



Cisco Expo KSA

Cisco Data Center Interconnect

Faadil Adam
Systems Engineer – Service Provider

5th December 2011

Agenda

- Data Center Interconnect (DCI):
 - Business Drivers
 - DCI Solution Components
 - SAN Extension
 - LAN Extension
 - Path Optimization
- LAN Extension Technologies
 - Overview
 - Ethernet Based
 - Overlay Transport Virtualization
 - A-VPLS
 - VPLS
- Conclusion and Q&A

Data Centers Under Increasing Pressure

Collaborative Services



Visual Networking



User Expectations



Economy

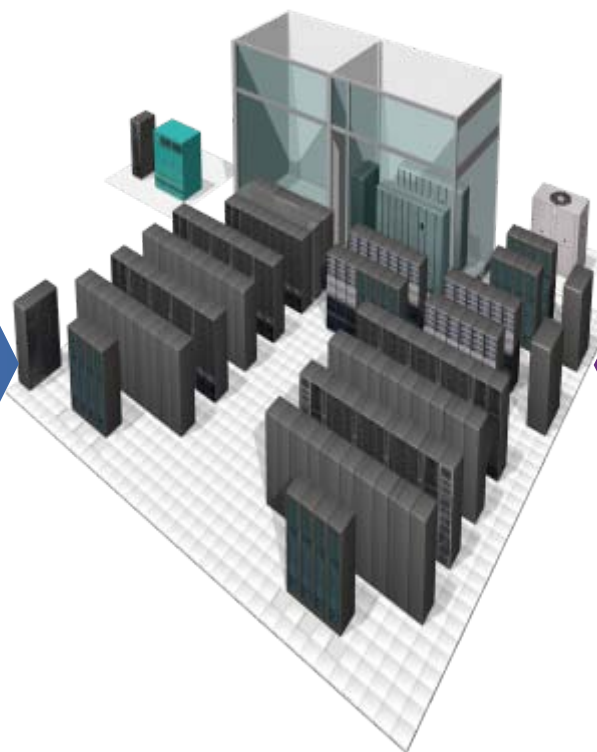


Regulation



New
Business
Pressures

Satisfying Increasing Demand
While Decreasing Cost



Operational
Limitations

Power and Cooling



Asset Utilization



Provisioning



Security Threats



Bus. Continuity



Data Center Interconnect

Business Drivers

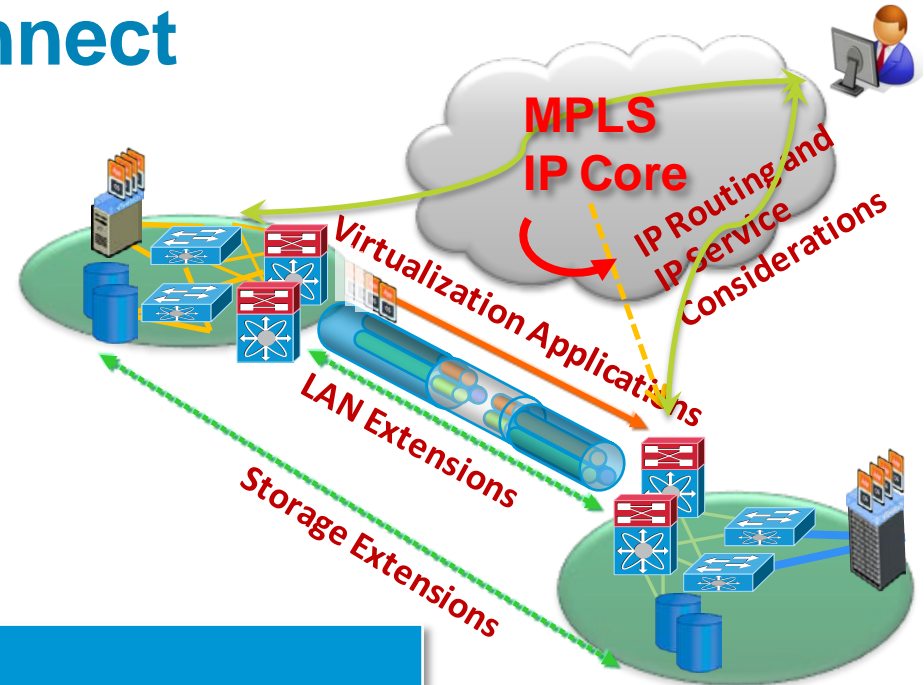
- Data Centers are extending beyond traditional boundaries
- Virtualization applications are driving DCI across PODs (aggregation blocks) and Data Centers



