Cisco live!

What You Make Possible

TOMORROW starts here.

CISCO

VMDC Architecture for Multi-Tennant IaaS Cloud Computing

BRKSPG-2663

Agenda

- VMDC Introduction
  - Architectural Baseline
- Hierarchical L3 Data Centre Designs
  - VMDC 2.x
  - Tiered Security Services
  - Service Differentiation
- Data Centre Interconnect
  - Scalable VPLS Design
- FabricPath L2 Data Centre Designs
  - VMDC 3.x
- Virtualised Services based Data Centre Designs
  - VMDC 4.x
- Cloud Orchestration & Assurance

© 2013 Cisco and/or its affiliates. All rights reserved.

VMDC Cloud Architecture

- Introduction

Cisco Validate Design Process

Innovation and Quality Through System Level Design and Validation

Key Customer Engagements
Consider end-to-end view

Product Development
Cross platform collaboration

System Development Fundamentals

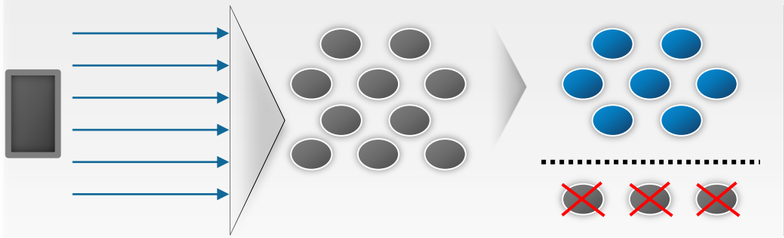
Thought Leadership
System level innovations

System Delivery
Tested and validated designs

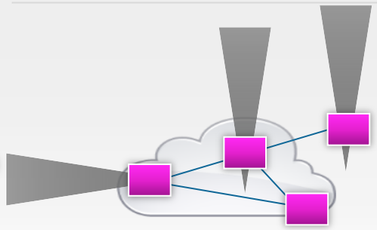
Cisco Validated Design

System Development Guidelines

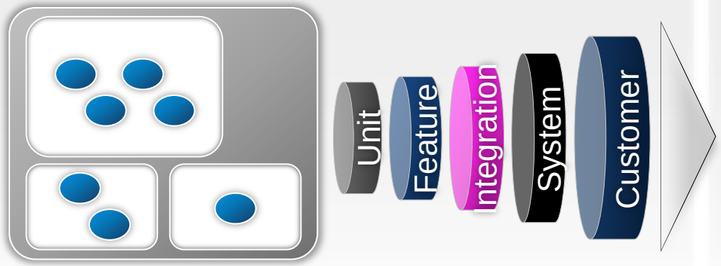
Planning



Design



End-To-End Validation

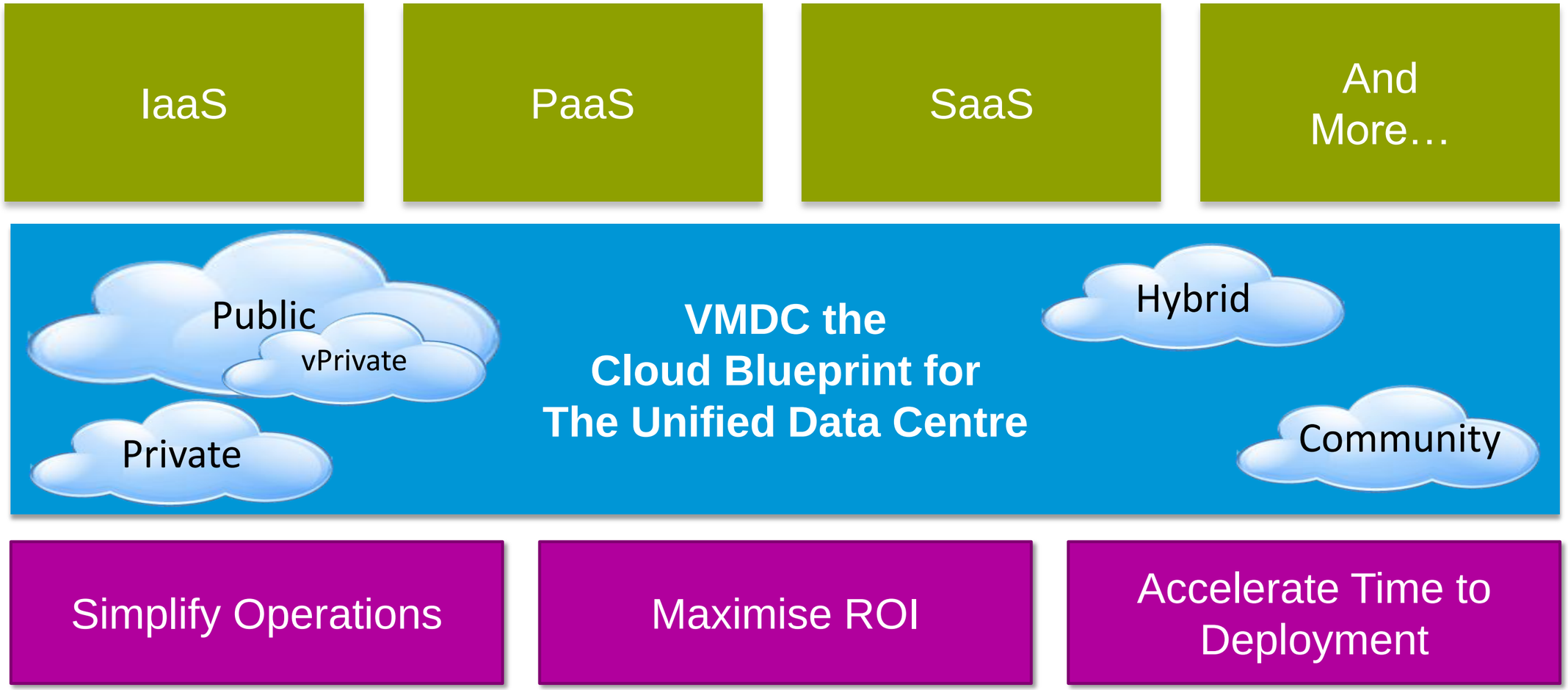


Documentation

© 2013 Cisco and/or its affiliates. All rights reserved.

VMDC – Cloud Blueprint for the Unified Data Centre

Foundation for Cloud Applications and Services



© 2013 Cisco and/or its affiliates. All rights reserved.

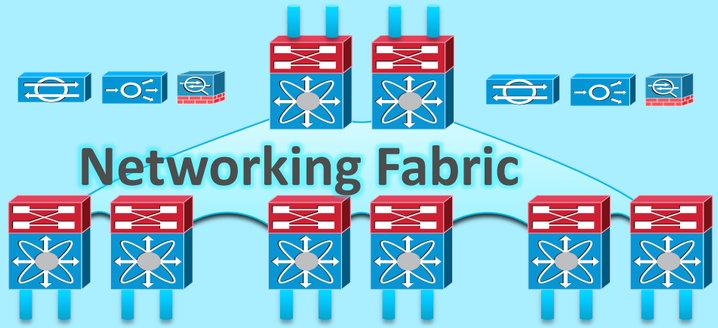
Virtualised Multiservice Data Centre

Inter-Data Centre Networking

WAN

Unified Fabric and Data Centre Networking

Providing Network and Services Virtualisation

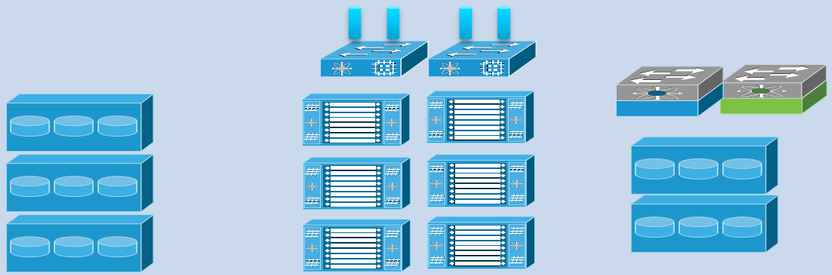


Services

Network Fabric

Unified Computing and Integrated Systems

Providing Server and Application Virtualisation



Access

Storage

Compute

Cloud Service Management

- Business Support
- Provisioning Configuration
- Portability/ Interoperability

VMDC

© 2013 Cisco and/or its affiliates. All rights reserved.

Building a Multi-Service Infrastructure

Key Considerations

Modularity
Pod based design
Scalability framework for manageable increments
Predictable physical and cost characteristics
Streamline Turn-up of New Services

High Availability
Carrier Class Availability
Platform/Network/Hardware/Software Resiliency
Minimise the probability and duration of incidents
Focus on your business, not fighting fires

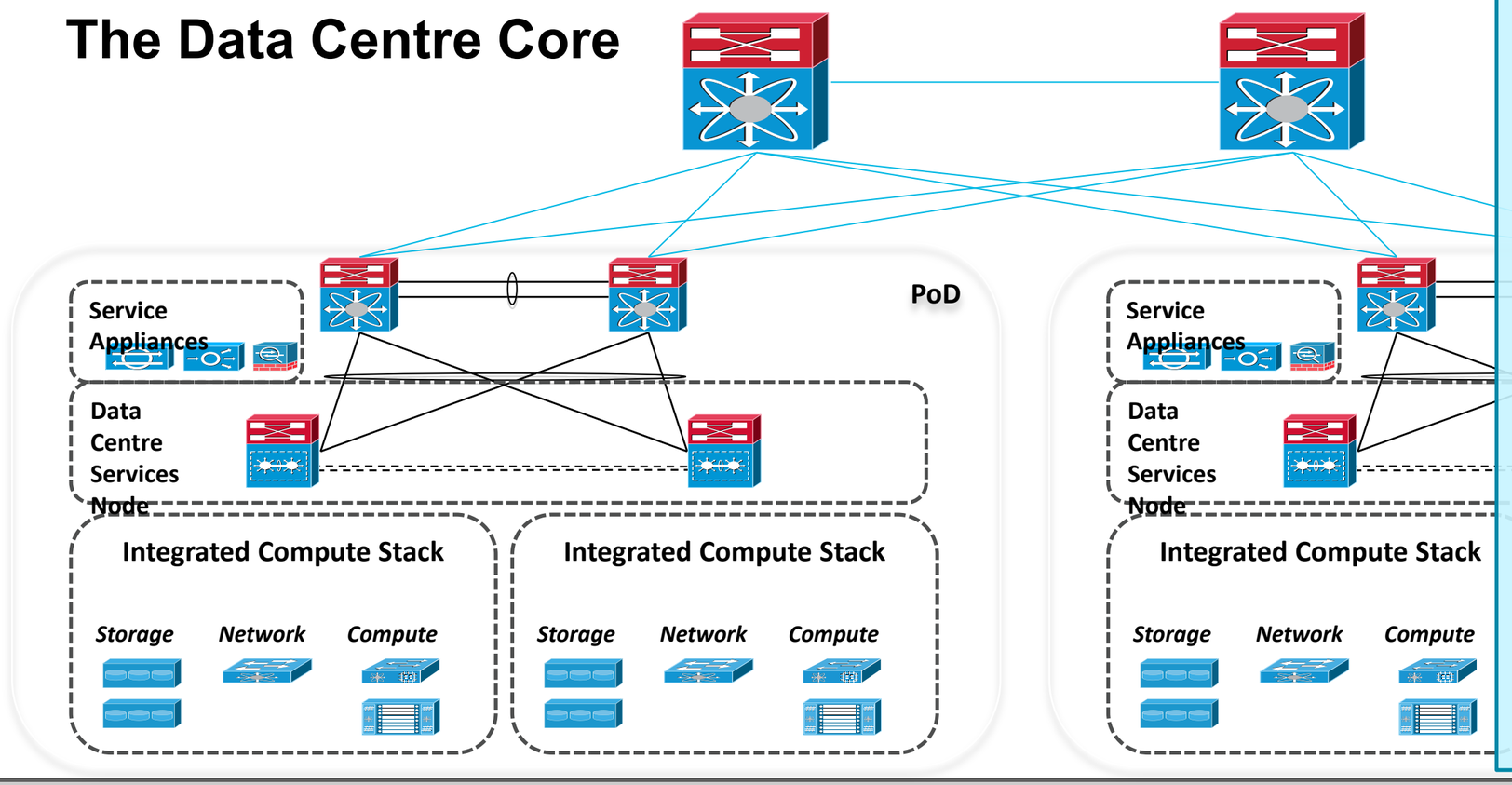
Secure Multi-tenancy
Shared Physical Infrastructure
Tenant Specific Resources
Use Cases
Comply with business policies

Differentiated Service Support
Design logical models around use cases
Services-oriented framework
Combines compute/storage/network
Resources are applied and tuned to meet needs

Service Orchestration
Dynamic application and reuse of resources
Automated service orchestration and fulfillment
Integration with Network Containers
Rapid Self Service IT

© 2013 Cisco and/or its affiliates. All rights reserved.

Scale Through Modularity



Factors to Consider

- **L2 Scale -** Virtual Machine Density, VMNics per VM, MAC Address Capacity,
- Cluster Scale, ARP Table Size, VLAN scale, Port Capacity, Logical Failure Domains L2 Control Plane
- **L3 Scale –** BGP Peering, HRSP Interfaces, VRF Instances, Routing Tables and Convergence, Services
- **Resource Oversubscription –** Network Compute, and Storage Oversubscription, Bandwidth per VM

The Solution

- PoD replication

Benefits

- Optimise CAPEX savings while maintaining SLAs
- Predicable performance and scale based on building blocks
- Effective way to add separate application environments

VMDC Security Control Framework

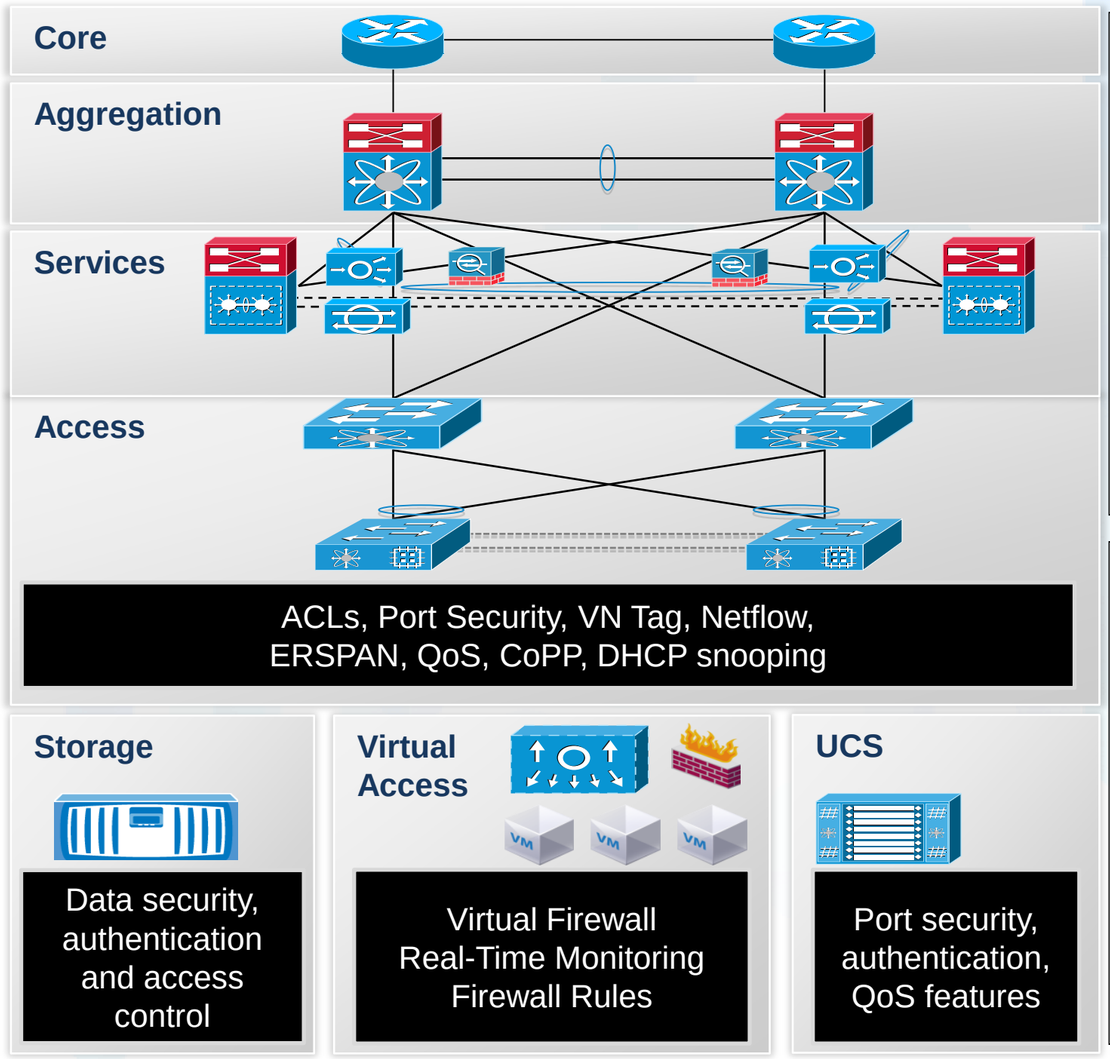
Security Management

- Visibility
- Event correlation, syslog, centralised authentication
- Forensics
- Anomaly detection
- Compliance

AD CSM ACS

Services

- Initial filter for DC ingress and egress traffic; Virtual Context used to split polices for server-to-server filtering
- Additional firewall services for server farm specific protection



Infrastructure Security

- Infrastructure Security features are enabled to protect device, traffic plane, and control plane
- 802.1ae provides separation through encryption

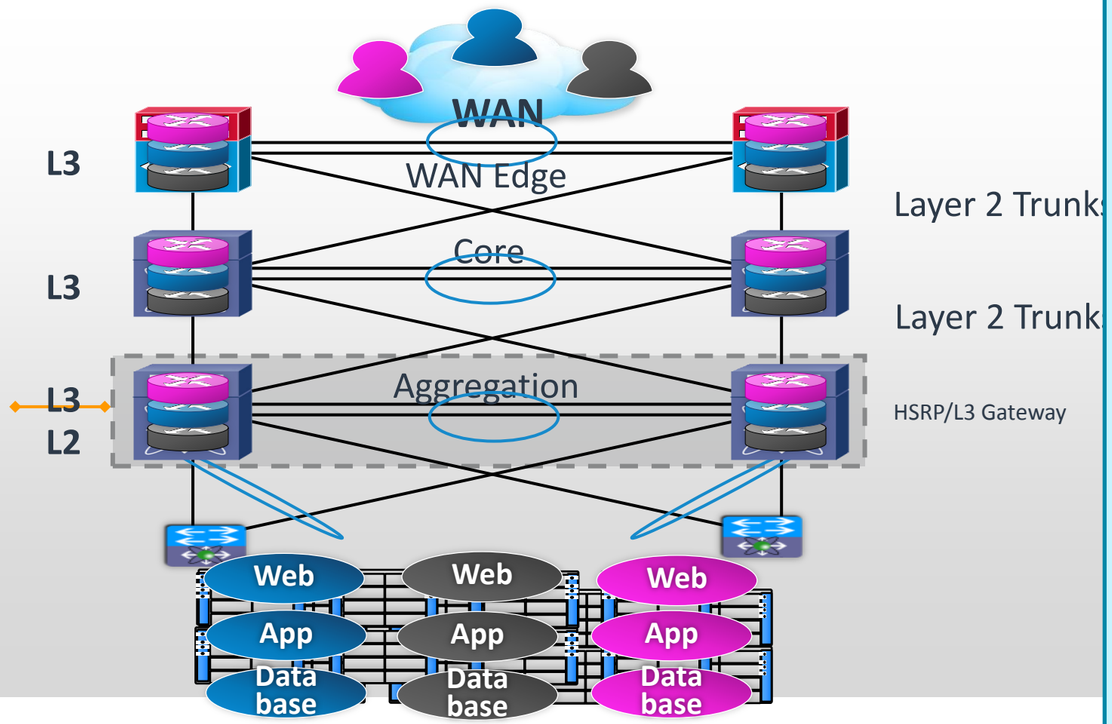
Services

- IPS/IDS provide traffic analysis and forensics
- Network Analysis provide traffic monitoring and data analysis
- Server load balancing masks servers and applications

© 2013 Cisco and/or its affiliates. All rights reserved.

Secure Tenant Separation

- VRF-lite implemented at core and aggregation layers provides per tenant isolation at L3
- Separate dedicated per-tenant routing and forwarding tables insuring that no inter-tenant (server to server) traffic within the data centre will be allowed, unless explicitly configured
- VLAN IDs and the 802.1q tag provide isolation and identification of tenant traffic across the L2 domain



- Compute Separation (vNICs, VLANs, Port Profiles)
- Storage Separation (Cluster File System Mgmt, VSAN and FC Zoning, LUN Masking, vFilers)
- Application Tier (Network Centric, Server Centric, VSG)

The Solution

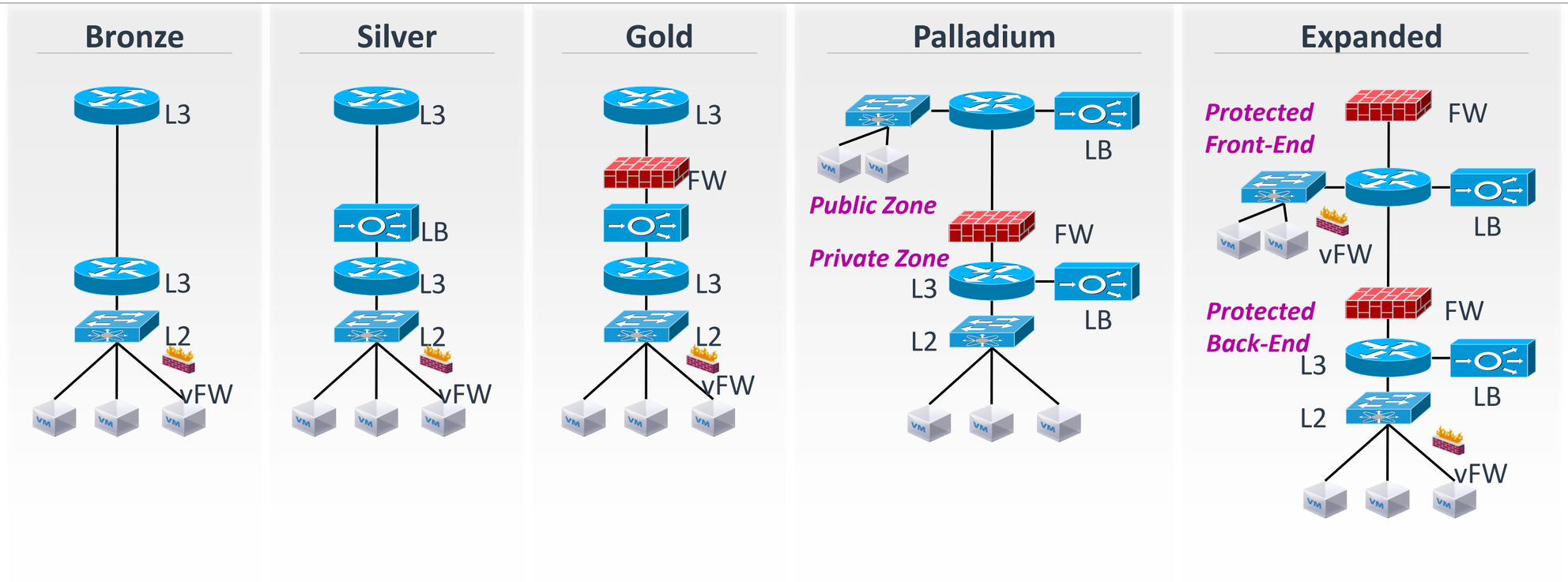
- Virtualise L3 network and map to VLANs

Benefits

- End to end secure separation across the data centre
- Overlapping IP addresses are allowed
- Automation tools to simplify deployment

