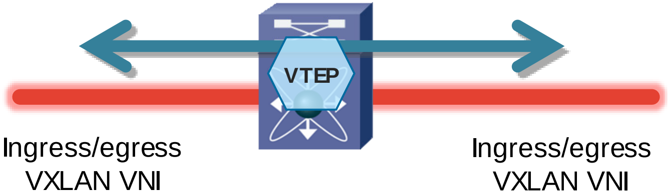
VXLAN-to-VLAN bridging



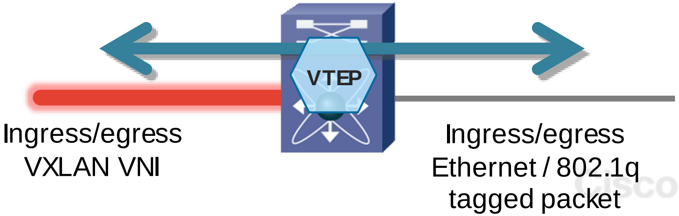
VXLAN-to-VXLAN routing



VXLAN bridging



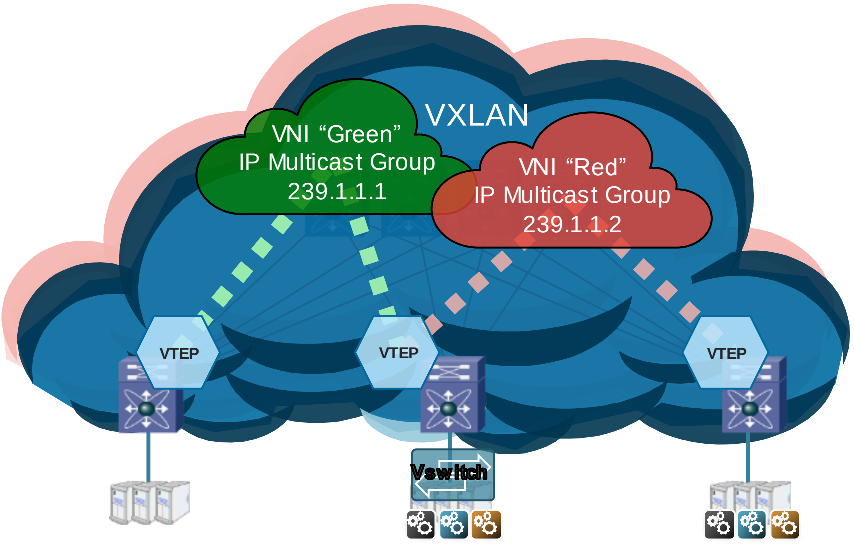
VXLAN-to-VLAN routing



VXLAN with Flood and Learn

Roadmap

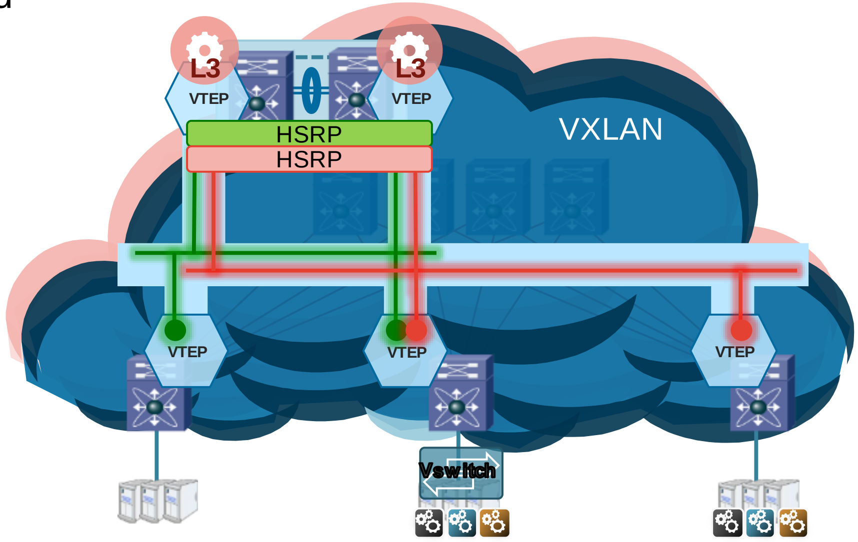
- Host learning on VTEPs based on flood and learn behaviour
- VTEPs join underlay IP multicast groups based on VNI ‘membership’
  - If VNI exists behind VTEP, join corresponding IP multicast group in underlay
- ARP (and other broadcast / unknown unicast / multicast traffic) in a given VNI flooded to all interested VTEPs