Drivers	Business Solution	Constraints	IT Technology
Business Continuity	<ul style="list-style-type: none"> ✓ Disaster Recovery ✓ HA Framework 	<ul style="list-style-type: none"> ✓ Stateless ✓ Network Service Sync ✓ Process Sync 	<ul style="list-style-type: none"> ✓ GSLB ✓ Geo-clusters ✓ HA Cluster
Operation Cost Containment	<ul style="list-style-type: none"> ✓ Data Center Maintenance / Migration / Consolidation 	<ul style="list-style-type: none"> ✓ VLAN Extension 	<ul style="list-style-type: none"> ✓ Distributed Virtual Data Center
Business Resource Optimization	<ul style="list-style-type: none"> ✓ Disaster Avoidance ✓ Workload Mobility 	<ul style="list-style-type: none"> ✓ Stateful ✓ Bandwidth ✓ Latency 	<ul style="list-style-type: none"> ✓ VM Mobility
Cloud Services	<ul style="list-style-type: none"> ✓ Inter-Cloud Networking ✓ XaaS 	<ul style="list-style-type: none"> ✓ Flexibility 	<ul style="list-style-type: none"> ✓ VM Mobility ✓ Automation

Data Center Interconnect

Solution Components



DCI Function	Purpose
Storage Extension	Providing applications access to storage locally, as well as remotely with desirable storage attributes
LAN Extension	Extend same VLAN across Data Centers, to virtualize servers and applications
Path Optimization	Routing users to the data center where the application resides while keeping symmetrical routing in consideration for IP services (e.g. Firewall)
Inter-DC Routing	Provide routed connectivity between data centers (used for L3 segmentation/virtualization, etc.)