Cloud Consumer Models

Example Validated Tenancy Models



The Solution

- Sample tenant containers

Benefits

- Quickly and securely onboard similar tenants
- Covers different levels of network services for a variety of needs
- Addresses varying security, QoS, and other requirements
- Solutions available to automate the process

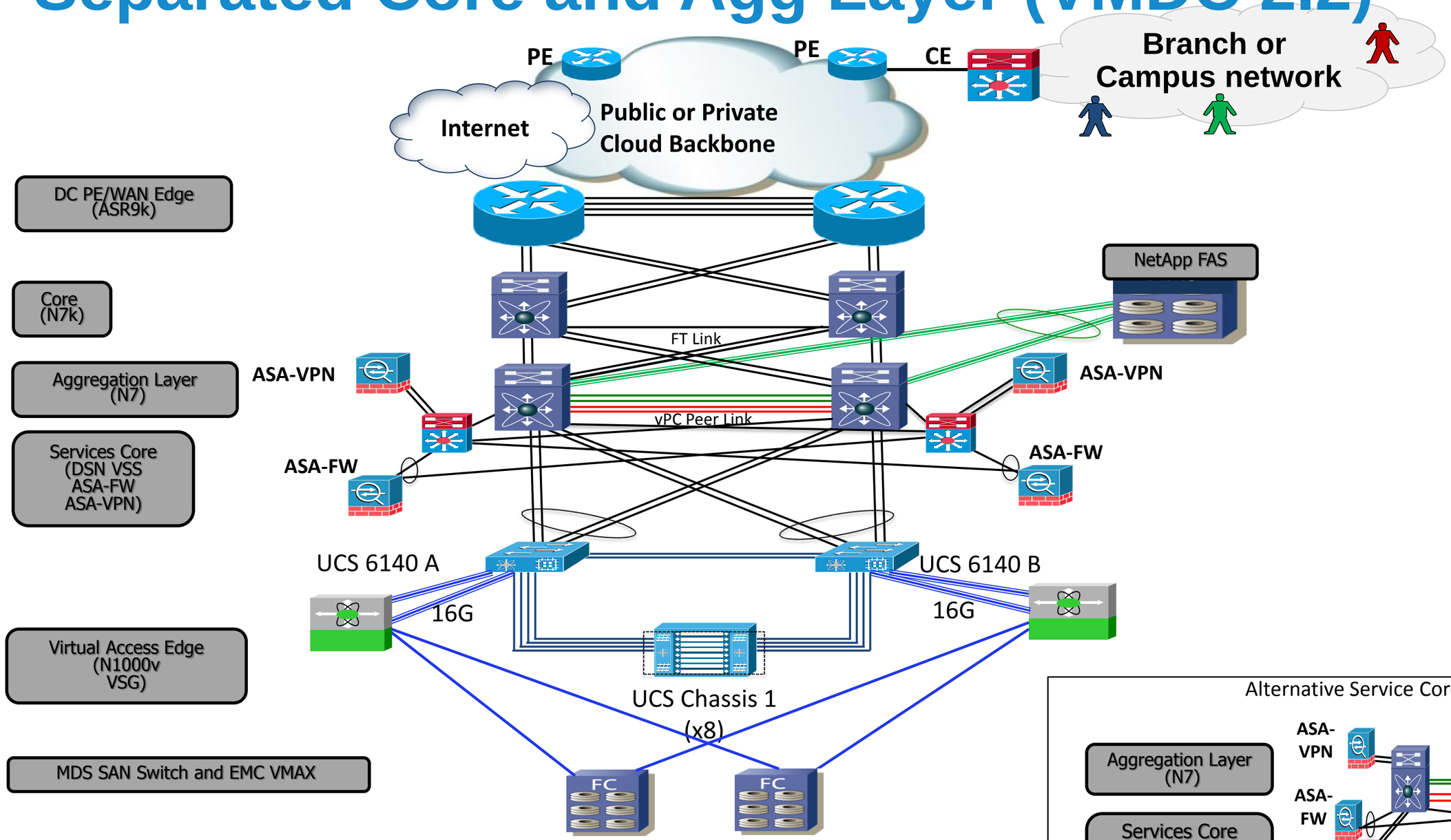
VMDC Solution Validation Scope

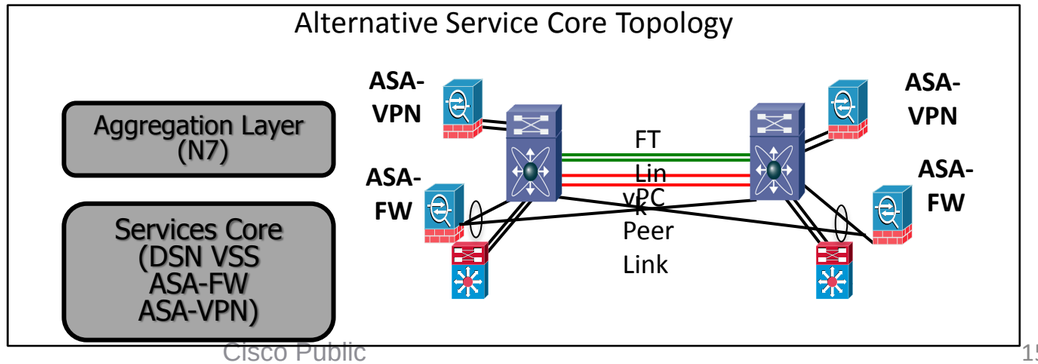
- End to end feature/integration testing to enable service delivery, and multi-tenancy / isolation
- Validation of Service Tier offering (network, compute, storage) and DC Services (VPN, FW, IPS, GSS etc)
- Baseline Steady State Traffic
  40G N-S, 40G E-W Stateful and Stateless traffic mix
- Failover/Negative tests to validate redundancy designs (with Baseline Steady State traffic) – Routing, vPC/MEC, ECMP, VSS, HSRP, Active-Active service modules, Clustering, SAN, Fabric, UCS blades, Storage controllers
- Network convergence measurements – HA tests
  Node, LC, Sup, Link,
- Stress/Load tests to validate end-end Service Flows, QoS, reliability, monitor cpu/memory
- Multi-dimensional Scalability (Tenants, VMs, VLANs, MAC, HSRP, Routes, Contexts)
- Validation of statistics and monitoring capabilities – SNMP, Sylog, Netflow, IO statistics on VC, N1k, UCS, N7k etc
- 3rd party systems including BMC CLM, Zenoss CSA, Vmware vSphere, vCloud Director, EMC , NetApp, Citrix Netscaler, Microsoft, RedHat

© 2013 Cisco and/or its affiliates. All rights reserved.

Hierarchical L3 Data Centre Designs -VMDC 2.x

Separated Core and Agg Layer (VMDC 2.2)

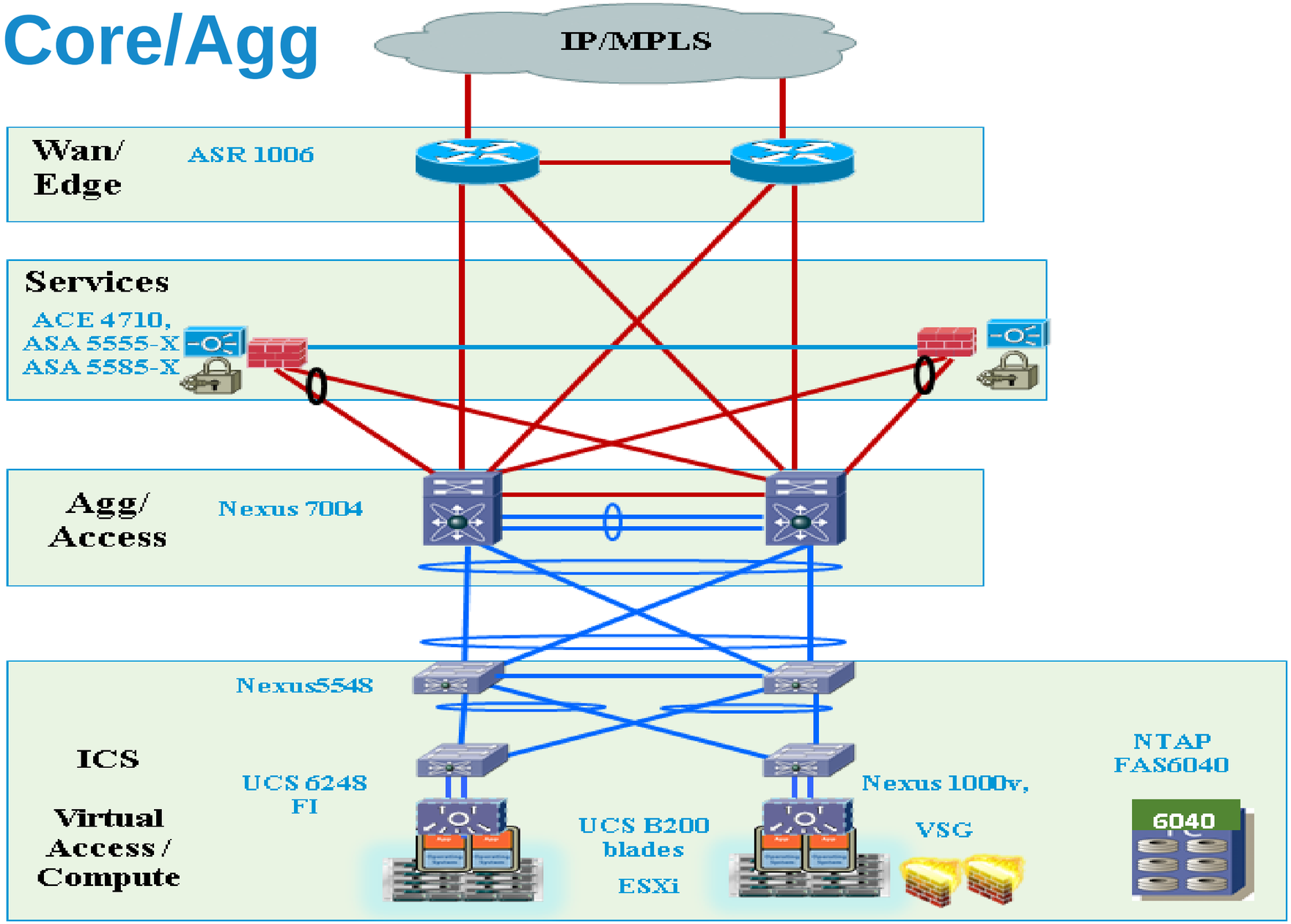




© 2013 Cisco and/or its affiliates. All rights reserved.

Collapsed Core/Agg

(VMDC 2.3)



© 2013 Cisco and/or its affiliates. All rights reserved.

Scale – Design Points

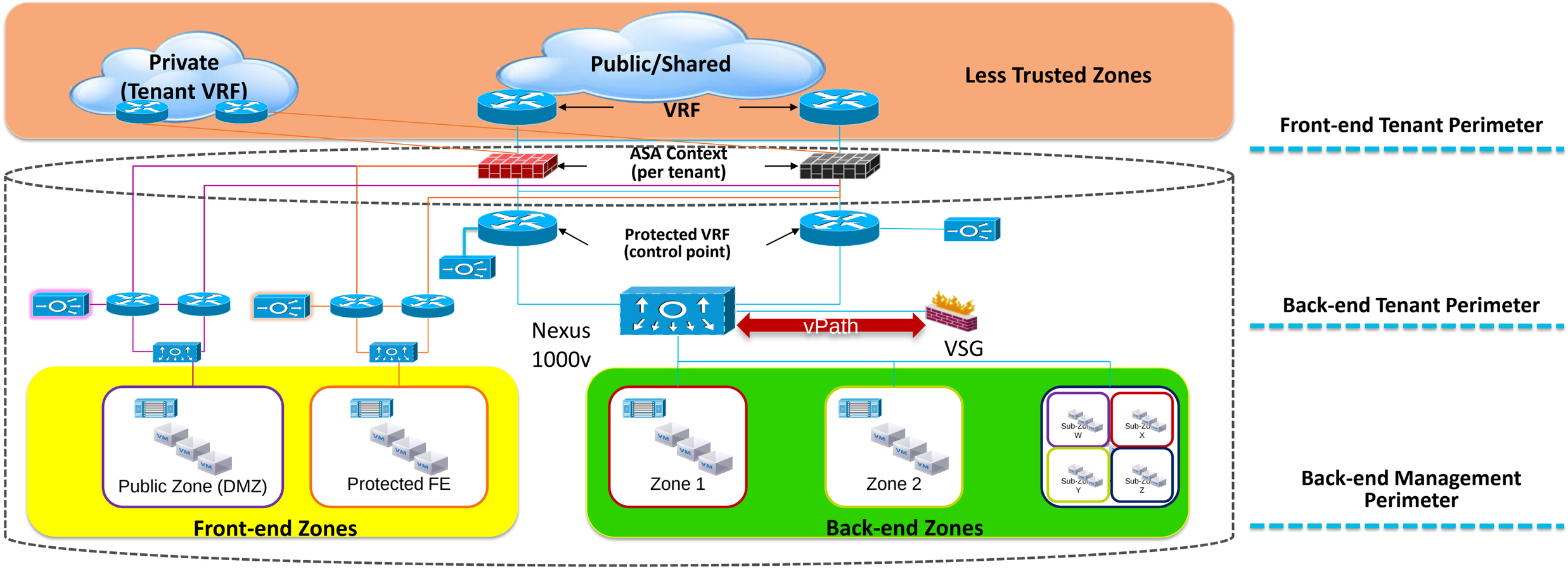
Scale Factors	VMDC 2.2	VMDC 2.3
# ICS in PoD	8	3
# UCS Blades in ICS	64	64
# PoD in DC	6	4
# VMs in PoD	12,000	6,000
# VMs in DC	72,000	24,000
# Tenants in PoD	200	500
# Tenants in DC	200	2000

© 2013 Cisco and/or its affiliates. All rights reserved.

Tiered Security Services

VMDC Consumer Model

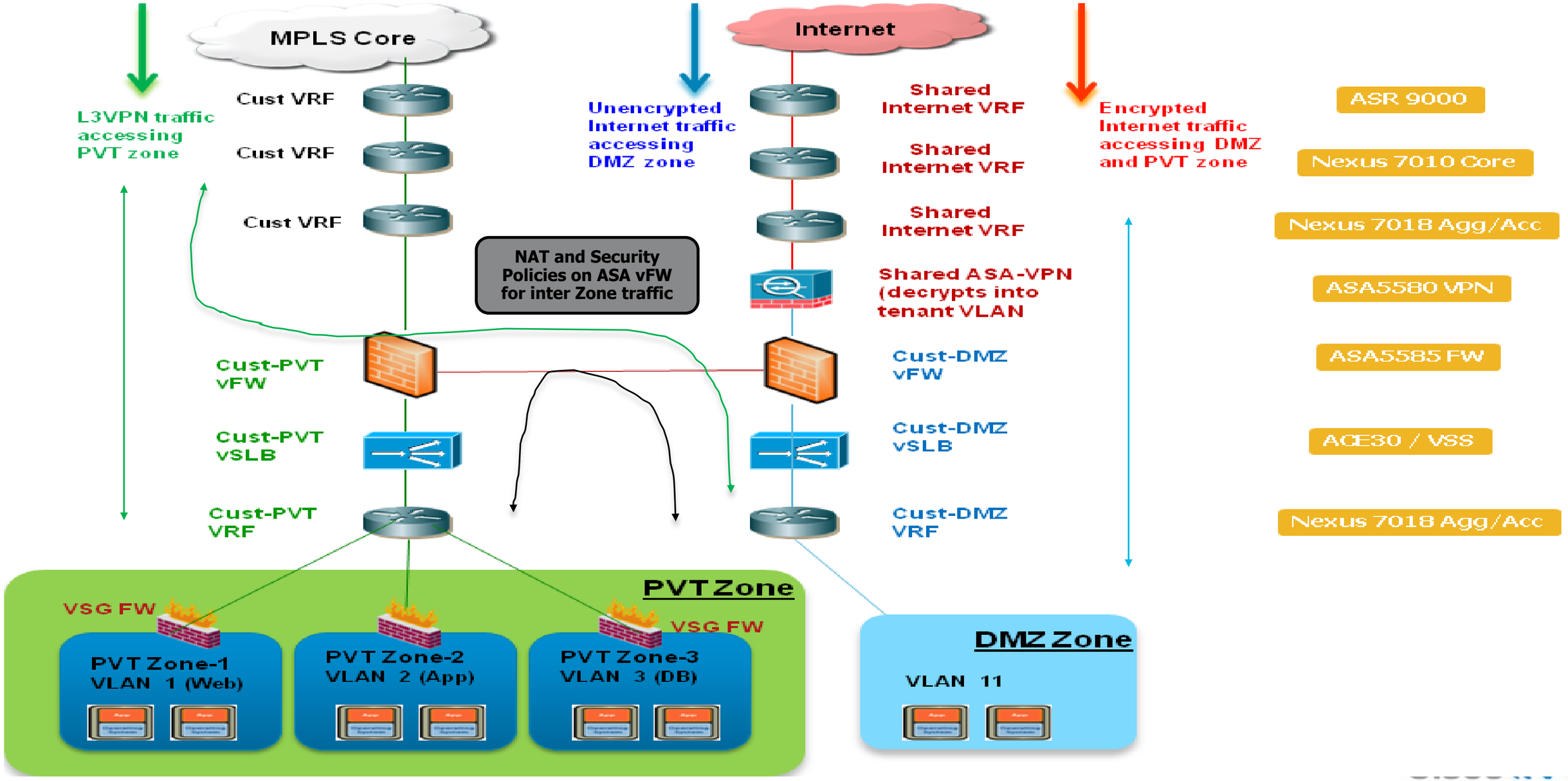
Logical Perimeters and Zones



Note: RA VPN Concentrators not shown

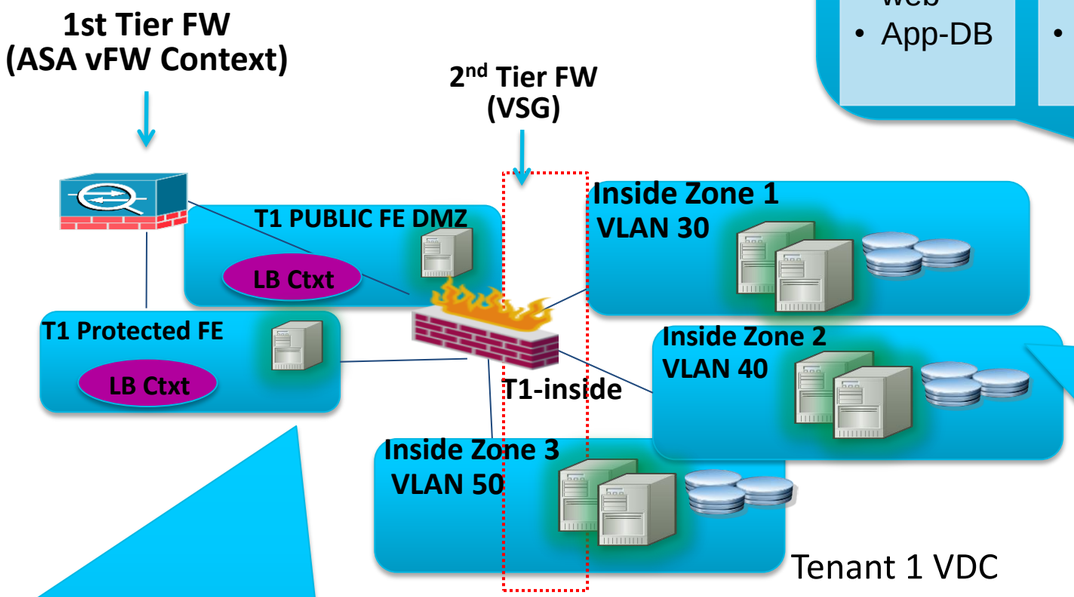
© 2013 Cisco and/or its affiliates. All rights reserved.

VMDC 2.2 Gold Service (PVT + DMZ Zones)



© 2013 Cisco and/or its affiliates. All rights reserved.

Back-End Security Zone Model in VMDC 2.2



Policy 2: App Tier

Name	Source	Dest	Protocol	Ethertype	Action
• Intra-app • App-web • App-DB	• vZone eq App • vZone eq App • vZone eq App	• vZone eq App • vZone eq Web • vZone eq DB	• Any • Any • Any	• Any • Any • Any	• Permit • Permit • Permit

Policy 3: DB Tier

Name	Source	Dest	Protocol	Ethertype	Action
• Intra-DB • DB-web • DB-app • Allow SSH	• vZone eq DB • vZone eq DB • vZone eq DB • Any	• vZone eq DB • vZone eq Web • vZone eq App • Network Port eq 22	• Any • Any • Any • Eq TCP	• Any • Any • Any • Any	• Permit • Permit • Permit • Permit

Policy 1: Web Tier

Name	Source	Dest	Protocol	Ethertype	Action
• Allow DNS • Allow Client HTTP	• Any • Port Profile name-not member T1-VM	• Network Port eq 53 • vZone eq web AND Network port member vm-http-port	• Eq UDP • Eq TCP	• Any • Any	• Permit • Permit

© 2013 Cisco and/or its affiliates. All rights reserved.

VMDC VSG Details

- VSG defined for all tenants. One VSG VM per tenant
- Per Tenant VSG protects Data vnics on VMs.
- Shared Service VLAN, and MGMT vlan for all tenant VSG's (to consume fewer VLANs) - only need 2 VLANs
- Gold/Silver tenants have 3 VLANs (Web, App, Db). Each tenant VSG has 3 "zones" defined – one for each VLAN
- Shared MGMT VSG protects VM mgmt vnic's (SP managed)
- For scalability reasons, defining separate MGMT VSG per ESX Cluster. One "mgmt" tenant, with separate vDC per ESX cluster
- Tenant VSG is a flat organisation - has 3 "zones" for web, app, db for Gold/Silver. No vDCs created in the VNMC.
  - Web zone allows all http/https/ftp traffic from Outside to VM
  - App zone allows only http/https traffic from Web-zone, and all SSH traffic from outside to VM
  - DB zone only allows http/https traffic from App-zone

© 2013 Cisco and/or its affiliates. All rights reserved.

VSG Tested Scale

- Nexus1000v 1.4a release, VSG/VNMC 1.2 release
- VMDC 2.2 testbed setup with 57 tenants (3 Gold, 6 Silver, 48 Bronze), and 504 VM's
- VSG defined for all tenants. So total of 57 VSGs
- Per Tenant VSG protects Data vnics on VMs.
- Gold/Silver tenants have 3 VLANs (Web, App, Db). VSG zone defined for each VLAN
- Each Gold VSG protects 8 VMs, Silver VSG 16 VMs, Bronze VSG 8 VMs
- MGMT VSG protectes VM Mgmt vnic's – so that different tenant VMs cannot talk to each other on the Mgmt vnic.
- One VSG for protecting all 504 VMs does not scale (300 limit per VSG)
- Defining separate MGMT VSG per ESX Cluster. So 3 MGMT-VSGs, each protecting 168 VMs
- Tested NW attributes, VM attributes, Zones. Goal was 4000 cps on each VSG
- VSG's were hosted in a separate ESX cluster (centralised deployment) of 10 UCS B200-M2 blades

© 2013 Cisco and/or its affiliates. All rights reserved.

VSG Best Practices

- Make sure to deploy VNMC to the VMware cluster with HA enabled.
- It is advisable to configure a unique HA ID for each VSG HA pair, even when VSGs are not sharing the same HA VLAN.
- In a HA VSG deployment, the primary and secondary VSG should not be placed on the same ESXi host; use DRS anti-affinity rules, or DRS Host Groups to ensure that the primary and secondary VSG nodes do not run on the same ESXi host.
- VSG does not support VMware DRS live vMotion. VMware DRS should be disabled or set to partially automated.
- In an HA VSG deployment, the primary and secondary VSG should not be placed on the same data store; create at least two data stores for the primary and secondary VSG nodes.
- Make sure to configure a unique data IP address for each compute firewall profile.
- In multi-tenant deployments, Cisco recommends that the compute firewall object be added directly at the tenant level.