Gateway Functions in VXLAN Flood and Learn

Roadmap

- Gateway functions centralised in VXLAN flood and learn
- Nexus 7000 / 7700 VPC pair with L2 + L3 VXLAN gateway capabilities
- VPC provides MAC state synchronisation and active-active HSRP forwarding
- Redundant VTEPs share Anycast VTEP IP address in underlay
- VXLAN bridging occurs directly between VTEPs

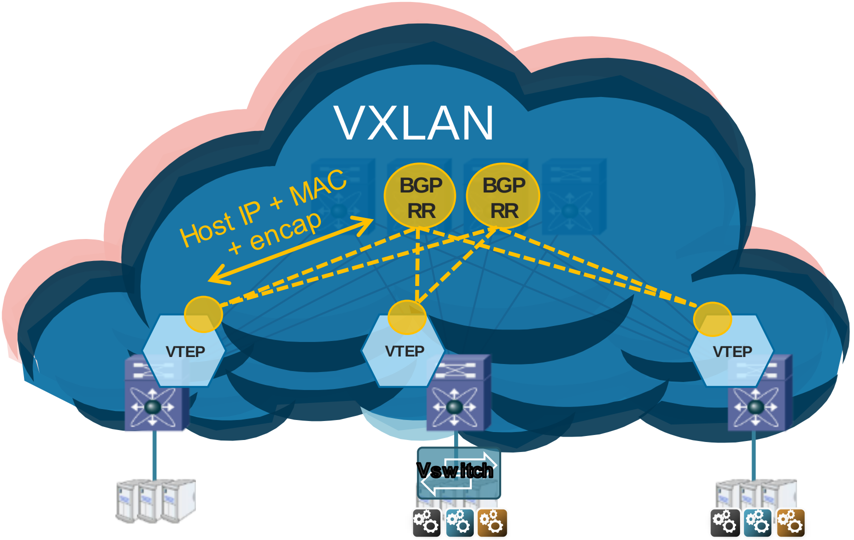


VXLAN + EVPN

Roadmap

- Host learning on VTEPs based on control-plane learning via MP-BGP using the EVPN address family
- VTEPs advertise new host MAC/IPs in BGP
- Route reflectors reduce number of peering sessions
- VTEPs still join underlay IP multicast groups to handle broadcast / multicast / unknown unicast traffic forwarding