Data Center Interconnect

SAN Extension



DC 1



ESX-A source



vmware



Edge



Director



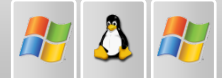
DC 2



ESX-B target



vmware



Edge



Director



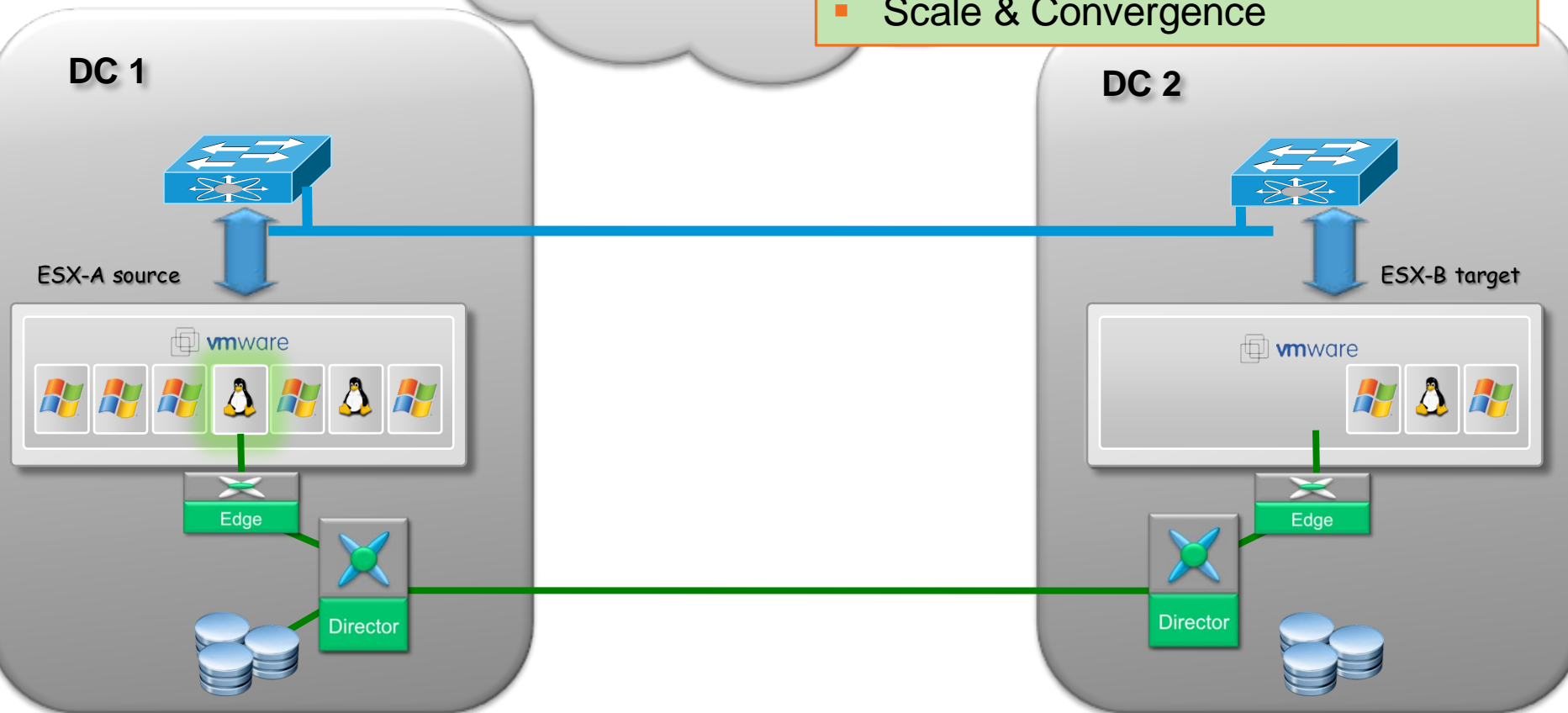
- SAN Extension typically over DWDM/ Dark Fiber
- Synchronous/ Asynchronous Replication
- Storage Elasticity

Data Center Interconnect

LAN Extension



- STP Isolation is the key element
- Multipoint
- Loop avoidance + Storm-Control
Unknown Unicast & Broadcast control
- Link sturdiness
- Scale & Convergence