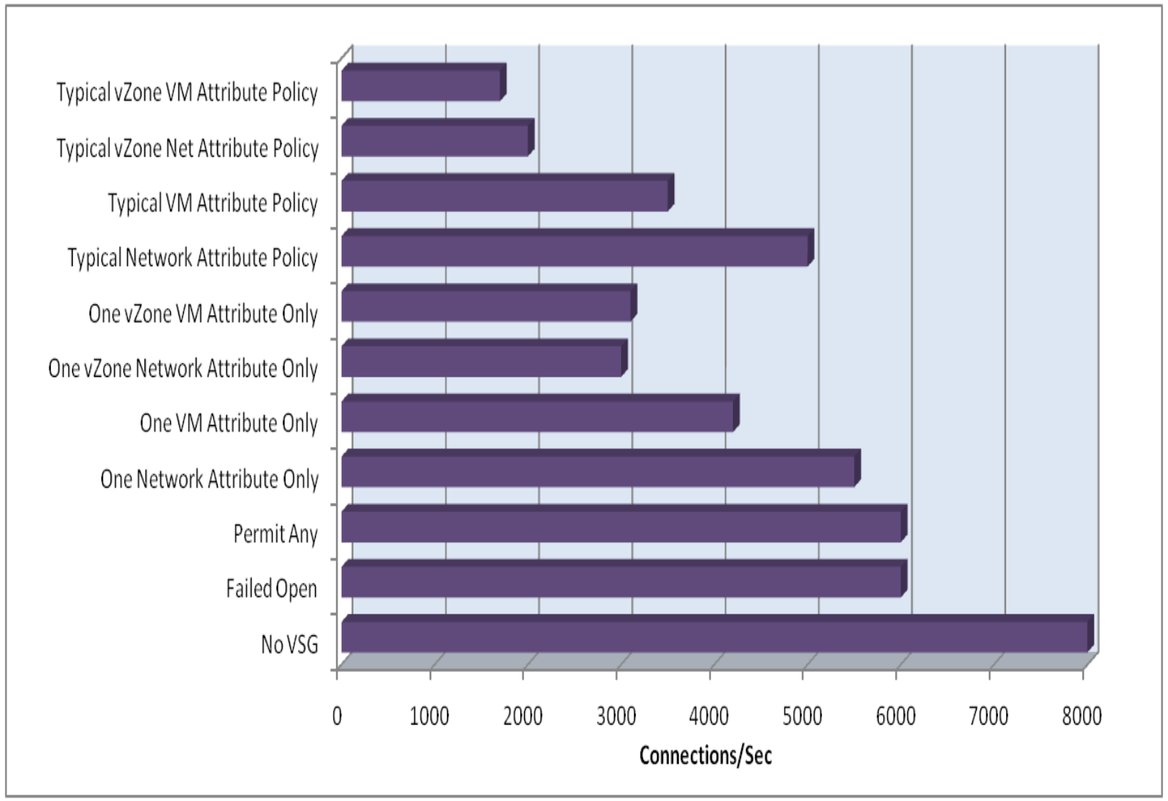
© 2013 Cisco and/or its affiliates. All rights reserved.

VSG Test Findings

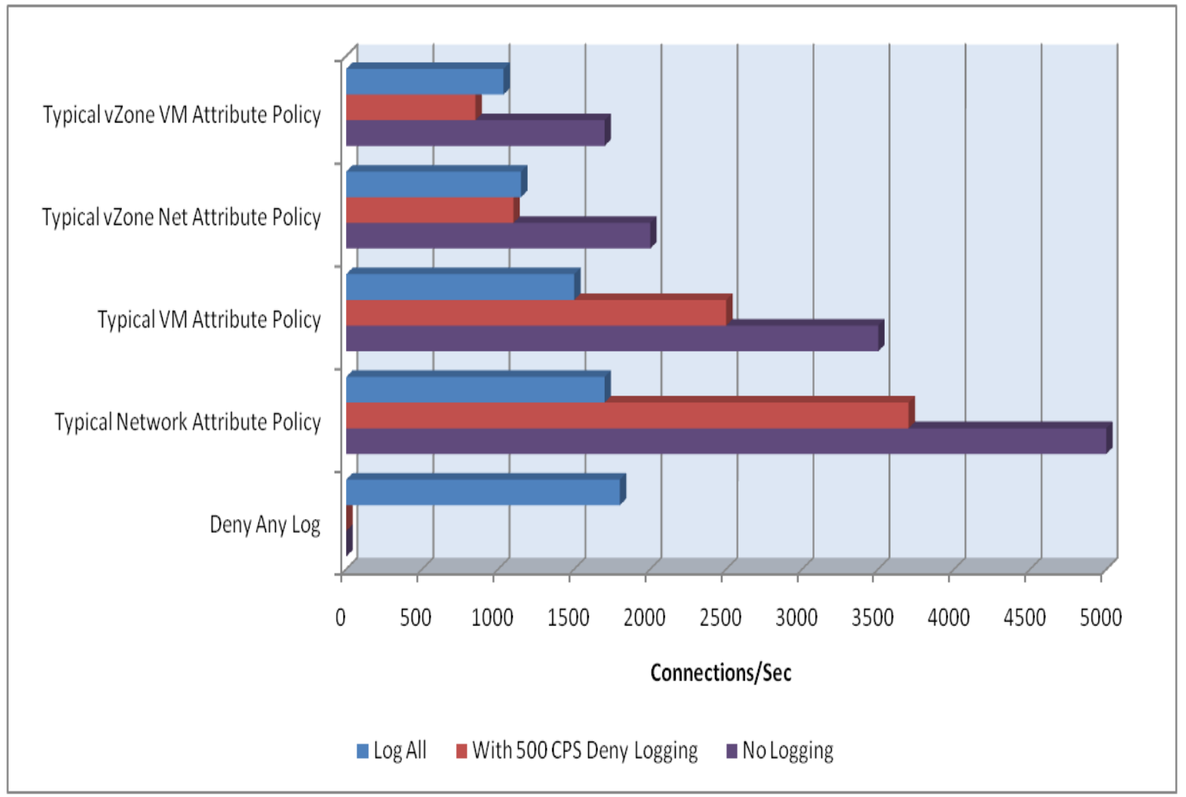
- Using zone in the VSG configuration impacts VSG performance
- Most zone configuration can be replicated using object-groups
- Using VM attributes (vm name, port-profile name, etc), as opposed to using network attributes (ip address, port, etc), will have performance impact on the VSG
- Using Logging/Audits etc will impact performance
- Marketing numbers about 4000 cps with NW Attributes (no zones, logging), and 1500 cps with VM Attributes
- The VSG performance depends on the policy configuration used; the test configuration achieved performance of ~1000 connection per seconds

© 2013 Cisco and/or its affiliates. All rights reserved.

VSG Performance Results



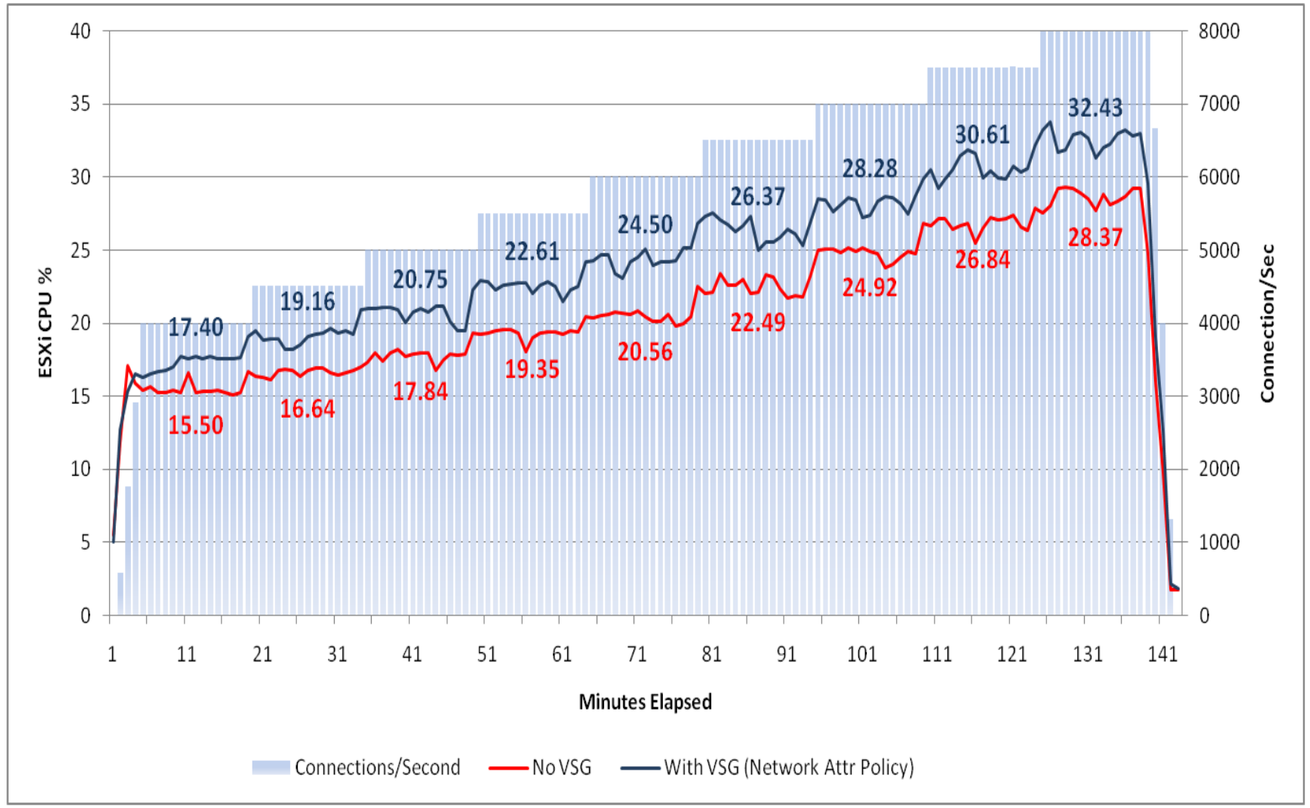
VSG Perf with different Policy Settings

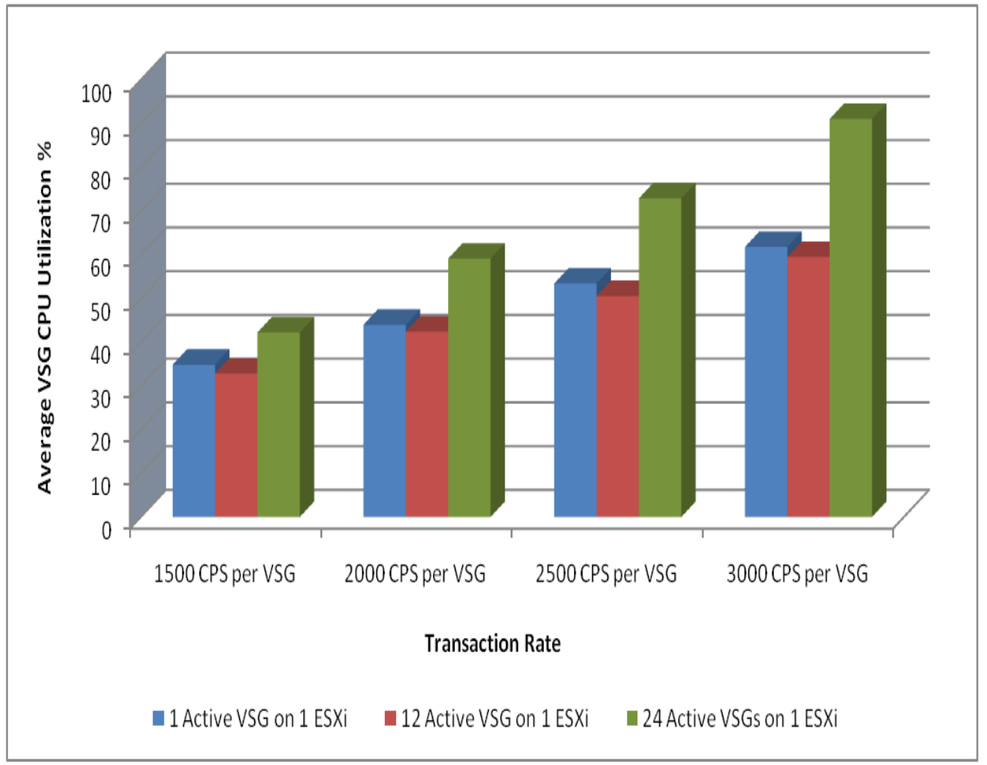


VSG Performance with Logging

© 2013 Cisco and/or its affiliates. All rights reserved.

VSG Performance Results





ESXi host CPU impact with multiple VSG

Perf with multiple VSG on a single ESXi host

© 2013 Cisco and/or its affiliates. All rights reserved.

VSM Port-Profiles for VSG

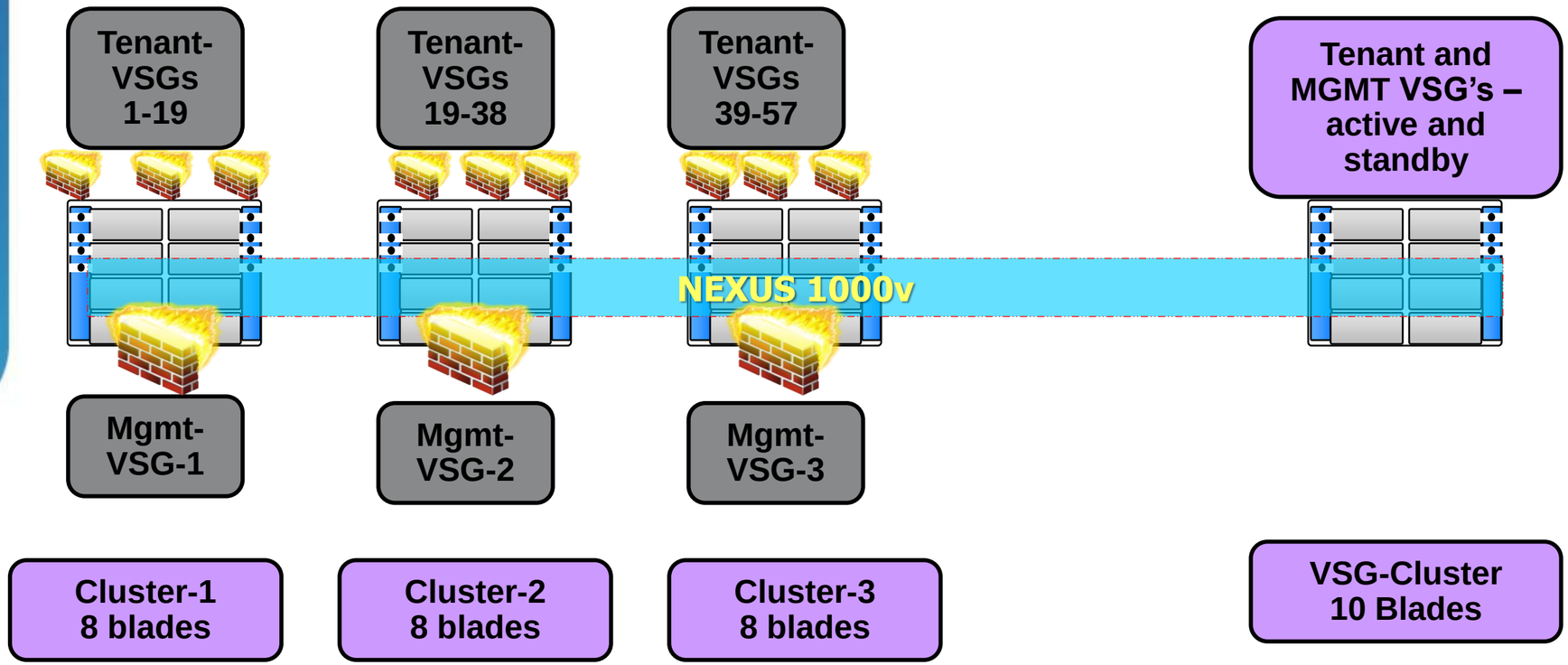
```
port-profile type vethernet vsg-mgmt
  vmware port-group
  switchport mode access switchport access vlan 107
  service-policy type qos input vsg-mgmt
  system vlan 107
  max-ports 64
  state enabled
port-profile type vethernet vsg-data
  vmware port-group
  switchport mode access switchport access vlan 109
  service-policy type qos input vsg-data
  system vlan 109
  max-ports 64
  state enabled
port-profile type vethernet vsg-ha
  vmware port-group
  switchport mode access switchport access vlan 108
  service-policy type qos input vsg-ha
  system vlan 108
  max-ports 64
  state enabled
```

```
ip access-list vsg-to-vnmc
  10 permit ip any 192.168.1.28/32
class-map type qos match-all vsg-mgmt
  match access-group name vsg-to-vnmc
policy-map type qos vsg-mgmt
  class vsg-mgmt
    set cos 6
    set dscp 48
  class class-default
    set cos 0
    set dscp 0
policy-map type qos vsg-data
  class class-default
    set cos 6
policy-map type qos vsg-ha
  class class-default
    set cos 6
vlan 107 name vsg-mgmt
vlan 108 name vsg-ha
vlan 109 name vsg-data
```

VMDC VSG Layout

Each VM has Data vnic and MGMT vnic.
Data vnic monitored by per Tenant-VSG.
Mgmt vnic monitored by shared Mgmt-VSG.

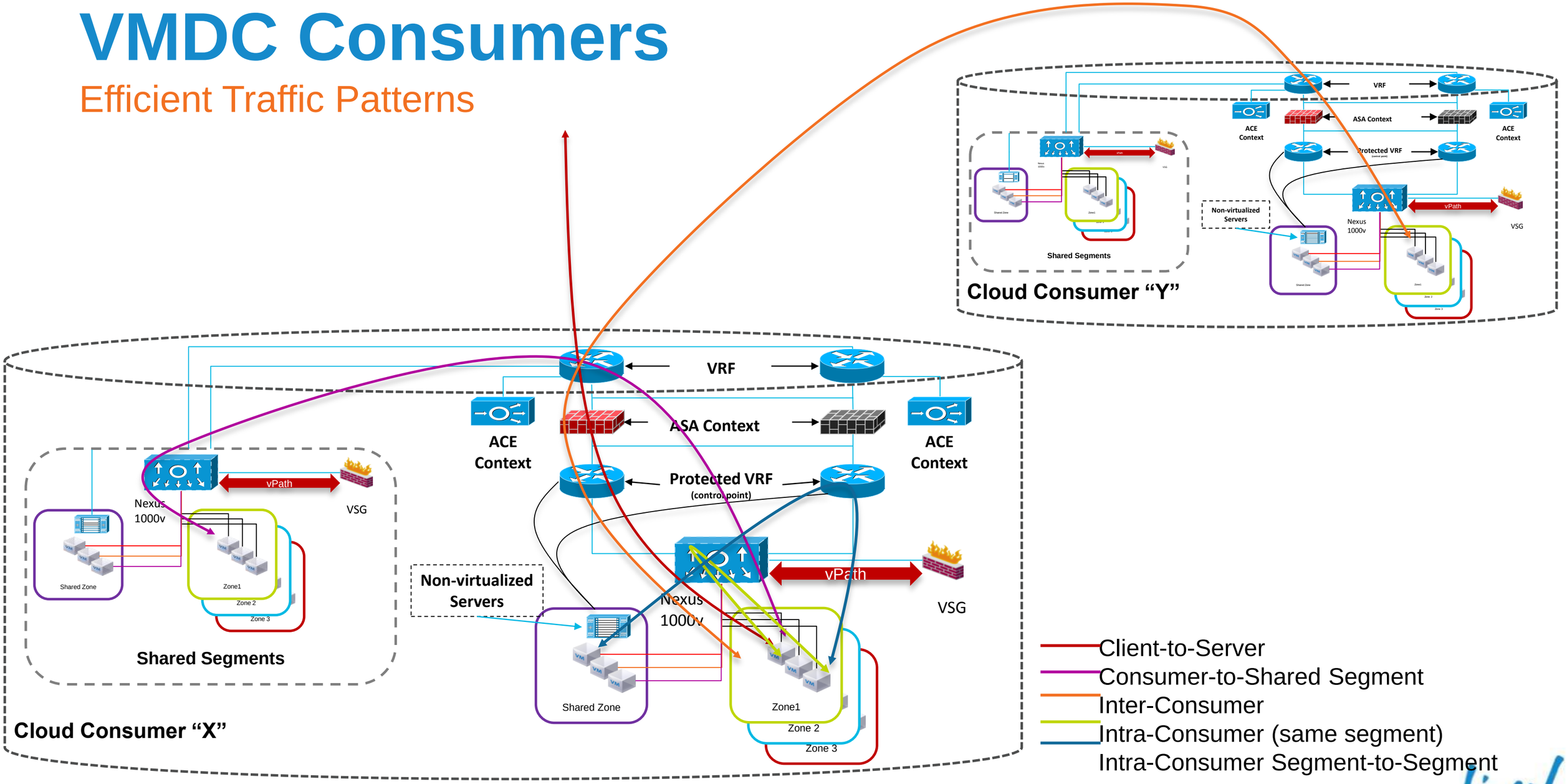
Mgmt-VSG per Cluster, to keep within scale.
VSG's hosted in separate ESXi Cluster.
Tenant Containers mapped to specific ESXi Clusters



© 2013 Cisco and/or its affiliates. All rights reserved.

VMDC Consumers

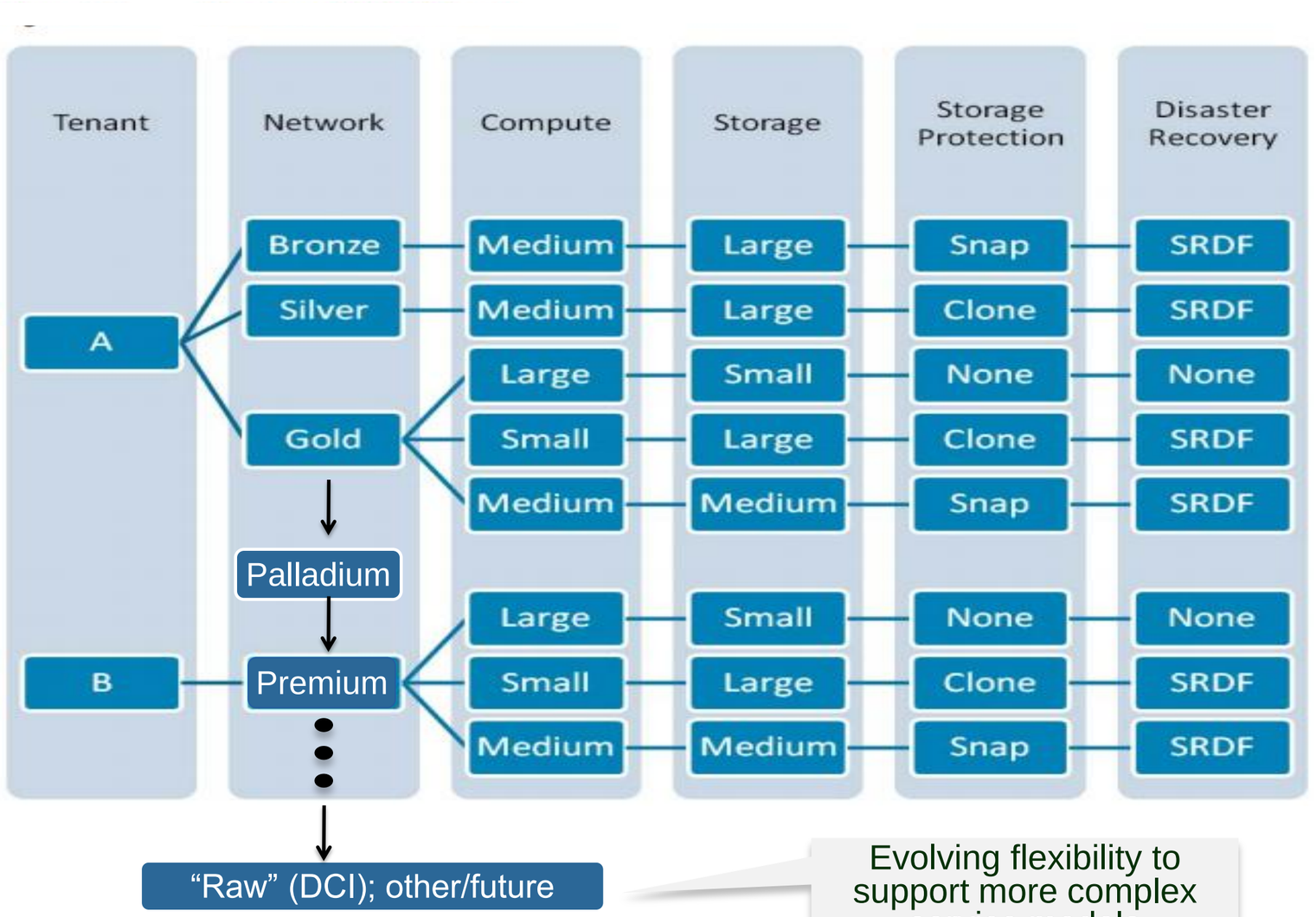
Efficient Traffic Patterns



© 2013 Cisco and/or its affiliates. All rights reserved.