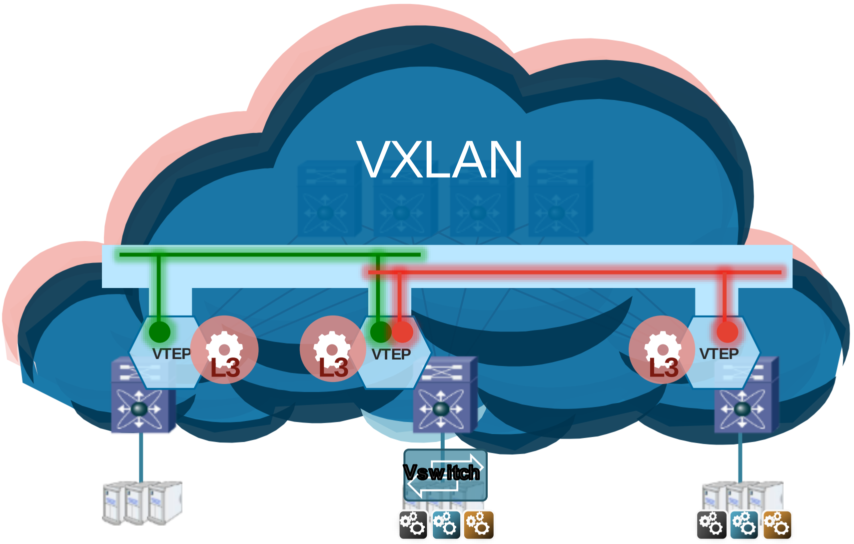
  Or, perform head-end replication…



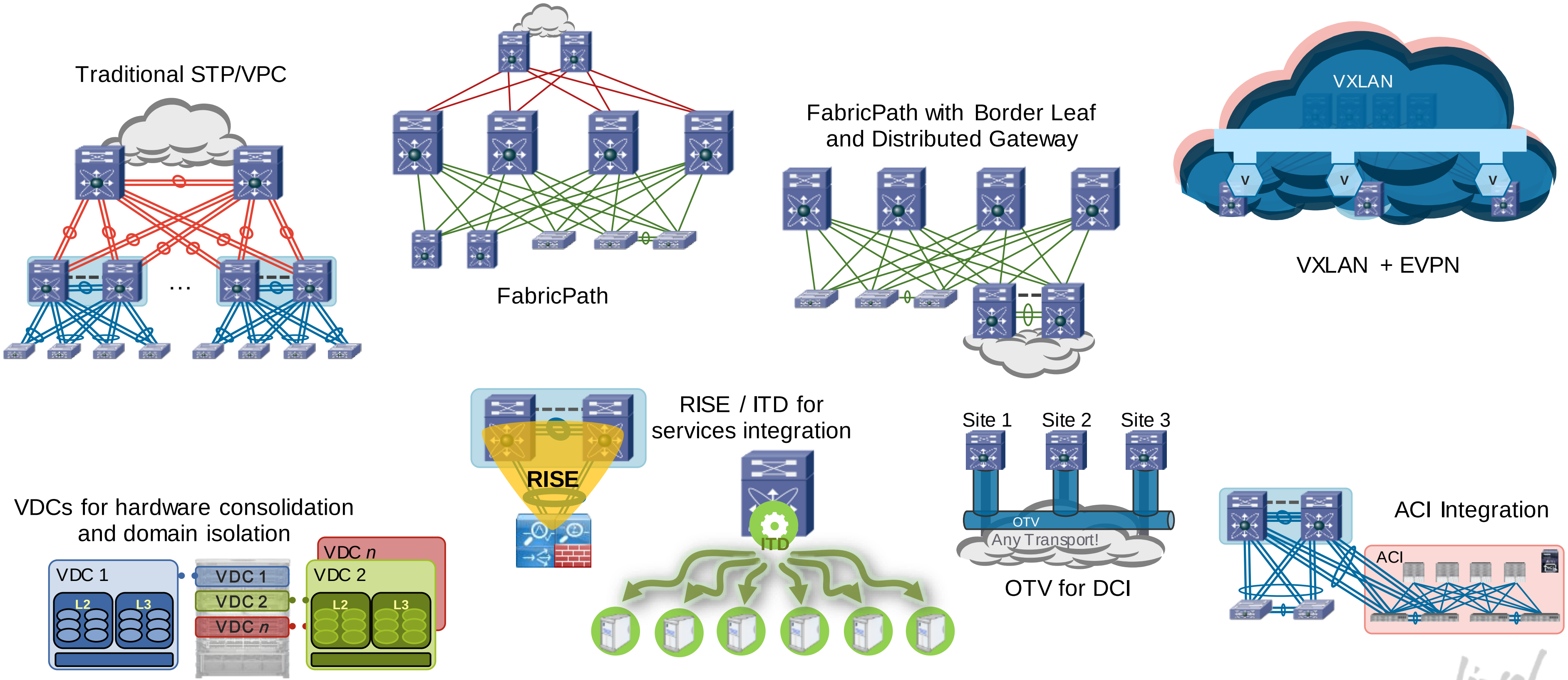
Gateway Functions in VXLAN + EVPN

Roadmap

- Gateway functions distributed in VXLAN + EVPN
  - Centralised gateways also possible
- VXLAN bridging and routing occurs directly between VTEPs
  - Host reachability known via MP-BGP
- All leaf switches share gateway IP and MAC for each subnet
  - No HSRP
  - No change to gateway when hosts move within fabric



Flexible Data Centre Options with Nexus 7000 / Nexus 7700



Key Takeaways

- Nexus 7000 / Nexus 7700 switching architecture provides **foundation** for flexible and scalable Data Centre network designs
- Nexus 7000 / Nexus 7700 design **building blocks** interwork and complement each other to solve customer challenges
- Nexus 7000 / Nexus 7700 platform continues to **evolve** to support next-generation/emerging technologies and architectures

Q & A

Complete Your Online Session Evaluation

Give us your feedback and receive a Cisco Live 2015 T-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 http://showcase.genie-connect.com/clmelbourne2015
- Visit any Cisco Live Internet Station located throughout the venue

T-Shirts can be collected in the World of Solutions on Friday 20 March 12:00pm - 2:00pm

Learn online with Cisco Live!
Visit us online after the conference for full access to session videos and presentations. www.CiscoLiveAPAC.com

Thank you.

cisco