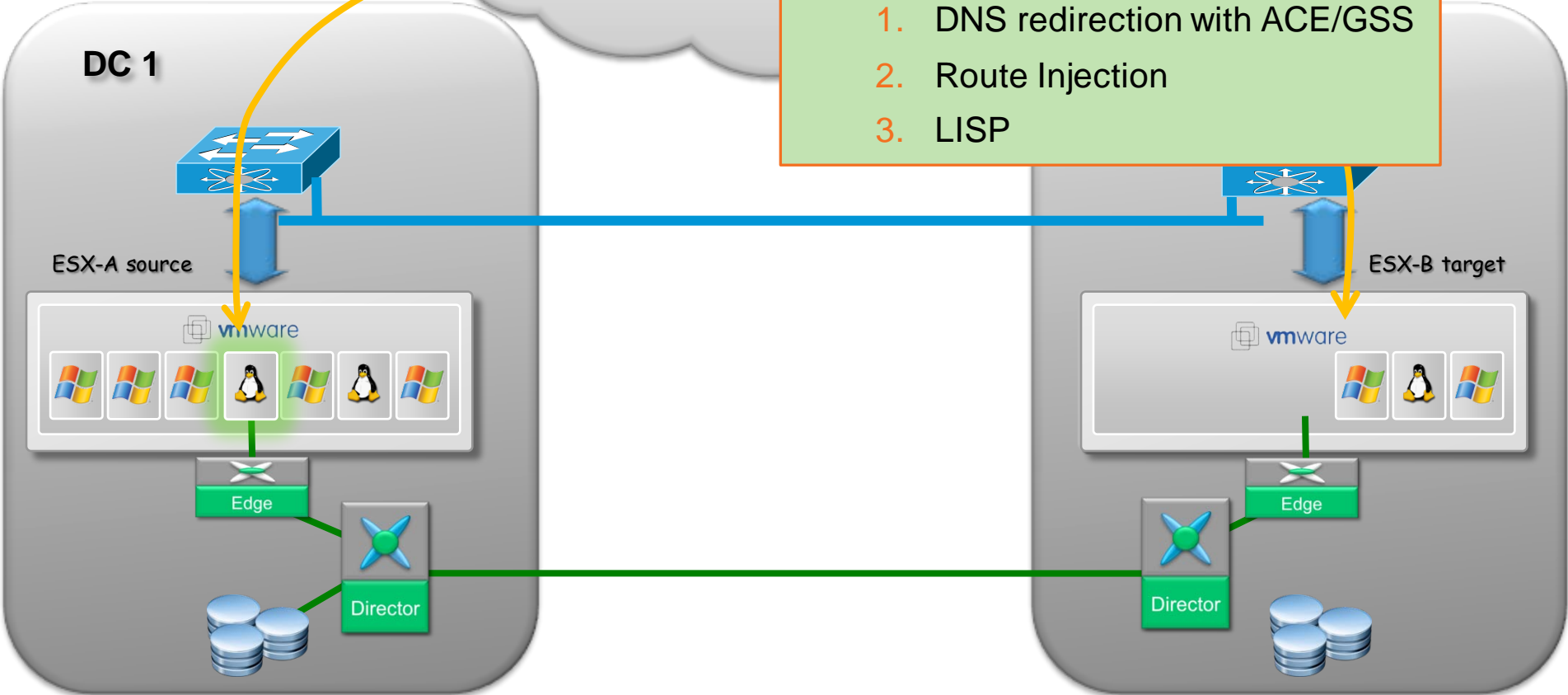
Data Center Interconnect

Path Optimization



Options

- Egress
 - ➔ Addressed by FHRP Filtering
- Ingress:
 1. DNS redirection with ACE/GSS
 2. Route Injection
 3. LISP



Agenda

- Data Center Interconnect (DCI):
 - Business Drivers
 - DCI Solution Components
 - SAN Extension
 - LAN Extension
 - Path Optimization
- LAN Extension Technologies
 - Overview
 - Ethernet Based
 - Overlay Transport Virtualization
 - A-VPLS
 - VPLS
- Conclusion and Q&A

LAN Extension Overview



Data Center Interconnect Technology Options

Ethernet

➤ *VSS & vPC, FabricPath/ TRILL*

- Applies easily for dual site interconnection
- Over dark fiber or protected DWDM
- Easy crypto using end-to-end 802.1AE