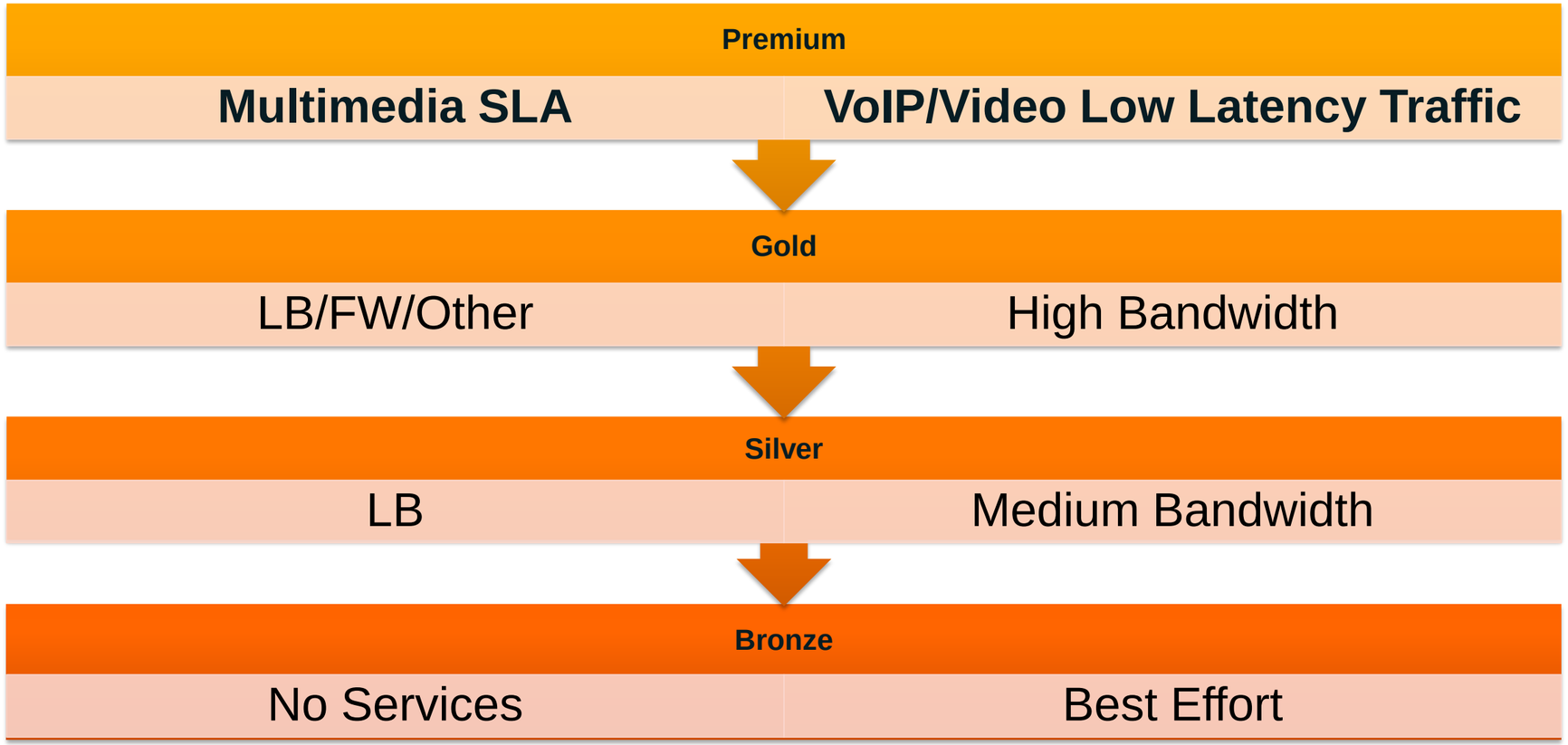
VMDC Service Differentiation

Differentiated Services



© 2013 Cisco and/or its affiliates. All rights reserved.

Model of Differentiated Service "Tiers" in VMDC



Tenants can mix and match to build a complete "data centre" and support multiple application types

© 2013 Cisco and/or its affiliates. All rights reserved.

VMDC Service Level Agreements

Across QoS Domains

- Inter-cloud QoS framework (not shown) but defined in VMDC
- SLA needs to resolve possible contention between Southbound and Northbound SLAs

  e.g. if sum of committed per VNIC/VM bandwidths exceeds corresponding per tenant commitments, the per tenant commitment may not be met

  This is normal for “hose” model SLAs
- SLAs committed from DC edge to edge

  At NGN edge, condition to enforce per-tenant aggregates (policing on ingress to NGN)

  L2 traffic is shaped/policed at aggregate (port) level per class to take a share of NGN bandwidth
- PE Southbound SLA

  SLA per tenant per class

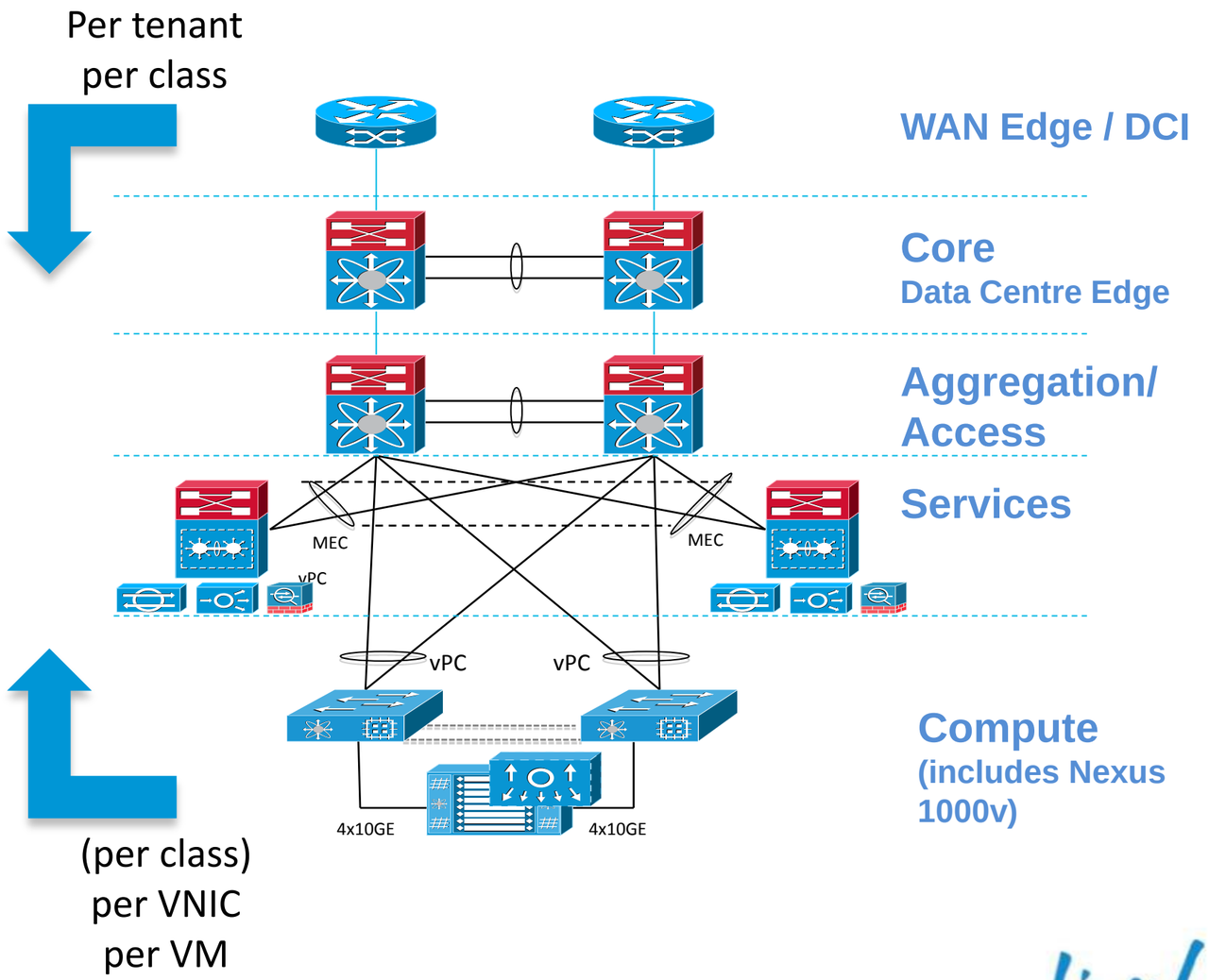
  i.e. aligns with NGN commitment

  HQoS to perform egress shaping to tenant aggregate
- Nexus 1000v Northbound SLA

  SLA per VNIC per VM

  Optionally per class per VNIC per VM

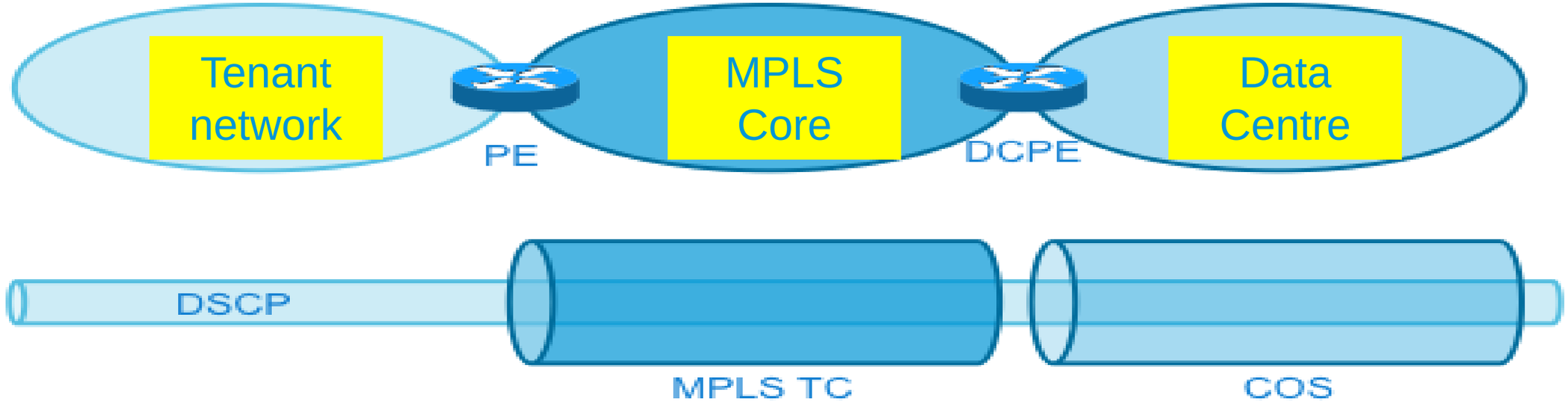
  Ingress policing + CBWFQ on egress (uplinks)



© 2013 Cisco and/or its affiliates. All rights reserved.
Cisco Public

The Hybrid End-to-End View: 3 QoS Domains

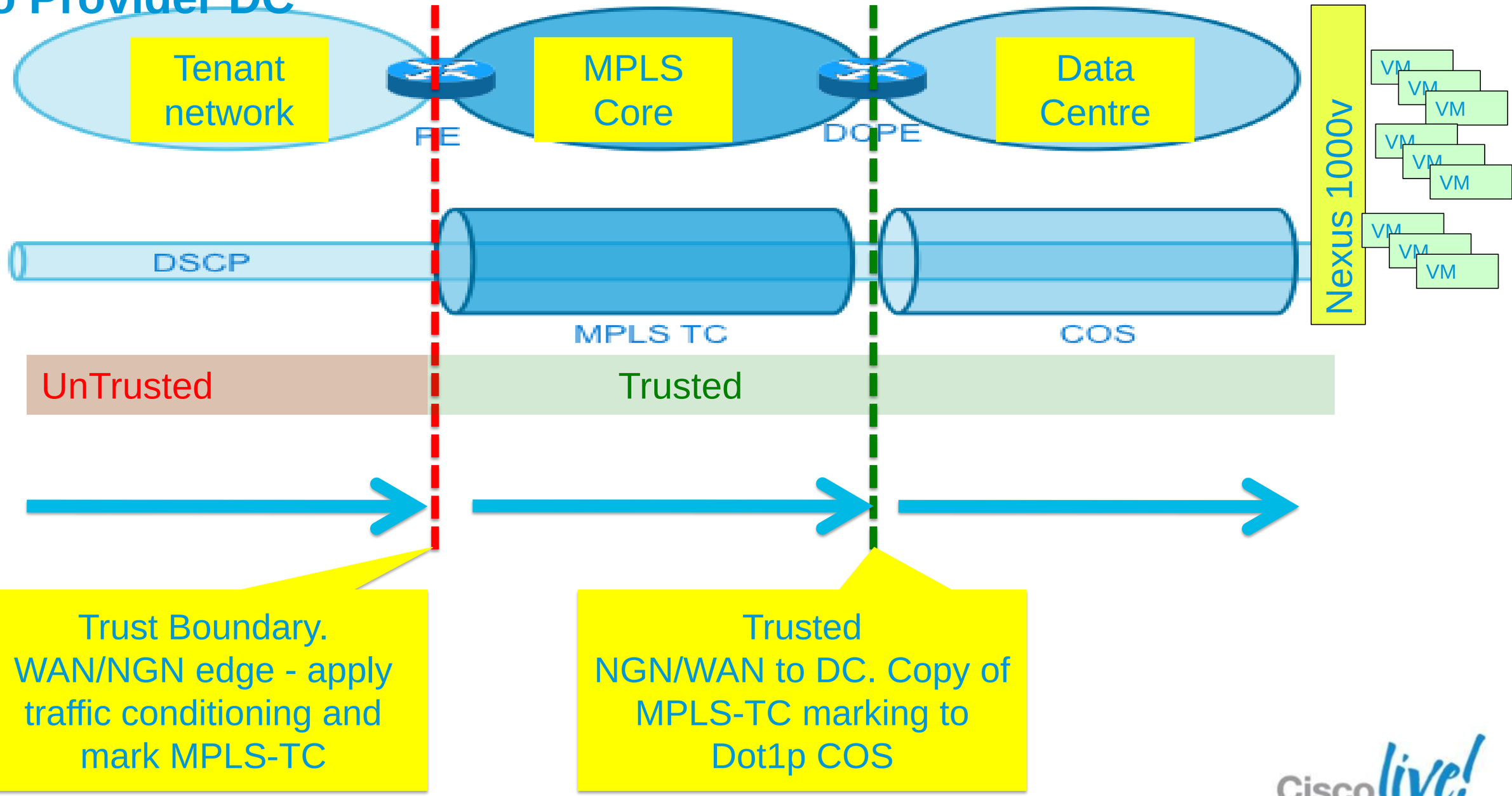
Enterprise Customer	SP-NGN	SP-DC	Public Cloud SP hosted IaaS
Tenant Dept	Enterprise MPLS WAN	Enterprise Data Centre	Private Cloud



- SLAs committed across each domain
- End-to-end SLAs are concatenation of domain SLAs
- Within each domain, the SLA needs to resolve possible contention between inbound and outbound SLAs

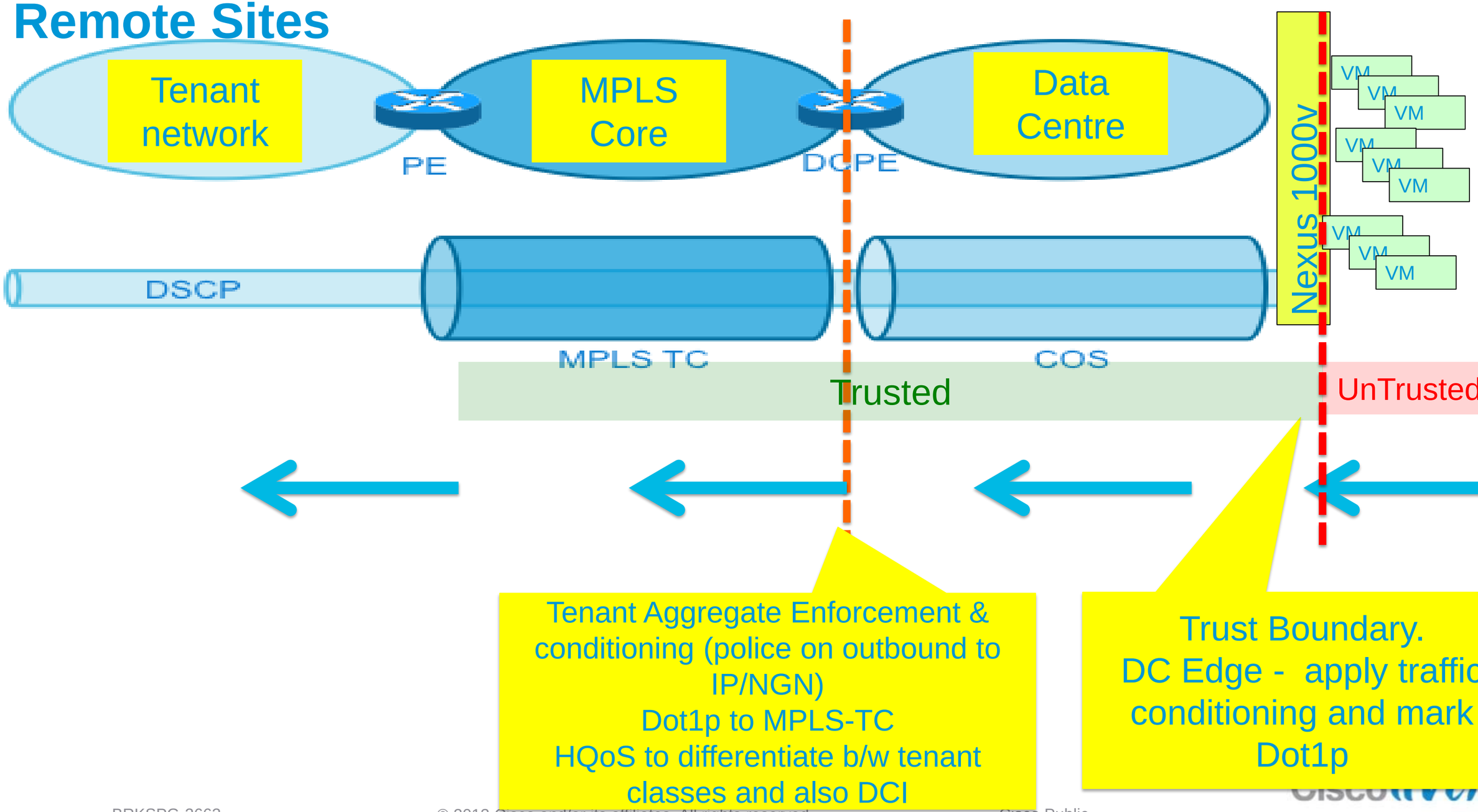
© 2013 Cisco and/or its affiliates. All rights reserved.

Trust Boundaries and Policy Enforcement Points from Remote Site to Provider DC



© 2013 Cisco and/or its affiliates. All rights reserved.

Trust Boundaries and Policy Enforcement Points – DC to Remote Sites



© 2013 Cisco and/or its affiliates. All rights reserved.

QoS Transparency

- QoS Transparency gives the capability of the end user organisation to define QoS markings independently from the DC provider organisation.
- DC provider will have independent markings based on contractual agreements with end-user organisation. The number of Provider classes and QoS treatment is different from the end tenant's QoS requirements.
- IP/DSCP (or ToS) is used by end tenants
- Outer header markings are used by SP/DC provider org. These are the MPLS TC bits or Dot1p COS bits. SP does not touch IP/DSCP/ToS settings and preserves it.
- End to End QoS Transparency requires IP/DSCP and Dot1p COS bits preservation on all platforms. For platforms that do not support QoS transparency, an interim workaround is to remark the IP/DSCP to the Provider setting instead of the tenant settings. This is done at the DC edge platforms - i.e. the PE and Nexus 1000v in VMDC2.2.

© 2013 Cisco and/or its affiliates. All rights reserved.

Multimedia Traffic Classifications (8 Class Model)

Traffic Class	EXP/CoS	DSCP	PHB
Utility Compute Data: Bronze-Standard	0	CS0	Default
Utility Compute Data: Silver-Business to Business & Webex Collaboration Data (Interactive)*	1	CS1	AF
Utility Compute Data: Gold – Business Critical	2	CS2	AF
Storage – FCOE & VoIP Call Control	3	CS3	AF42,AF43
Video Streaming (Future)*	4	CS4	AF41
VoIP Bearer & Video Conference	5	CS5	EF
Network Control	6	CS6	AF
Network Mgmt & Service Control	7	CS7	AF

*Webex , Video Streaming and NFS flows not included in 2.2 test scenarios

- Number of Classes should align with IP/NGN core
- Constrained by the number of CoS markings available (8)
- Constrained by the number of queues/thresholds supported by DC platforms

© 2013 Cisco and/or its affiliates. All rights reserved.

Class to Queue Mapping

VMDC 8 class model	COS	VMDC HCS Aligned 8 Class Model	VMDC NGN Aligned 8 Class Model	VMDC (UCS 6xx0) 6 class model	HCS 6 class model	4 class model N7k fabric
Network Mgmt + Service control	7	Network Mgmt + VM control	Network Mgmt + VM control	Network Mgmt (COS 7) + Service control (COS 7) + Network control (COS 6)	Network Mgmt (COS 7) + Service control (COS 7) + Network control (COS 6)	Queue 1
Network control	6	Network control	Network control			
Priority #1	5	Voice bearer	Res VoIP / Bus Real-time	Priority #1	Voice bearer	
Bandwidth #1 (Priority 2)	4	Interactive Video	Video streaming	Bandwidth #1	Interactive Video	Queue 2
Bandwidth #2	3	Call Control	Video interactive / FCOE	FCOE (Bandwidth #2)	Call Control	
Bandwidth #3 “Gold”	2	FCOE	Bus critical in-contract (COS 2) Bus critical out-of-contract (COS 1)*	Bus critical in-contract (COS 2) Bus critical out-of-contract (COS 1)*	FCOE	Queue 3
Bandwidth #4 “Silver”	1	Webex collaboration data (interactive)	Silver In-contract (COS 2) Out-of-contract (COS1)	Silver In-contract (COS 2) Out-of-contract (COS 1)	Webex collaboration data + Standard data	Queue 4
Standard (Bandwidth #5) “Bronze”	0	Standard data	Standard data	Standard		

* Different drop thresholds for in- and out-of-contract

© 2013 Cisco and/or its affiliates. All rights reserved.

Bandwidth Reservations

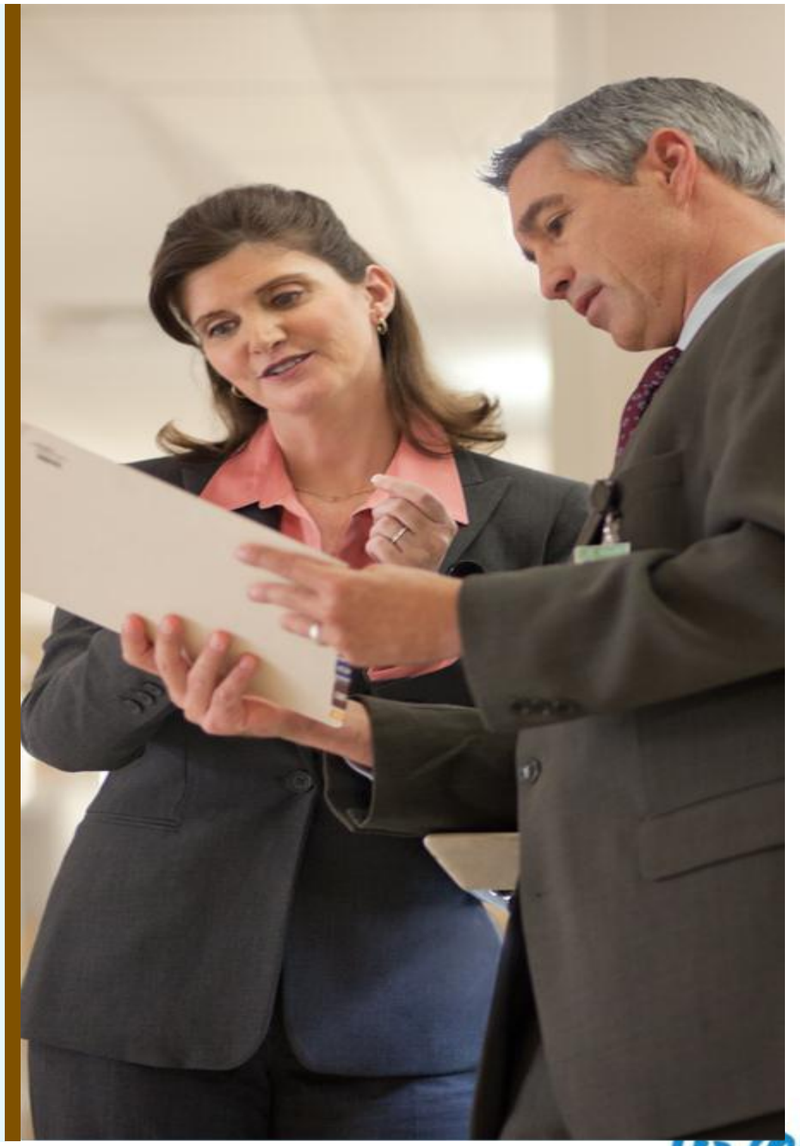
Traffic Class	EXP/CoS	BW Reserved (Remaining After Priority)	Actions
Utility Compute Data: Bronze-Standard	0	15% (17%)	WRED
Utility Compute Data: Silver-Business (In/out of contract) & Webex Collaboration Data (Interactive)*	In (COS 2) Out (COS 1)	60% (70%)	WRED Out of Contract dropped before in contract
Utility Compute Data: Gold – Business Critical (In/Out of contract)	In (COS 2) Out (COS 1)		WRED Out of Contract dropped before in contract.
Storage – FCOE & VoIP Call Control	3	3% (4%)	
Video Streaming (Future)*	4	x% (x%)	WRED, egress policing per tenant
VoIP Bearer & Video Conference	5	15%	Priority, egress policed per tenant
Network Control	6	4% (5%)	
Network Mgmt & Service Control	7	3% (4%)	

© 2013 Cisco and/or its affiliates. All rights reserved.

CloudVerse Megatest (LightReading / EANTC)

http://www.cisco.com/en/US/solutions/ns341/eantc_cloud.html

- **VMDC 2.2 based architecture** – each test overlaid as tenant in multi-tenant cloud
- 70+ 10G IXIA ports, 75+ VRFs/tenants, 600+ VLANs, 1500+ VMs
- **80 Gbps** of north-to-south (next-generation network [NGN] to cloud) traffic: 1 million clients to 50,000 servers
- **300 Gbps** of east to west (within data centre) traffic: switched & routed - with Cisco® FabricPath 2-tier design – showing 15,000 MAC addresses and 256 VLANs
- **67 million** NAT64 sessions simulated, at 80Gbps, 4 million/sec
- 1 million residential gateways shown for 6rd, at 80Gbps
- 40Gbps of video streaming – using Cisco CDS Internet streaming appliance, and on a Cisco ASR 9000 Series Cisco Integrated Services Module (ISM)
- PCRF for throttling mobile subscribers in real time
- Any video format, any device, any where: iPad, Android, PC, etc.
- Cisco VM-FEX in VMDirectPath performance demonstrated:
  - VM-FEX compared to software DVS in 4 ways: iSCSI read-write, L3 IMIX traffic, HTTP traffic, and video encoding
  - **20 to 30% performance improvement** in throughput, CPU, and IOPS with Cisco Data Centre VM-FEX DirectPath I/O

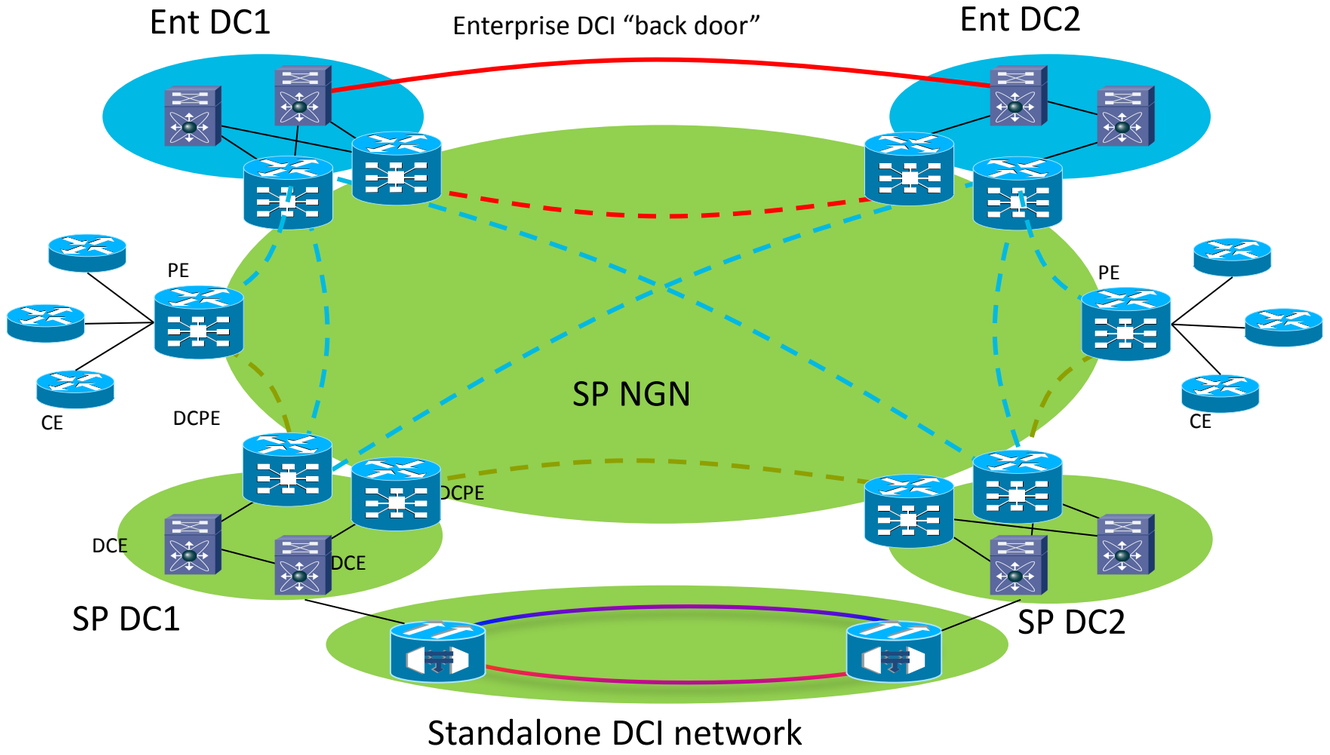


© 2013 Cisco and/or its affiliates. All rights reserved.

Data Centre Interconnect

Data Centre Interconnect

- Interconnection models:
  - Enterprise to Enterprise (E2E)
  - Enterprise to Service Provider (E2SP)
  - Service Provider to Service Provider (SP2SP)
- Overlay based techniques
  - OTV, LISP, VXLAN
  - Suitable for intra-Ent DC interconnect
- NGN based DCI solution:
  - Addresses E2SP for workload migration
  - Addresses SP2SP for regional or distributed data centres
- Standalone DCI network
  - Provides interconnection between main SP DCs
  - Owned by SP DC team
  - Addresses SP2SP only
  - Very high bandwidth – packet / optical solution likely the most cost effective



	Overlay solution	PE-based solution
Ethernet	(e)TRILL /802.1ad Future	
MPLS		VPLS, A-VPLS, EVPN, EoMPLS
IP	OTV, LISP	

© 2013 Cisco and/or its affiliates. All rights reserved.

VPLS Attachment Circuit Redundancy Options

MC-LAG vs. MST-AG vs. nV

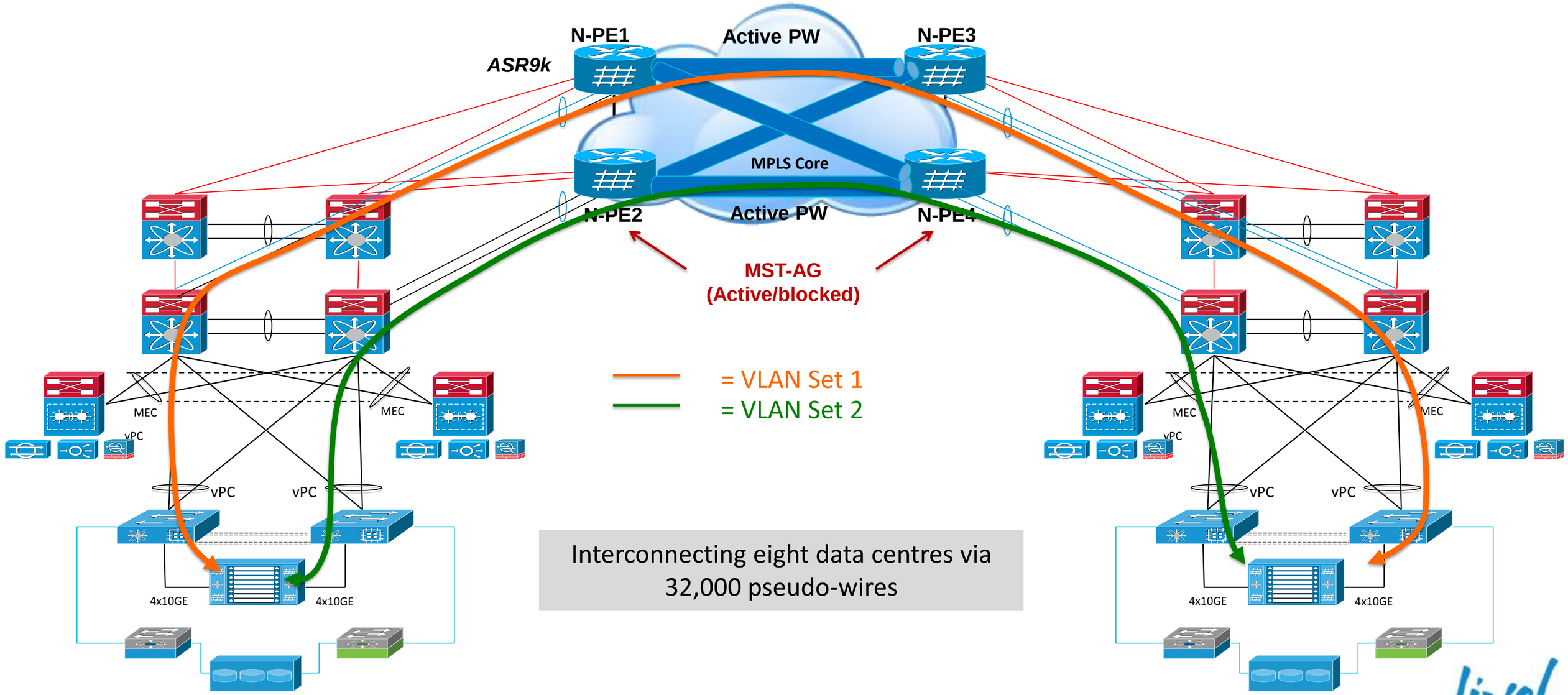
MST-AG	• Standard based solution as long as access network support MST/PVST • Works for any access network topology but requires to move the STP Root • Good access domain isolation • Work with 802.1ah PBB • Convergence time depends on access network STP
MC- LAG	• Simple solution for spoke-and-hub topology, works for both bridging and non-bridging access device • Standard based solution by using 802.3ad • Sub-second convergence • Can be active/standby mode, or Act/Act with Per VLAN load balancing • MC-LAG to vPC is optimal for redundancy, • VLAN scale limits per vPC on Nexus7000 – suitable for less than 2000 VLANs
nV Cluster	• Dual ASR9000 appear as single Virtual Chassis (Act/Act) • Synchronised service state between the two nodes • Improved convergence times (<50 msec), independent of service scale • Improved L3 scalability (single routing adjacency across the 2 nodes) • Support for ISSU

VMDC High Scale DCI Usecases

- VPLS to provide multipoint LAN extension functionalities
- SP Data Centres Design
  - ASR 9000 as PE devices
  - MST-AG for attachment circuit redundancy, resilience and load balancing
  - VLAN Translation for operational efficiency and scale
- vMotion for live workload migration between sites
  - ESX/N1k extended on both sides
  - VC, VSM, VNMC deployed on the same Data Centre site
  - Storage interconnect and replication needed
- Cold Migration for Server and VM Maintenance purposes
  - Storage interconnect and replication needed
  - Migration tools like vCloud Director, VM cloning utilities

© 2013 Cisco and/or its affiliates. All rights reserved.

VMDC SP-SP DCI Topology



Interconnecting eight data centres via 32,000 pseudo-wires

© 2013 Cisco and/or its affiliates. All rights reserved.

VMDC VPLS DCI: Validated Scale

VMDC Item	Scale	ASR 9000	Nexus 7000
# Customers	2000	BD: 2000	VLANs: 4000
VLANS/Customer	2	EFP: 6000 4000 for N7k 2000 for IXIA to add MAC scale	VDC: 2, 2000 vlan/VDC 1 VDC/Line Card exclusive Vlan-port:28k/VDCTotal: 56k
Sites	8 Sites	16 remote PE BGP AD: 16 peers	-
Pseudo-Wires		32000	
MST-AG	2	2 (1 per 1000 cust/2000 vlan)	-
MSTI	8 per MST-AG	8 per MST-AG Total : 16	8 per VDC Total 16
MAC ***		470k (1 node fail) A9K-1 = 331K, A9K-2 = 335K (normal)	128k (1 node fail) per VDC, 256k total N7K-1 = 62K, N7K-2 = 66K (normal)
Line card		1 for DC, 1 for WAN A9K-8T-E	1 per VDC M132
Port-Channel		1 bundle/1000 customers (2x)	1 uplink(2x), 8 downlink (1x), 12 crosslink (1x) – to keep vp count high

*** normally traffic is load balanced hence nodes learn half the # MAC. Node fail causes all MACs on surviving node

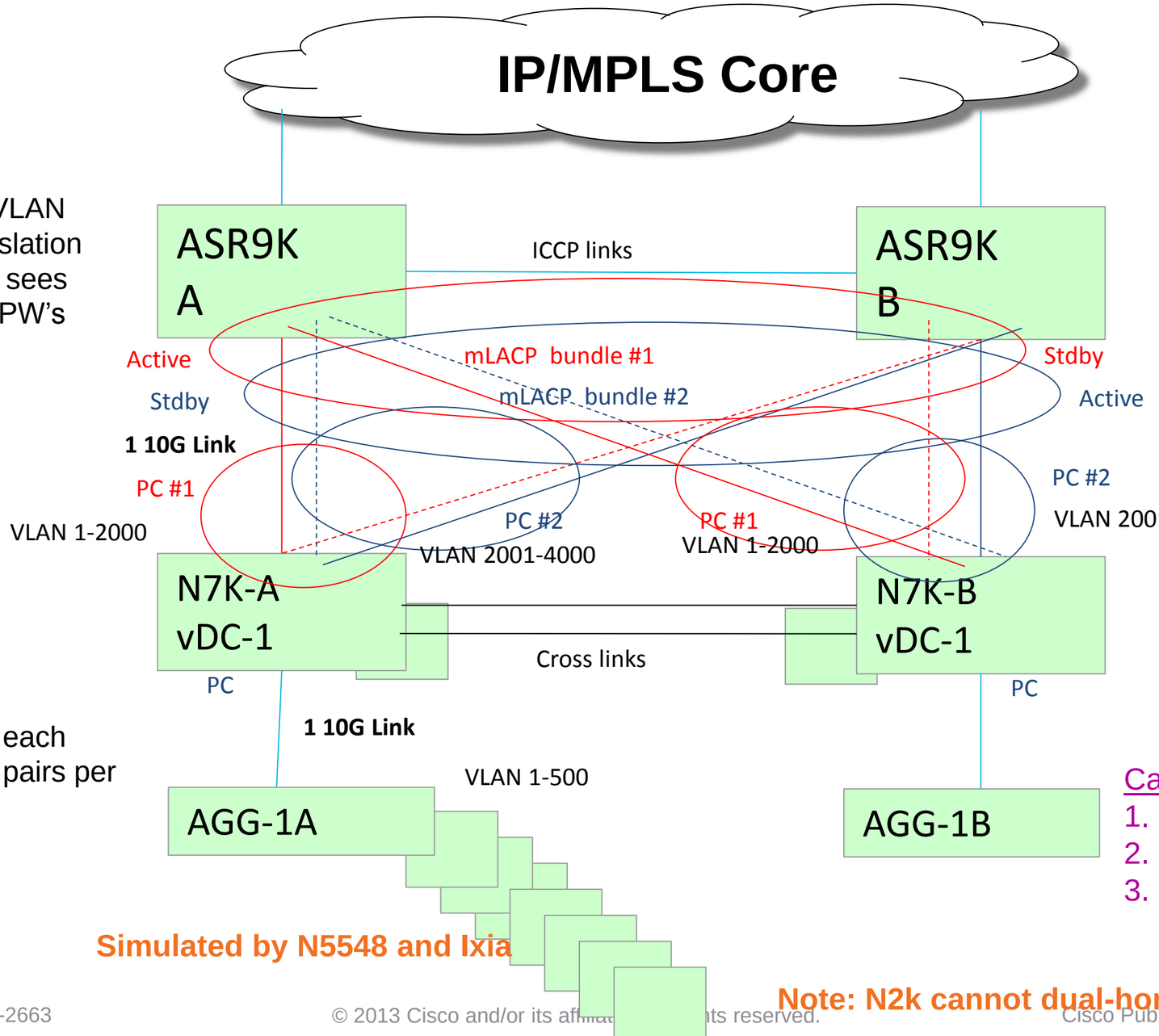
© 2013 Cisco and/or its affiliates. All rights reserved.

VMDC VPLS DCI: Test Results

Test Type	Test Results
ASR9000 Node Failure	Convergence time is between 3.1 to 4.5 in Chassis Reload - Down
	Convergence time is between 1.2 to 1.7 sec in Chassis Reload - Up
Nexus7000 Node Failure	Convergence time is typically less than 5 sec
	Node failure is impacted by CSCty00169 and CSCth37522 will address this.
Link Failure	Link Failure Converges within 3-5 second
	ASR9k sends MAC withdraw for all affected BD
	Link Restore Convergence is good
	For bidirectional flows, both directions have same convergence time
	Additional convergence delay up to 8-9s when traffic crosses from one N7k to other due to CSCtt96971

VPLS DCI Scale Testbed

2000 BD's (no QinQ), 4 VLAN per customer. VLAN translation for all customers. ASR9k sees 8,000 vlans, and 32,000 PW's (mesh to 8 DC sites)



All PC's have 1 10G link to each device (need more ports to scale higher)

2 vDCs 4000 VLAN per vDC, 8,000 total
2 upstream PC to ASR9k (distribute load)
Showing 1 vDC in diagram

Assuming 500 VLANs to each downstream pair. 8 such pairs per N7k.
So 8 downstream PC's.

Simulated by N5548 and Ixia

Can be:
1. N2k's for FEX to 1/10G servers
2. UCS6100s to 10G blades
3. N7k's (Agg N7k's connecting to pair of DCI N7ks)
   Will simulate test with IXIA

Note: N2k cannot dual-home to N7k

© 2013 Cisco and/or its affiliates. All rights reserved.

VMDC 2.2 DCI: Key Takeaways

- ASR1k Attachment Circuit Redundancy
  - Currently does not support MC-LAG for Attachment Circuit Redundancy
  - ASR9k or 7600 can be positioned in the meantime for redundant designs
- VPLS scaling with MC-LAG and vPC
  - Active/Standby Technology
  - Recommendation on max VLAN scalability limited to 1200 because of vPC current limitation (CSCtd80191 causing long convergence time with higher VLAN number)
- VPLS Deployment with MST-AG
  - Dependency on Spanning Tree (Root must be moved)
  - Traffic load-balancing per VLAN set
  - Convergence time can be up to 10-15 sec (worst case)

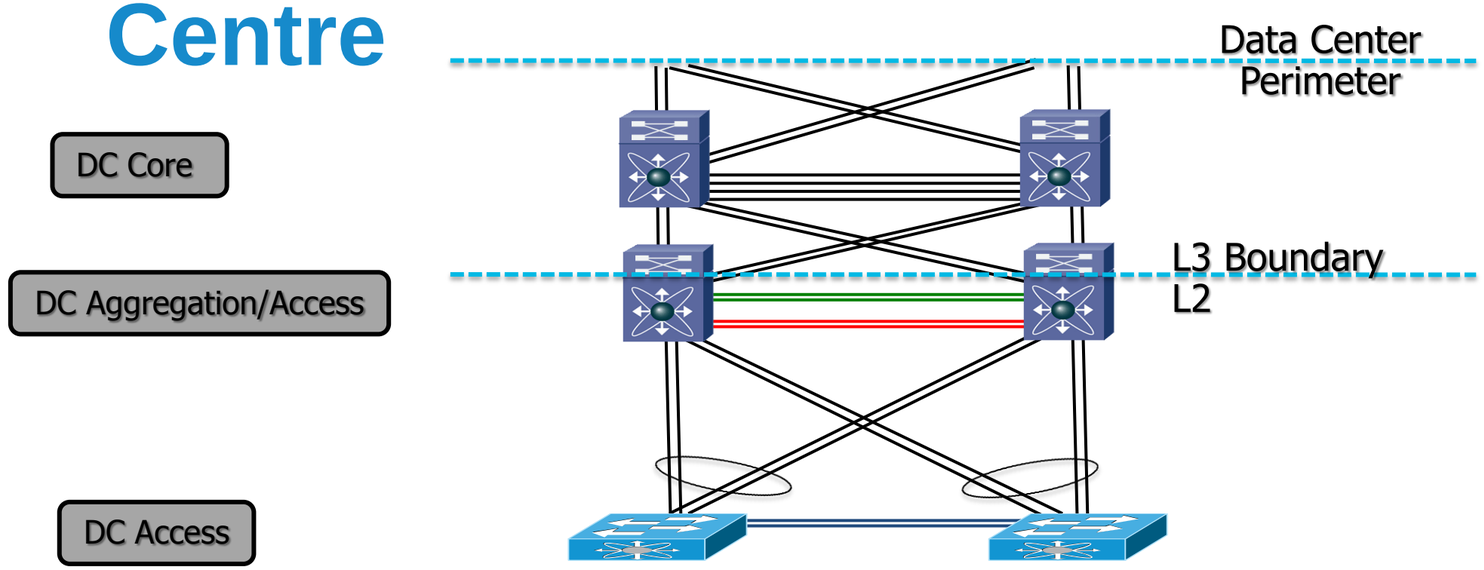
    2 VDCs, 4000 VLANs, 40k+ VLAN ports validated

    Convergence will be faster in smaller deployment
- nV Clustering will be recommended solution moving forward
  - Better scalability (fewer pseudo-wires needed for redundancy, fewer L3 adjacencies needed)
  - Faster convergence

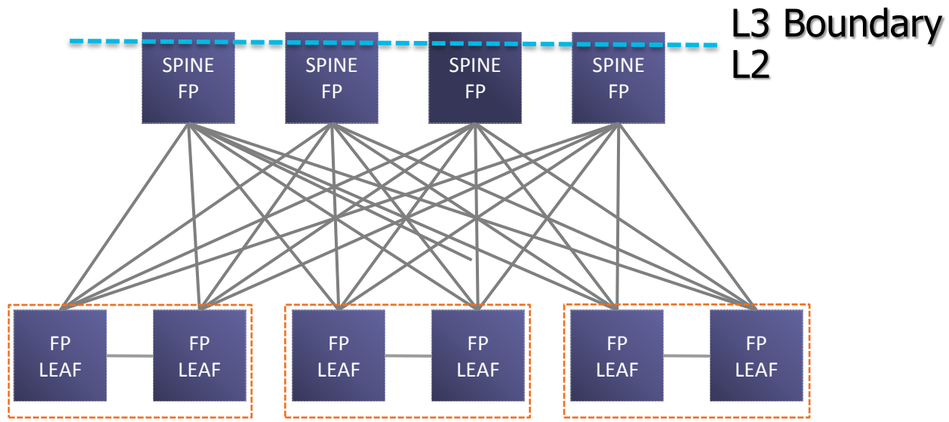
© 2013 Cisco and/or its affiliates. All rights reserved.