IP

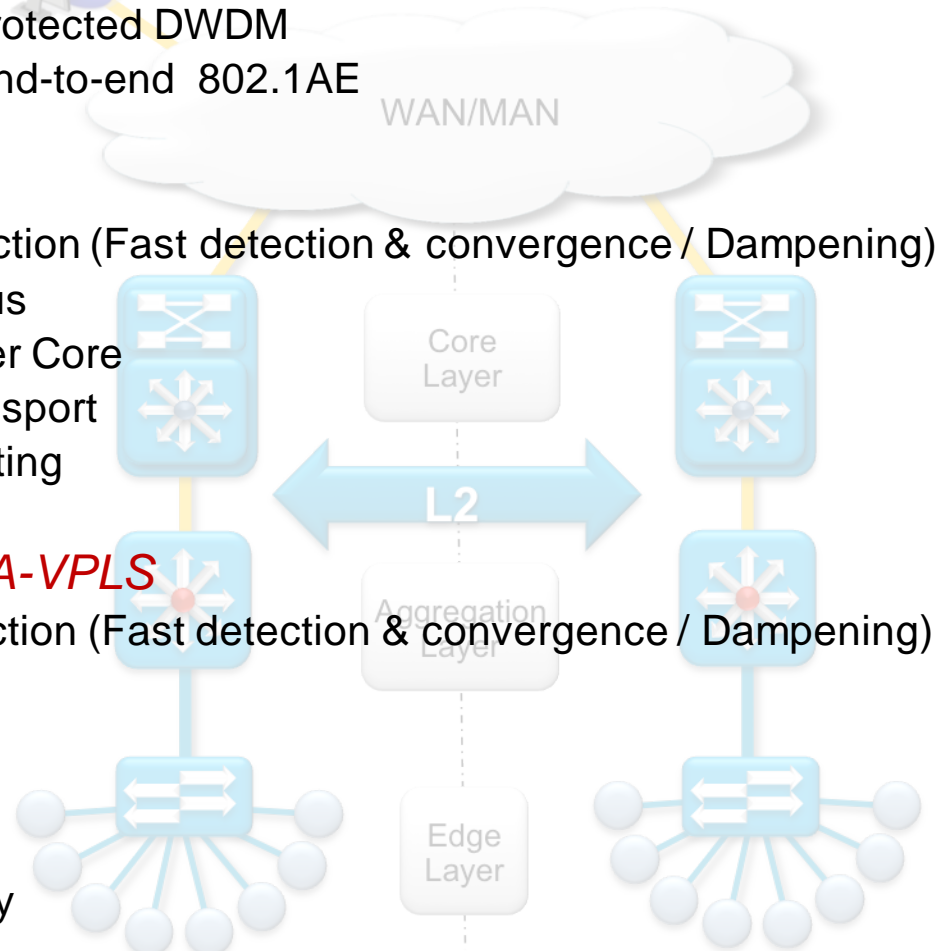
➤ *OTV*

- L2oL3 for link protection (Fast detection & convergence / Dampening)
- Enterprise / DC focus
- Easy integration over Core
- Works over any transport
- Innovative MAC routing

➤ *EoMPLS & VPLS & A-VPLS*

- L2oL3 for link protection (Fast detection & convergence / Dampening)
- PE style
- Large scale
- Multi-tenants
- Works over GRE
- Most deployed today