FabricPath L2 Data Centre Designs - VMDC 3.x

CLOS Based Model as a New L2 Option for the Data Centre



Multi-Layer L3 Hierarchical Design



L2 CLOS Design

- VMDC 2.x releases validated topology variants (i.e., collapsed core/aggregation, as the L2/L3 boundary)
- VMDC 3.x releases validated with FabricPath based architecture – for intra-POD or inter-POD VM Mobility

© 2013 Cisco and/or its affiliates. All rights reserved.

VMDC Fabric Path Design Considerations

Port density and oversubscription

How well does the design scale to support more edge ports?

What level of oversubscription is acceptable?

L2/L3 boundary options (localised or at aggregation-edge nodes)

Where will the L2/L3 boundary be placed?

How much capacity for inter-VLAN and N↔S routing does it provide?

Flexibility and scope

How flexible is the design with respect to VLAN extension, server mobility, direct communication paths, etc.?

System Resilience

Reconvergence

Host scale per FabricPath domain

What is the total number of unique MAC entries and/or ARP entries the design can support?

Fabric Path Modules – M1/F1 (Mixed VDC, and Split VDC)

Conversational MAC address learning

System Scale

VLAN scale: constrained by HSRP and GLBP FHRP

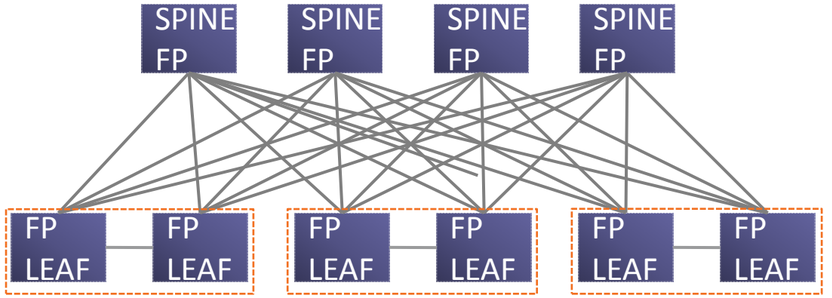
ARP learning rate

Service Attachment Options

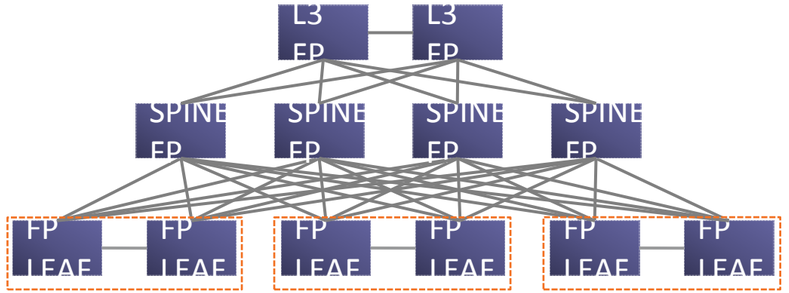
Local, Central, Distributed

Hybrid Physical/Virtual

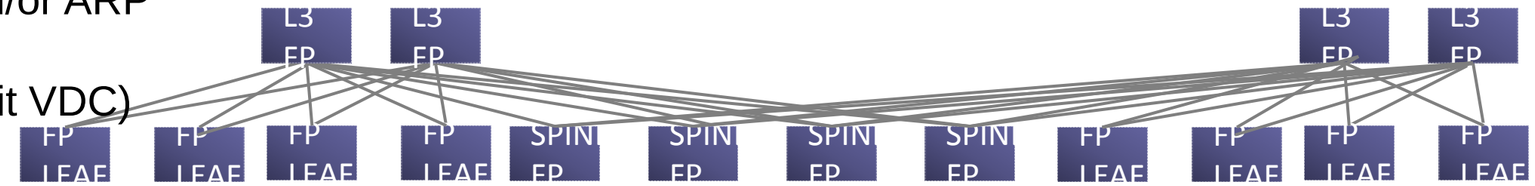
How to handle resilience and failover?



Typical DC
Base
Base with VPC+
L3 Leaf



Switched DC
Dedicated Spine

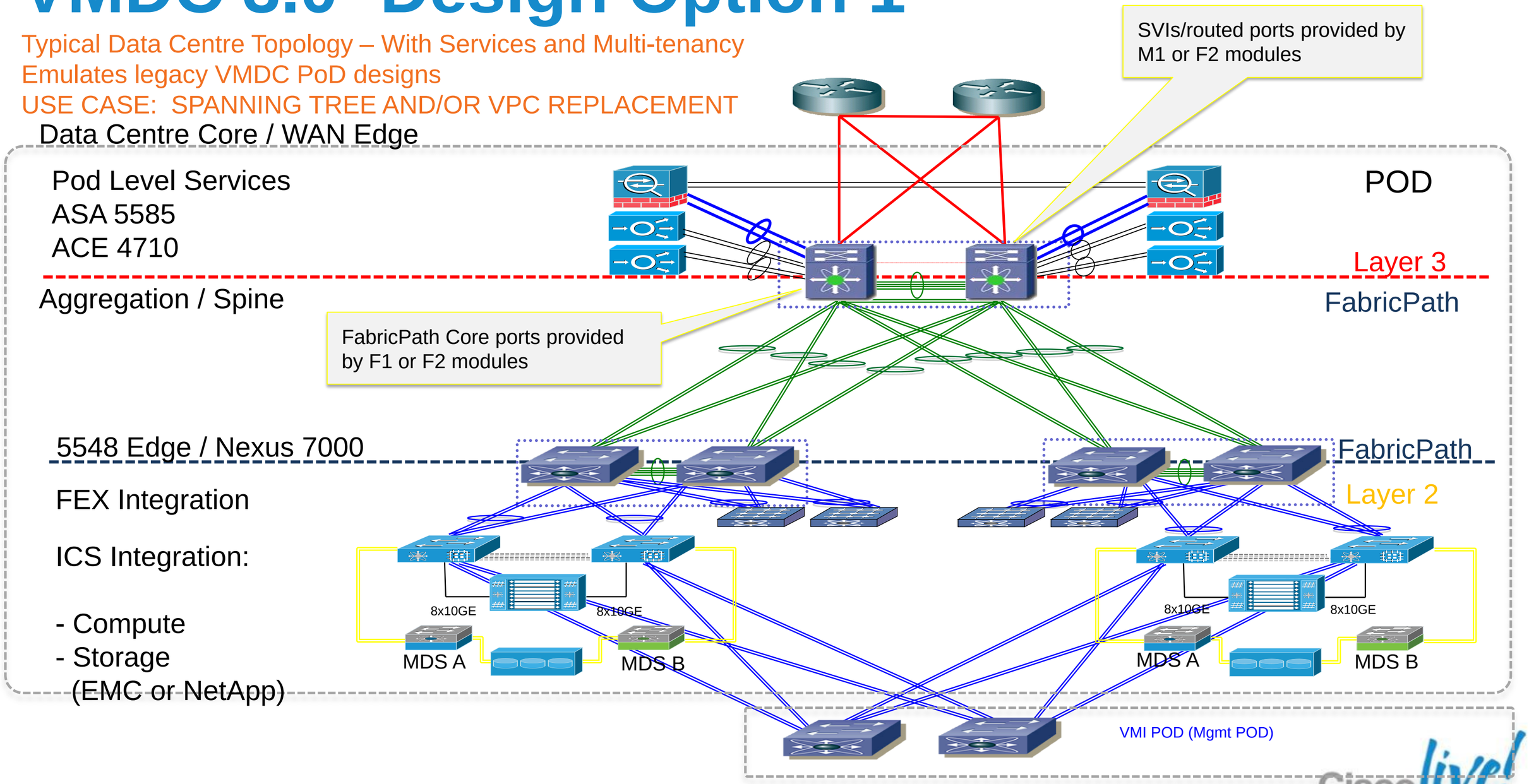


Extended Switched
Dedicated L2 Super-Spine
Inter-pod/Inter-Building

© 2013 Cisco and/or its affiliates. All rights reserved.

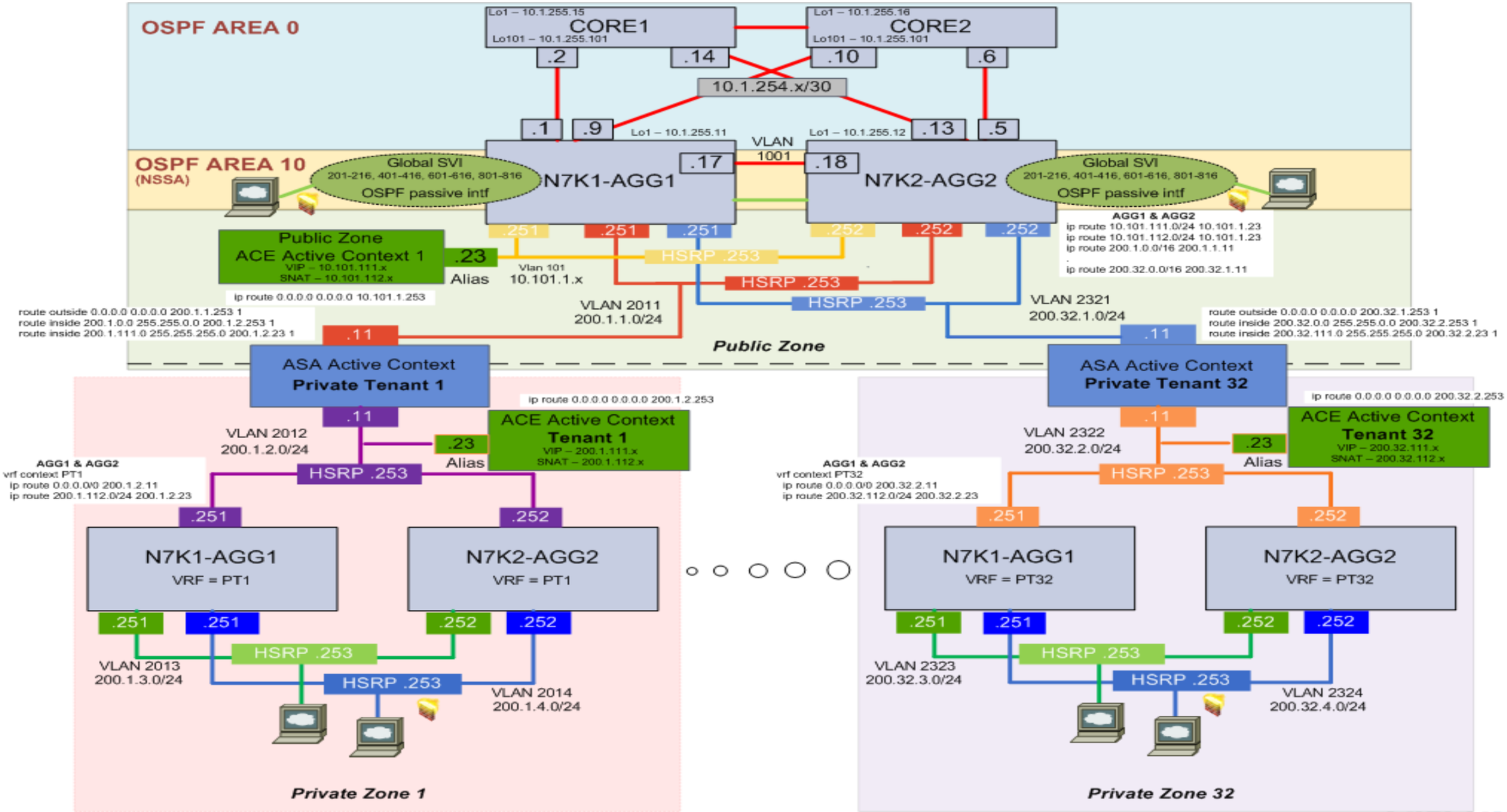
VMDC 3.0- Design Option 1

Typical Data Centre Topology – With Services and Multi-tenancy
Emulates legacy VMDC PoD designs
USE CASE: SPANNING TREE AND/OR VPC REPLACEMENT



© 2013 Cisco and/or its affiliates. All rights reserved.

Logical Tenancy Topology

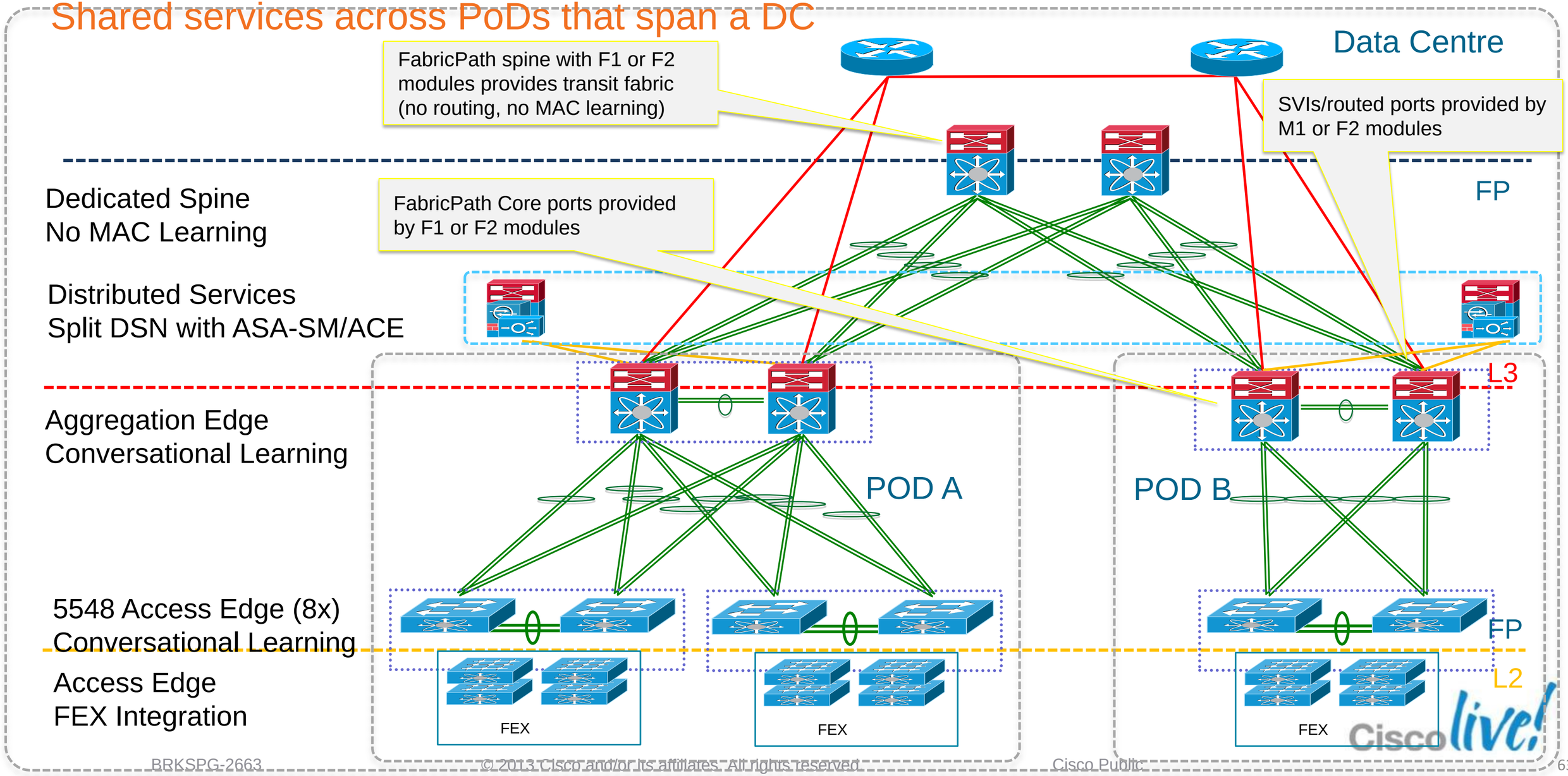


© 2013 Cisco and/or its affiliates. All rights reserved.

VMDC 3.0- Design Option 2

Extended Switched Fabric Data Centre Topology

Shared services across PoDs that span a DC



© 2013 Cisco and/or its affiliates. All rights reserved.

Scalability Considerations

- Aggregation-Edge
  - M1-F1 Mixed VDC with vPC+ has a limitation of 16K MAC currently. This means <16K hosts in the POD. NX-OS 6.2 release will have up to 128K MAC support in mixed VDC.
  - 500 HSRP interface per M1 module. (*tested with 200 HSRP interfaces*).
  - Aggressive PIM timer support is limited (*tested 5sec hello timer on 64 interfaces with 400 mroutes*).
  - ARP learning rate @ 200-300pps on M1/F1 mixed VDC. (vPC+ with ARP sync is recommended to improve unicast convergence).
  - Cisco Nexus 7000 Series NX-OS Verified Scalability Guide: *http://cco/en/US/docs/switches/datacenter/sw/verified_scalability/b_Cisco_Nexus_7000_Series_NX-OS_Verified_Scalability_Guide.html*
- Services
  - ACE 4710 supports up to 20 contexts, need additional pair of ACE 4710s as contexts grow.
  - <4G throughput per ACE pair considering HA.
- Access-Edge
  - 24K MAC entries (32K in hardware) per N5500 Access pair. This includes local and remote MAC.

© 2013 Cisco and/or its affiliates. All rights reserved.

Convergence Results Summary

Typical DC model

- Most Nexus 7000/ASA/ACE related failure and restoration scenarios yielded sub-4 sec unicast convergence
- AGG restoration has up to 20sec convergence for Load balancer and HSRP GW bound flows.
- L3 multicast failover times are mostly influenced by PIM Hello timer. *(with 5sec pim hello, convergence is <20sec during PIM DR failure)*
- The Nexus 5500 access-edge failure and restoration scenarios yielded sub-second unicast/multicast convergence.

Extended DC model

- Unicast convergence numbers are mostly in the range of 30-60 sec for the most rigorous (i.e., Agg. Edge node) failure cases.

© 2013 Cisco and/or its affiliates. All rights reserved.

Virtualised Services based DC
- VMDC 4.x

Trend Towards Virtualised Services

- Insertion of services {load-balancing, firewalling, tenant routing} within the tenant container fundamentally drives the logical design (both L2 and L3) within the data centre
  - Services are typically L3
- Industry transition underway from network-based services to virtualised services
- How should we provide logical connectivity between a tenant's virtual data centre container and their private cloud and/or remote users, given this transition?
- VMDC 4.x focus
  - Address transition to virtual services
  - Address tenancy scalability constraints of current solution
  - Routing as a Service (RaaS) for Cloud providers
  - Highlight service chaining considerations/issues
  - Highlight new scalability considerations (virtual appliances in compute tier)

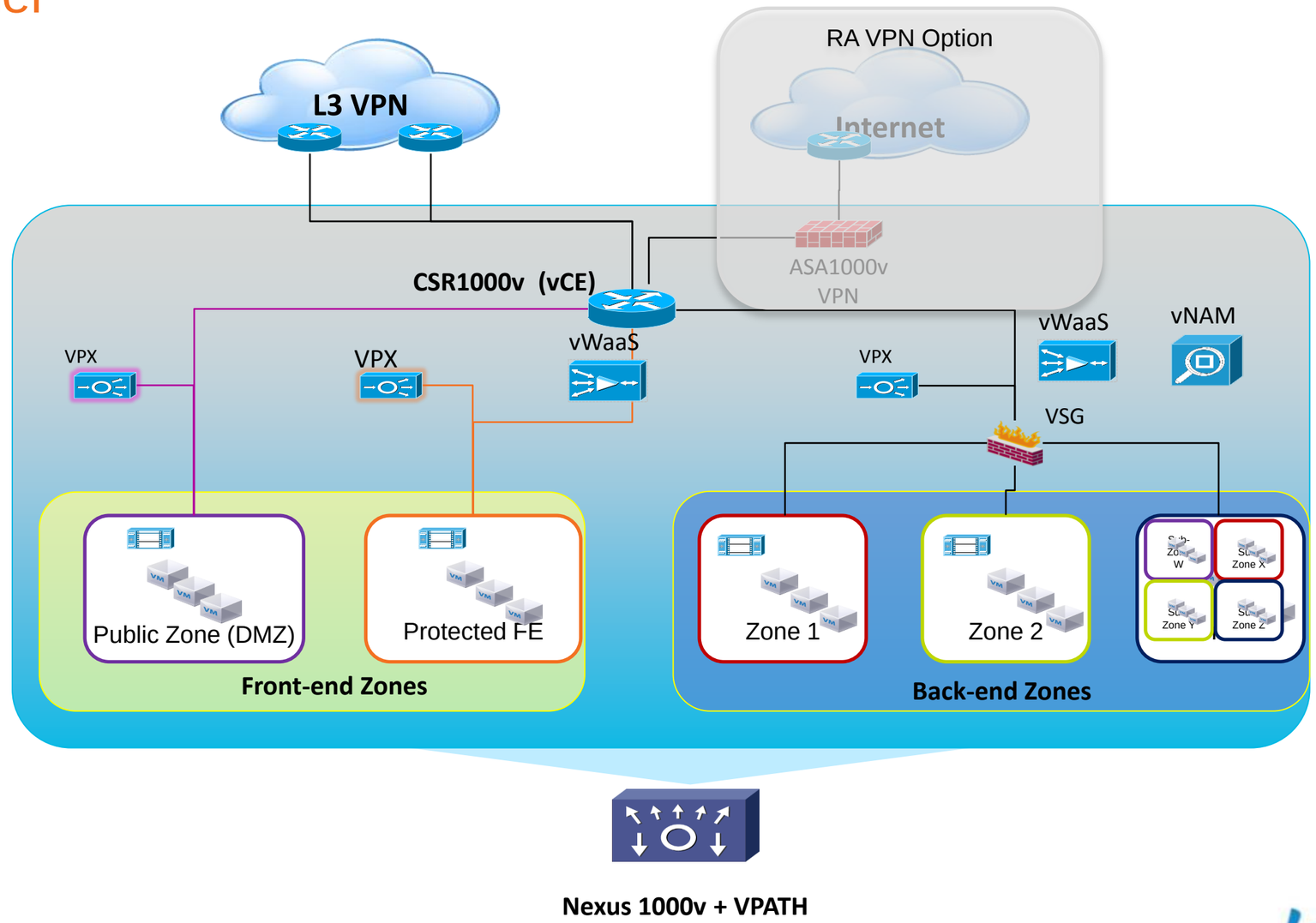
© 2013 Cisco and/or its affiliates. All rights reserved.

VMDC vCE Architecture:

Sample Virtualised Container

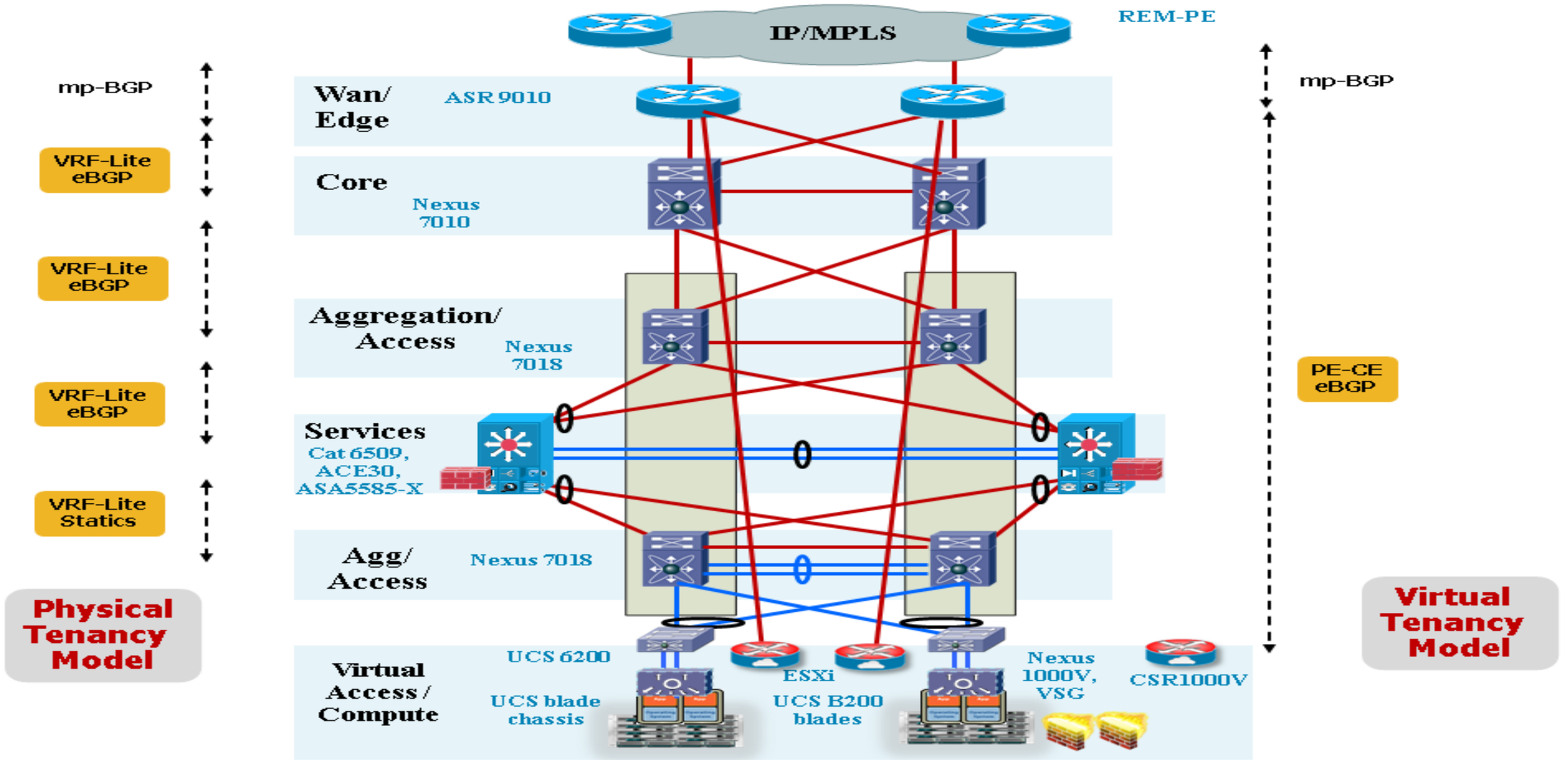
Components:

- CSR XE 3.9
- Netscaler VPX
- vWaaS 5.2 (vPath and WCCP redirection)
- vNAM 6.0
- VXLAN on N1kV



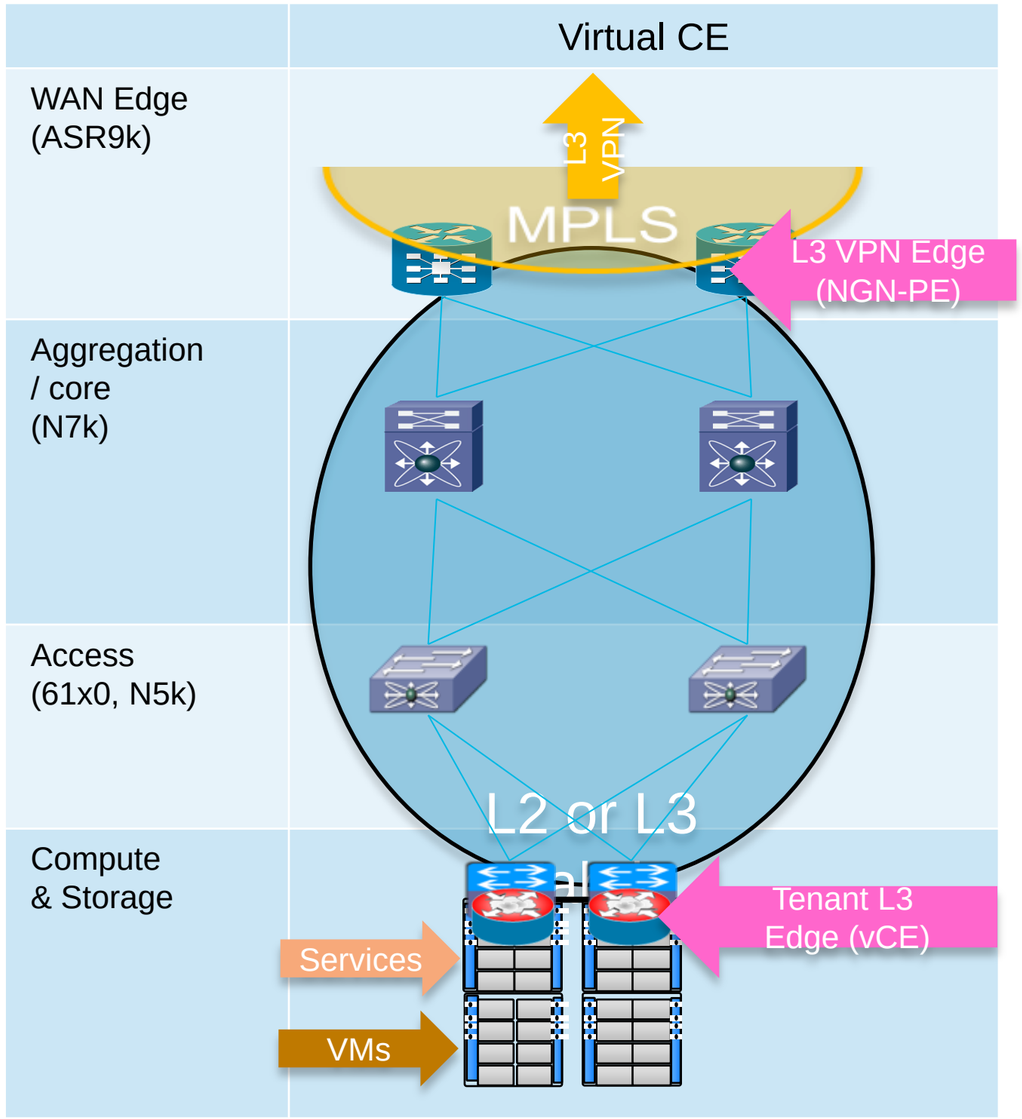
© 2013 Cisco and/or its affiliates. All rights reserved.

Impact on Intra-DC L3 Design



© 2015 Cisco and/or its affiliates. All rights reserved.

Virtual CE: Benefits



- Alleviates need to extend L3 VPN natively into the data centre
  - E-W L3 via vCE
  - E-W L2 extension via L2 overlay (VXLAN)
- PAYG solution – virtual CE per tenant follows same model for tenant routing as for other tenant services, i.e.. RaaS
  - Could be multiple vCEs per tenant
- Mirrors branch CE model, i.e. can support same features and management models
  - Allows for end-to-end services with enterprise sites (WaaS, LISP, IPSEC, etc)
- No cross-tenant dependencies, simplifies change management etc.
- Simplifies management and orchestration
  - Cisco working through dynamic PE VRF provisioning models
- Requires scalable DC WAN gateway and PE-vCE segmentation technology

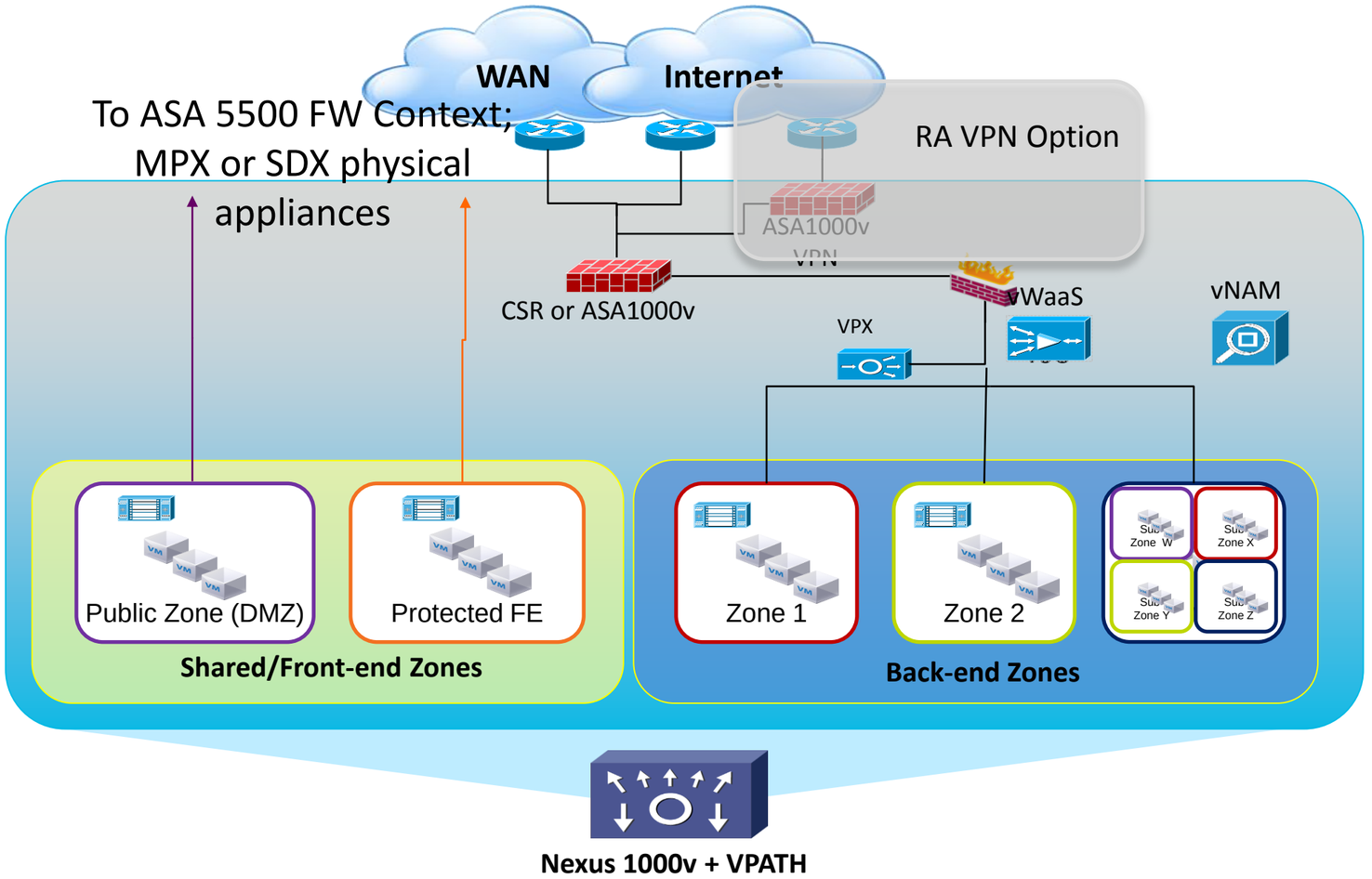
© 2013 Cisco and/or its affiliates. All rights reserved.

VMDC Hybrid Physical/Virtual Services Architecture:

Sample Virtualised Container

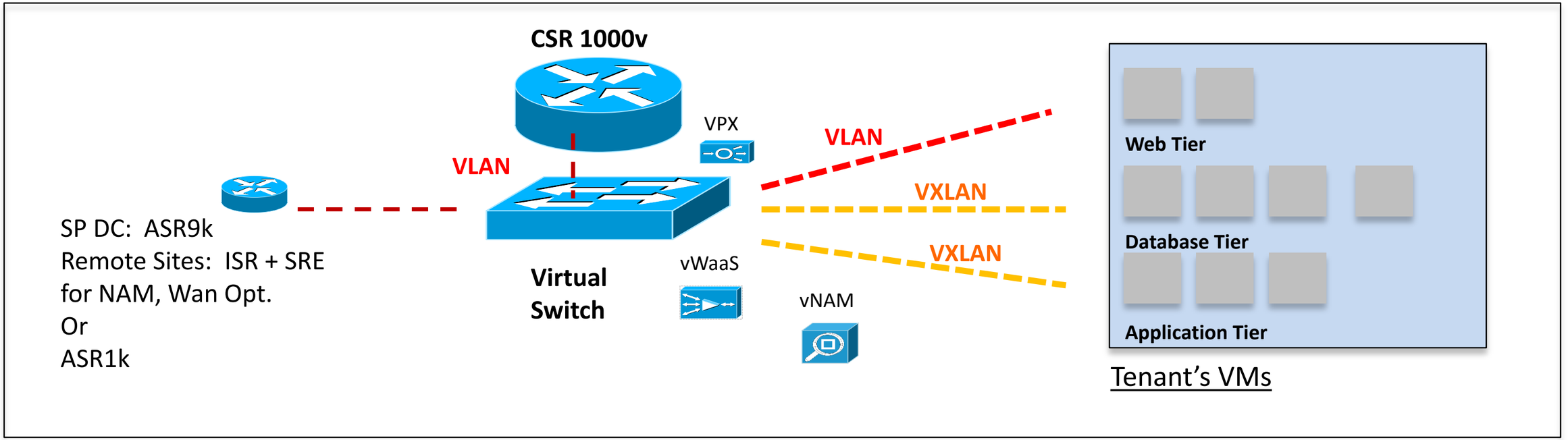
Components:

- ASA 5500 physical FW (and RA VPN Termination)
- Netscaler MPX or SDX (SLB)
- Bare metal UCS servers
- + Virtual appliances
- NAM 2300
- VXLAN for Scale



© 2013 Cisco and/or its affiliates. All rights reserved.

Application of VXLAN Virtual Overlays

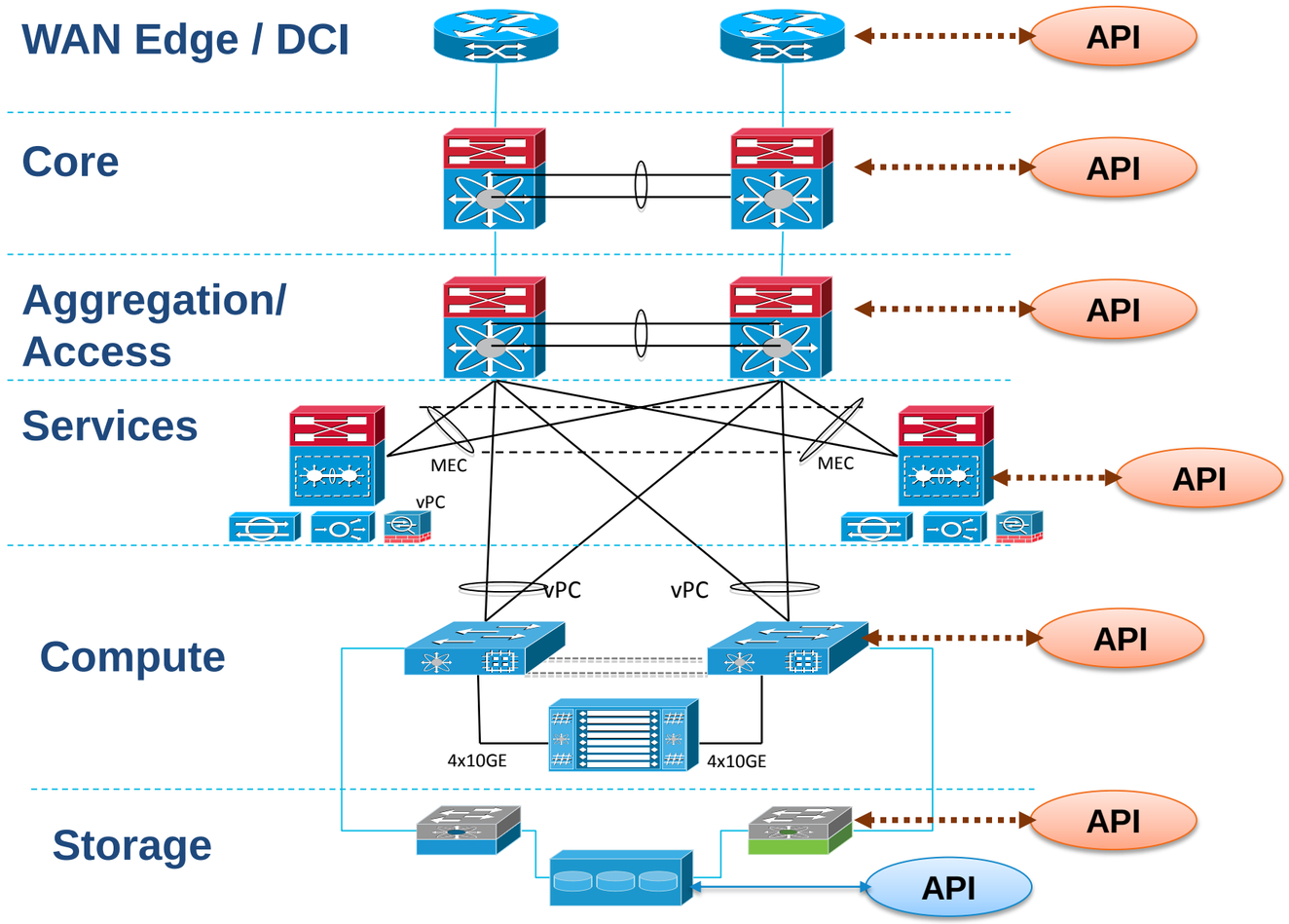


- VXLANs applied to extend segmentation scale within tenant containers
- Mapping of VXLAN to VLAN occurs on Nexus 1000v port profile

© 2013 Cisco and/or its affiliates. All rights reserved.

Cloud Orchestration & Assurance

Extensible Open Management



- VMDC offers an open management framework through a well documented set of component APIs
- The VMDC open management solution avoids vendor lock-in
- An open framework expedites VMDC integration into existing management solutions
- Cisco offers domain element management and network specific offerings such as:
  - Cisco Network Services Manager (NSM)
  - Cisco Data Centre Network Manager
  - Cisco UCS Manager
- **Storage solutions vary by vendor**

© 2013 Cisco and/or its affiliates. All rights reserved.

Cisco Cloud Management Solutions

- High Scale & Multi-tenant Apps
- Significant Complexity
- Established Market Position
- Complex Cloud Target/SP

- Automation of IT processes
- Integration of apps to the business process
- Private Cloud/ Large Enterprises

- Others like OpenStack
- Leveraging partner company assets

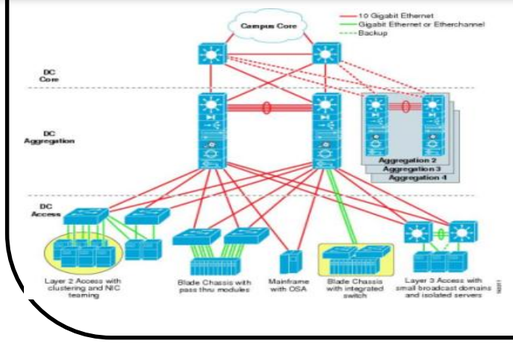
Today → Future

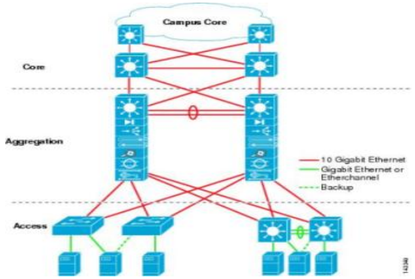
bmc

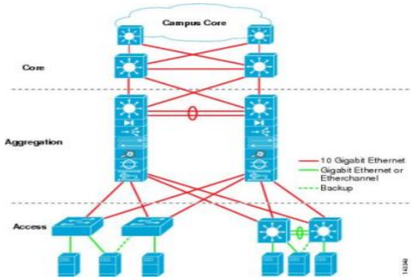
cisco

OpenStack/
Others

CISCO COMMON TECHNOLOGY
Network Services Mgr
CCN, etc.



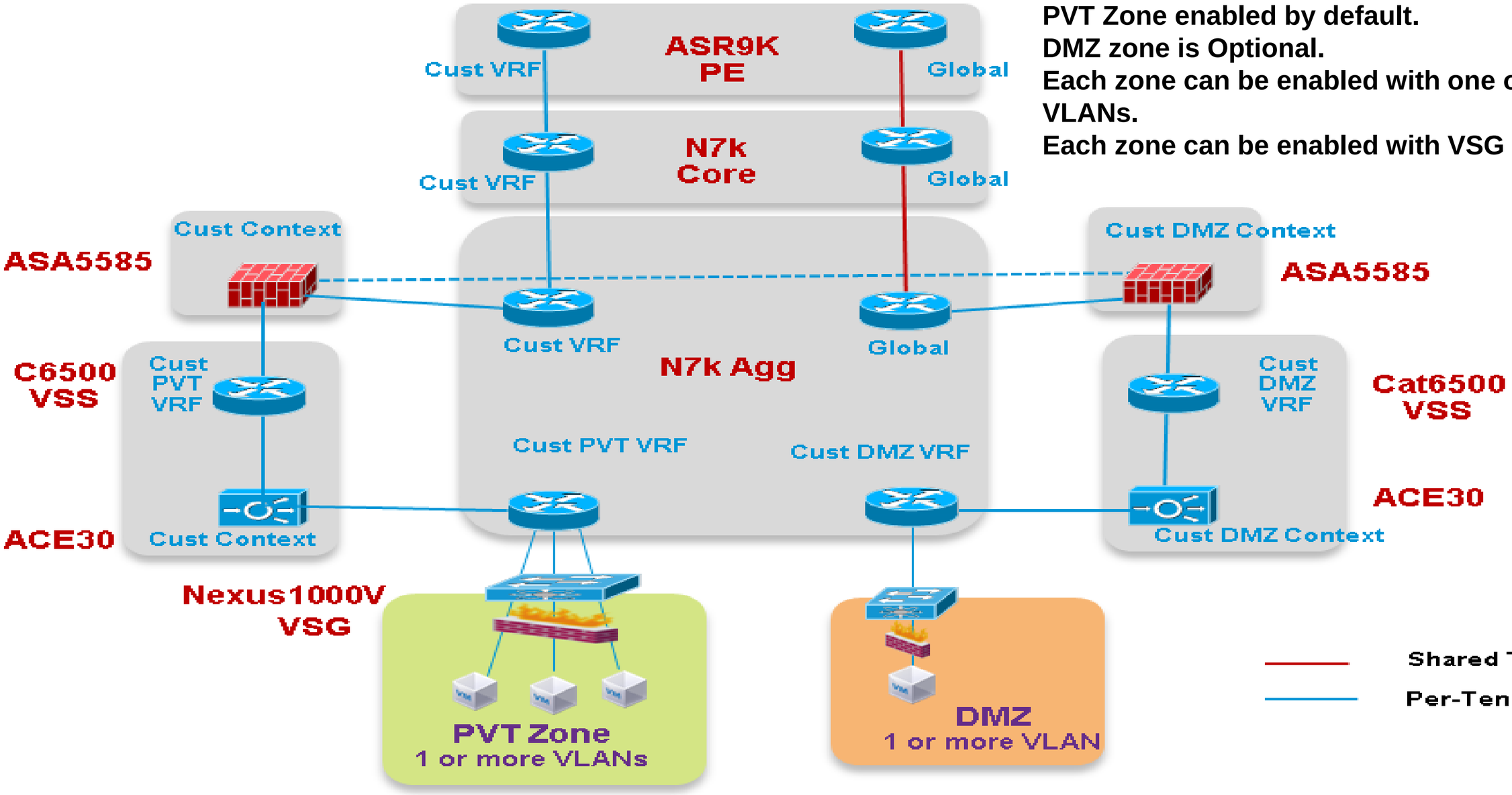




© 2013 Cisco and/or its affiliates. All rights reserved.

VMDC 2.2 Expanded Gold Container

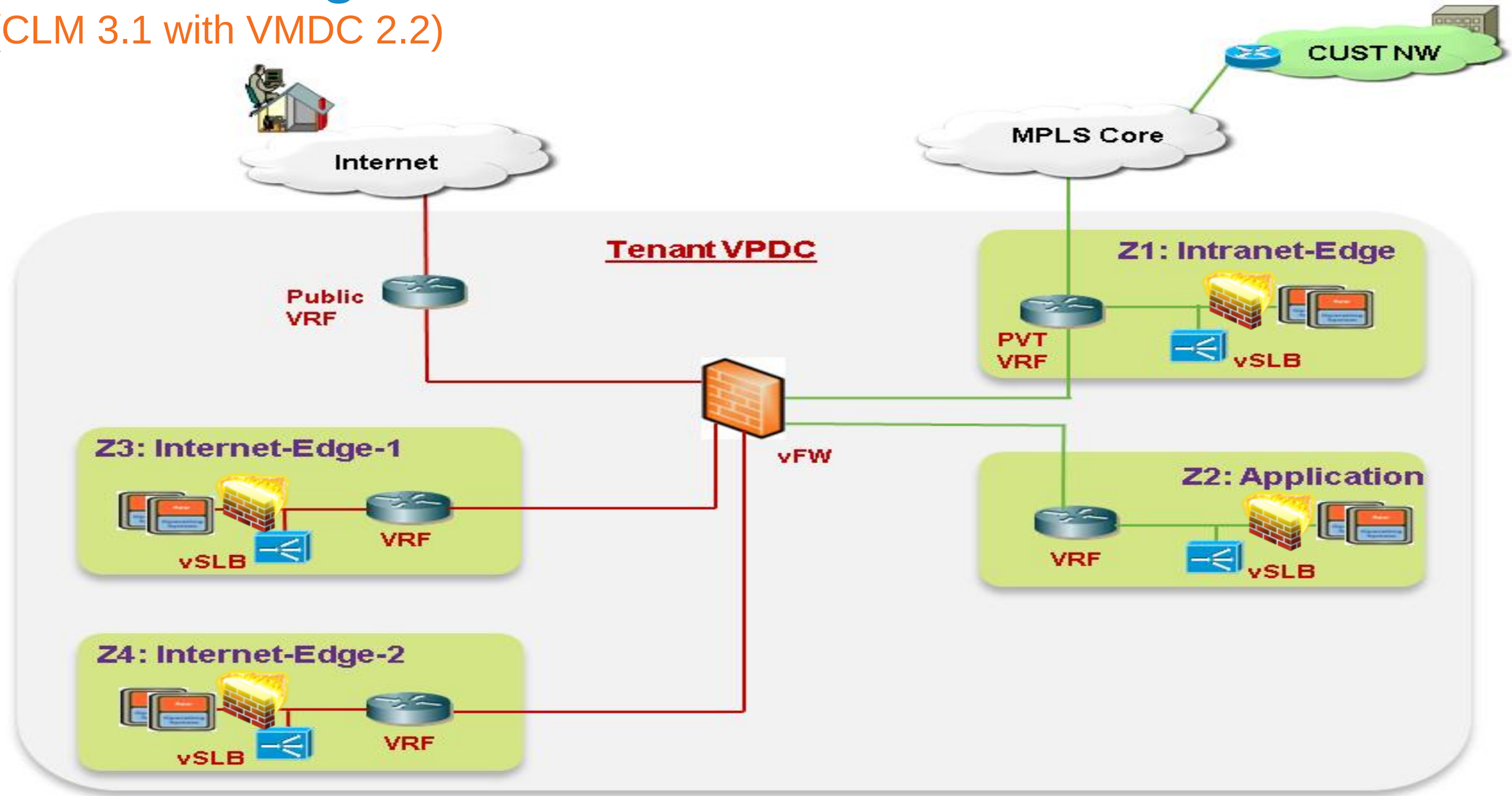
Flexibility built into Bluprints
PVT Zone enabled by default.
DMZ zone is Optional.
Each zone can be enabled with one or more VLANs.
Each zone can be enabled with VSG if needed.



© 2013 Cisco and/or its affiliates. All rights reserved.

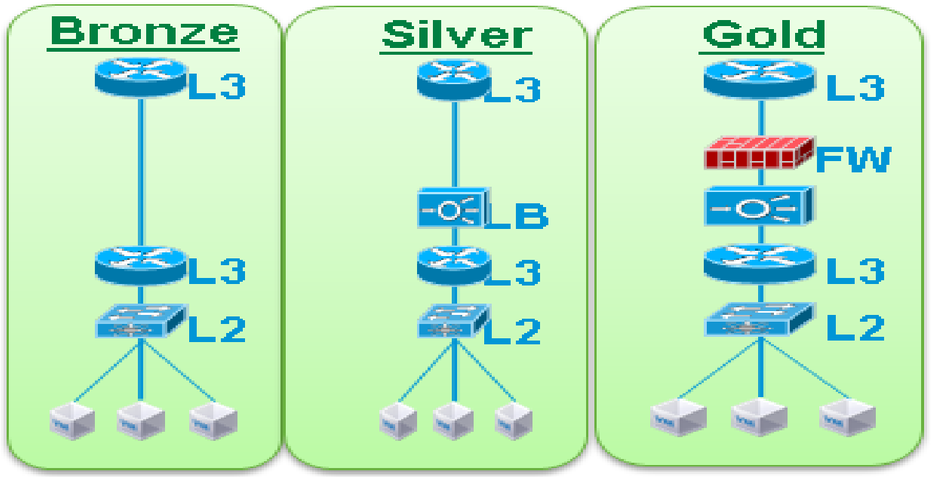
Transitioning to More Flexible Container Abstractions

(CLM 3.1 with VMDC 2.2)



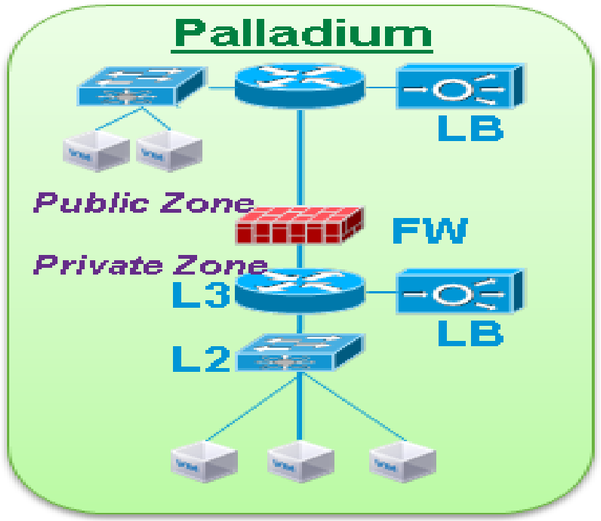
© 2013 Cisco and/or its affiliates. All rights reserved.

BMC CLM Supported Network Containers



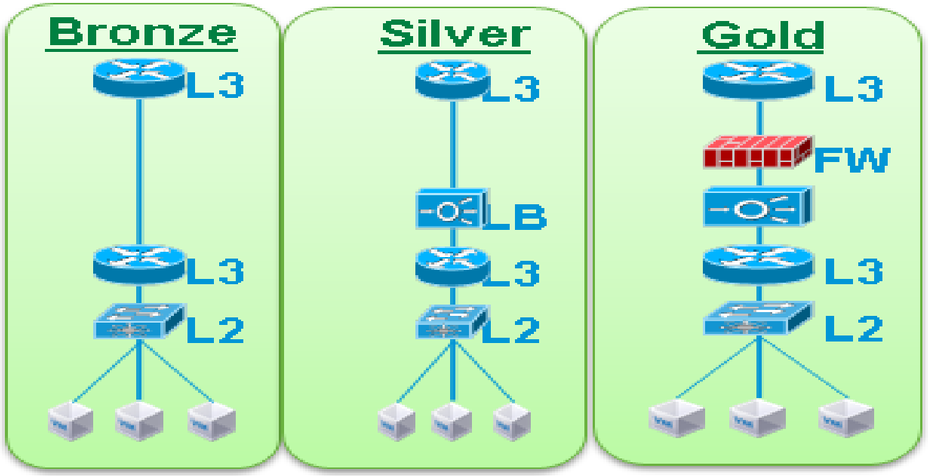
VMDC 2.0 + CLM 2.1

CRS, N7k, C6k, FWSM, ACE20



VMDC 2.1 + CLM 2.1

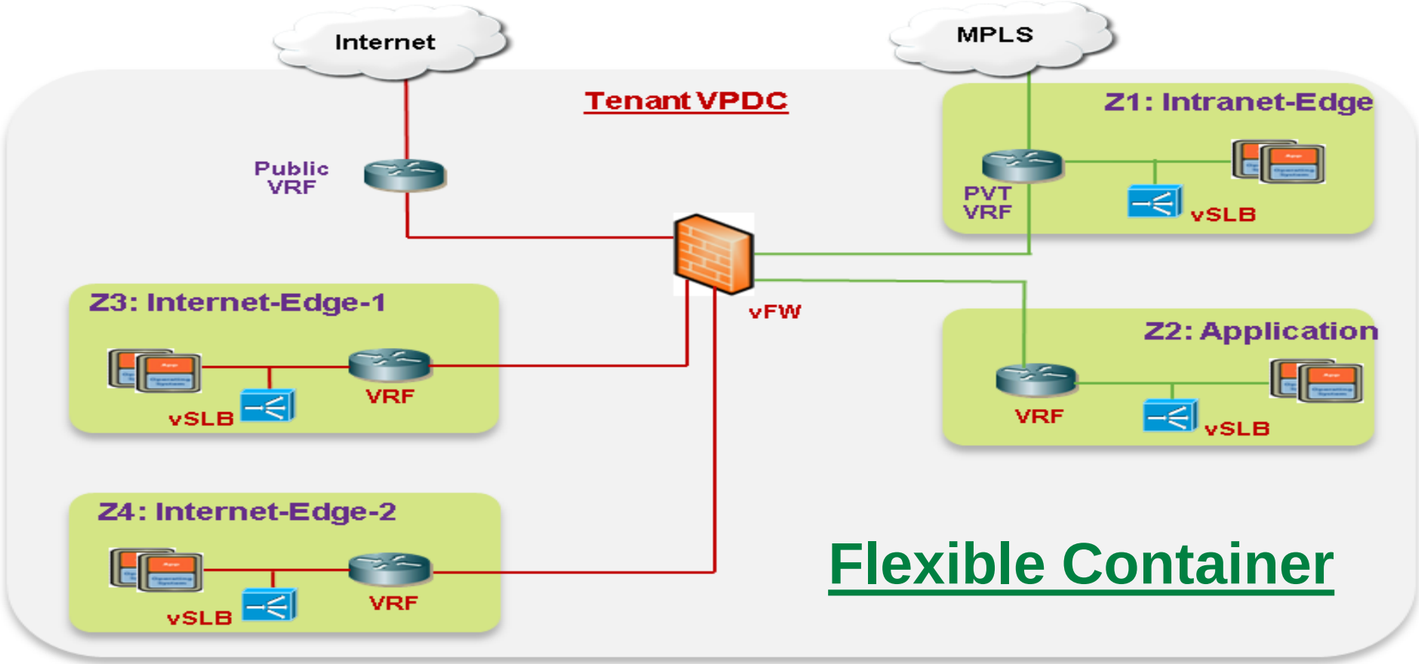
C6k, N7k, FWSM, ACE20



VMDC 2.2 + CLM 3.0

(Minus VSG & Expanded Gold container)

A9k, N7k, C6k, ASA, ACE30



VMDC 2.2 + CLM 3.0

4-Zone Flexible Container (No VSG)

A9k, N7k, C6k, ASA, ACE30

© 2013 Cisco and/or its affiliates. All rights reserved.

Automation Lessons Learned

- Automating with sufficient flexibility (multiple service models) is a key challenge
- Standardised POD design, baseline set of services in POD – homogenous, repeatable
- Simplify designs for topologies, service tiers etc – simpler workflows

  Simple Design: 1-2 Zones, 2-5 VLANs, 1 VRF, 1VFW, 1 vSLB etc.

  Complex Design: 5-6 Zones, 5-10 VLANs, 3-5 VRFs, 2-4 vFWs, 2-4 vSLB etc – complex routing between tiers/zones.
- Minimise number of Service Tiers, Network Containers/Zones, Service Offerings – service catalogue
- Simpler design leads to fewer touchpoints - efficient orchestration
- Faster customisation and deployment of Orchestration Systems
- Identify scale limits within each layer of the POD – Resource Pools, and Capacity/Resource Management
- Easier to maintain, troubleshoot, identify faults, provide service assurance.

© 2013 Cisco and/or its affiliates. All rights reserved.

Why Cloud Service Assurance?

Accelerates Cloud Adoption

- The inability to assure SLAs is a key barrier to entry to cloud services for enterprises (Source: IDC)

Differentiates Service

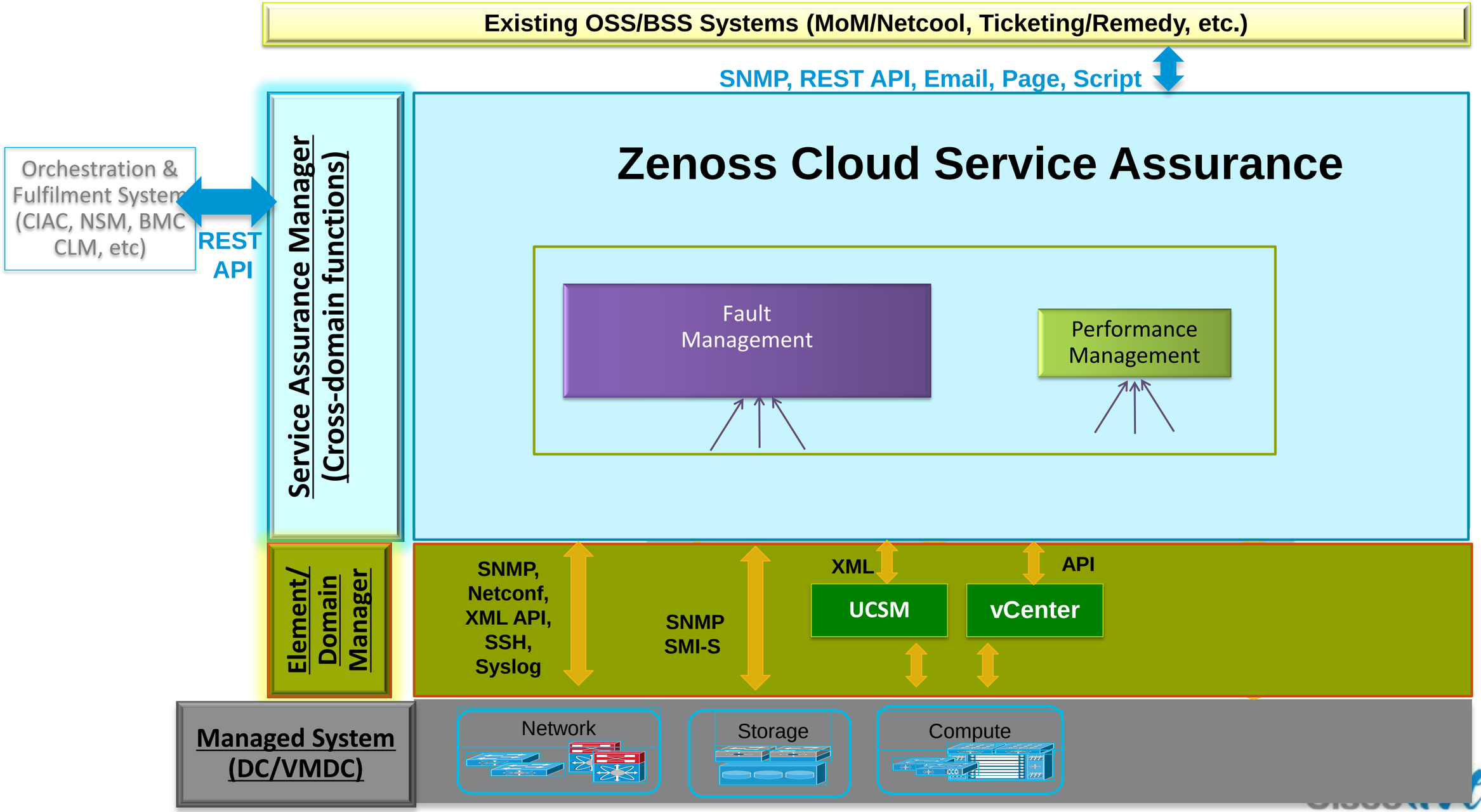
- The ability to assure SLAs is a key differentiator of (virtual) private cloud services over public cloud services; (Source: Cisco IBSG)
- Improved SLAs, Support & App Response Time Are Largest Drivers of Revenue; (Source: Schireson Associates)

Reduces cost

- Complexity & Cost of Day 2 operations significantly contributes to cost of overall cloud service
- Reducing Mean Time To Repair (MTTR) :
  - Unified monitoring: reduces swivel-chair
  - Root-cause analysis: faster answer to what's really broken
  - Service-impact analysis: prioritises faults by business impact

© 2013 Cisco and/or its affiliates. All rights reserved.

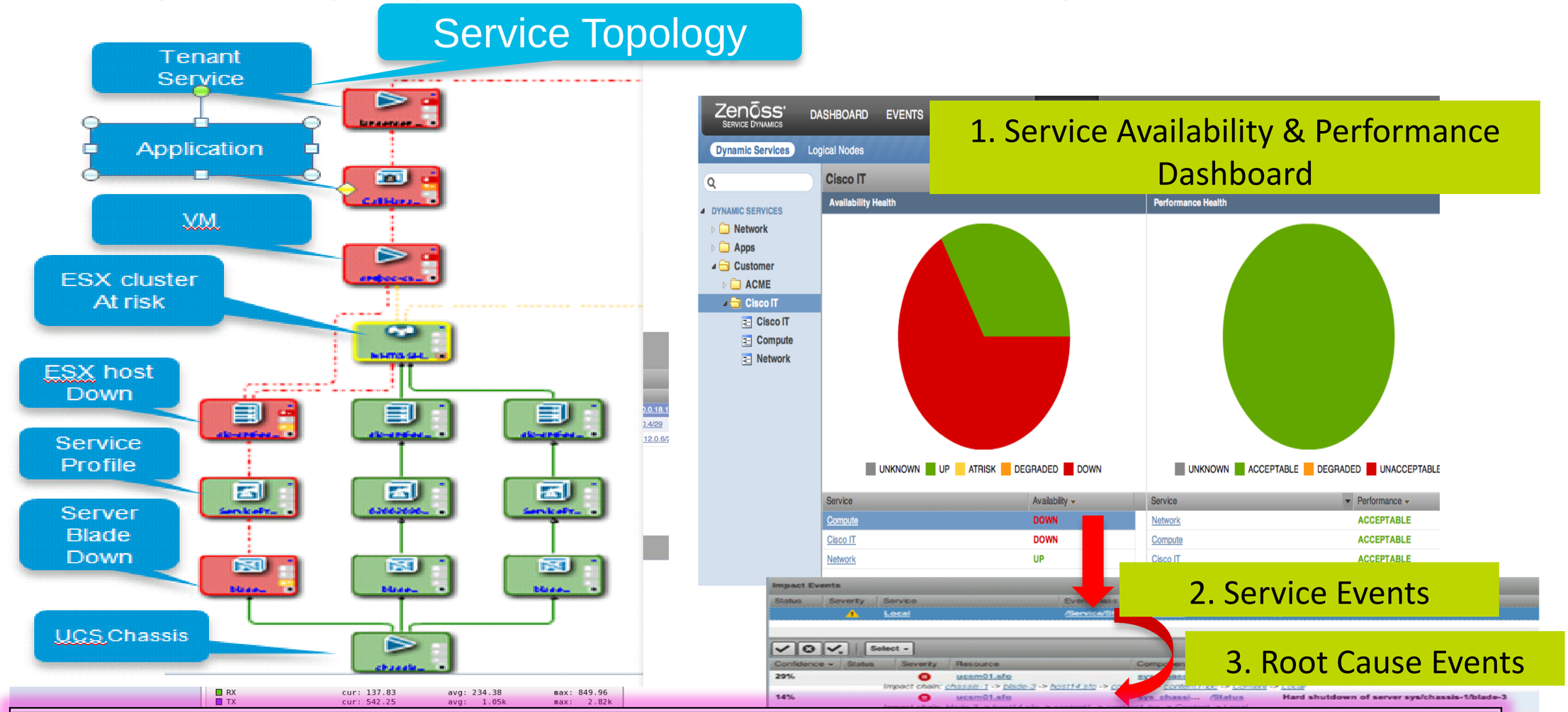
Cloud Service Assurance for VMDC - Architecture



© 2013 Cisco and/or its affiliates. All rights reserved.

VMDC Use Case Example – Summary

With service impact analysing raw device level data becomes optional



© 2013 Cisco and/or its affiliates. All rights reserved.

VMDC Resource Links

- VMDC Design Zone
  http://www.cisco.com/go/vmdc
- Questions: ask-vmdc-external@cisco.com
- VMDC 2.2 CVD
  http://www.cisco.com/en/US/partner/docs/solutions/Enterprise/Data_Center/VMDC/2.2/implementation_guide/vmdcImplementationGuide22.html
- VMDC 3.0 CVD
  http://www.cisco.com/en/US/partner/docs/solutions/Enterprise/Data_Center/VMDC/3.0/IG/VMDC_3.0_IG.html
- Cisco Cloud Megatest (based on VMDC)
  http://www.cisco.com/en/US/solutions/ns341/eantc_cloud.html
- Data Centre Interconnect Design Zone
  http://www.cisco.com/en/US/partner/netsol/ns749/networking_solutions_sub_program_home.html
- VMDC Orchestration with BMC CLM
  http://www.cisco.com/en/US/partner/solutions/ns340/ns414/ns742/cloud_orchestration_bmc_clm.html#~entitled
- VMDC Assurance with Zenoss CSA
  http://www.cisco.com/en/US/partner/solutions/ns340/ns414/ns742/dz_cloudservice.html
- Cloud Enablement Services Website
  http://www.cisco.com/en/US/products/ps11104/serv_home.html

© 2013 Cisco and/or its affiliates. All rights reserved.

Q & A

Complete Your Online Session Evaluation

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

Complete your Overall Event Survey and 5 Session Evaluations.

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

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

Cisco live! 365

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

Cisco live!

© 2013 Cisco and/or its affiliates. All rights reserved.

© 2013 Cisco and/or its affiliates. All rights reserved.