MPLS

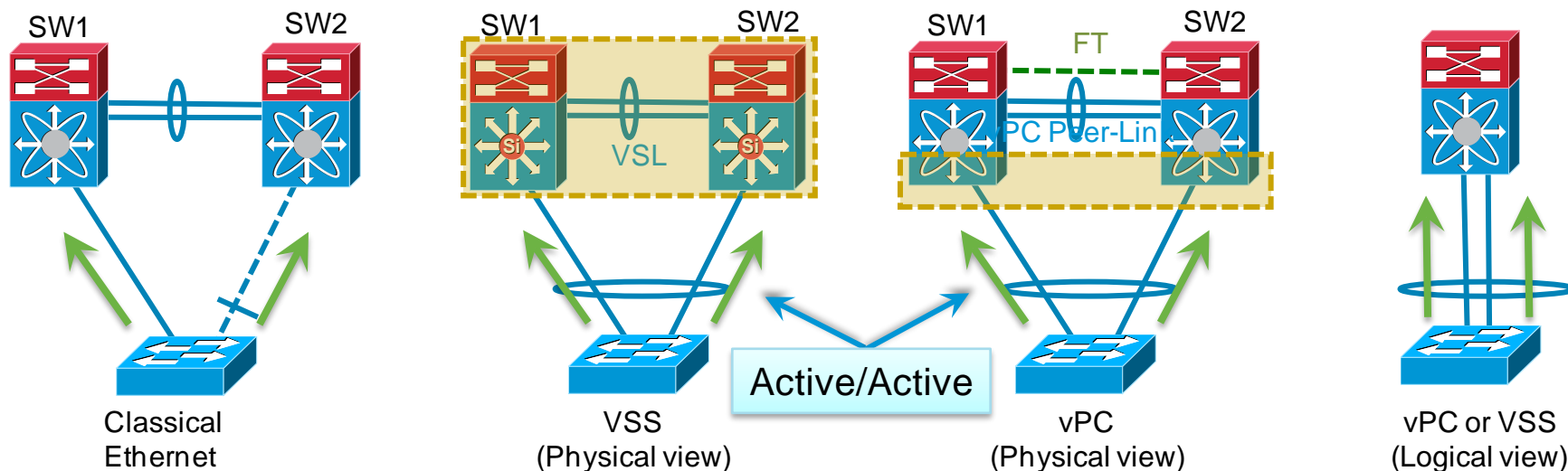


Ethernet VSS & vPC



Multi-Chassis EtherChannel (MEC)

Using Multi-Chassis Link Aggregation Control Protocol (mLACP)



Virtual Switching System on Cat6k (VSS)

Virtual Port Channel on Nexus (vPC)

Both VSS-MEC and vPC are a Port-channeling concept extending link aggregation to two separate physical switches

Allows the creation of resilient L2 topologies based on Link Aggregation.

Eliminates the dependence on STP in the L2 access-distribution Layer

Scale Available Layer 2 Bandwidth

Simplify Network Design

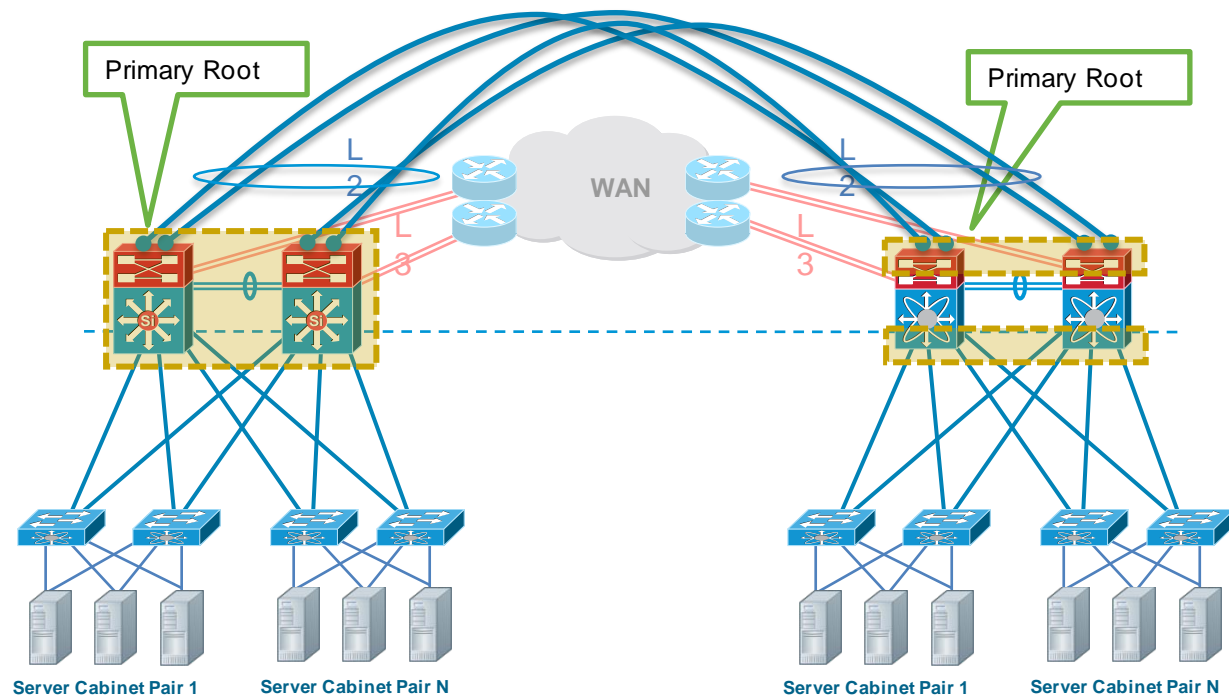
Dual Sites Interconnection

Leveraging MECs Between Sites

At DCI point:

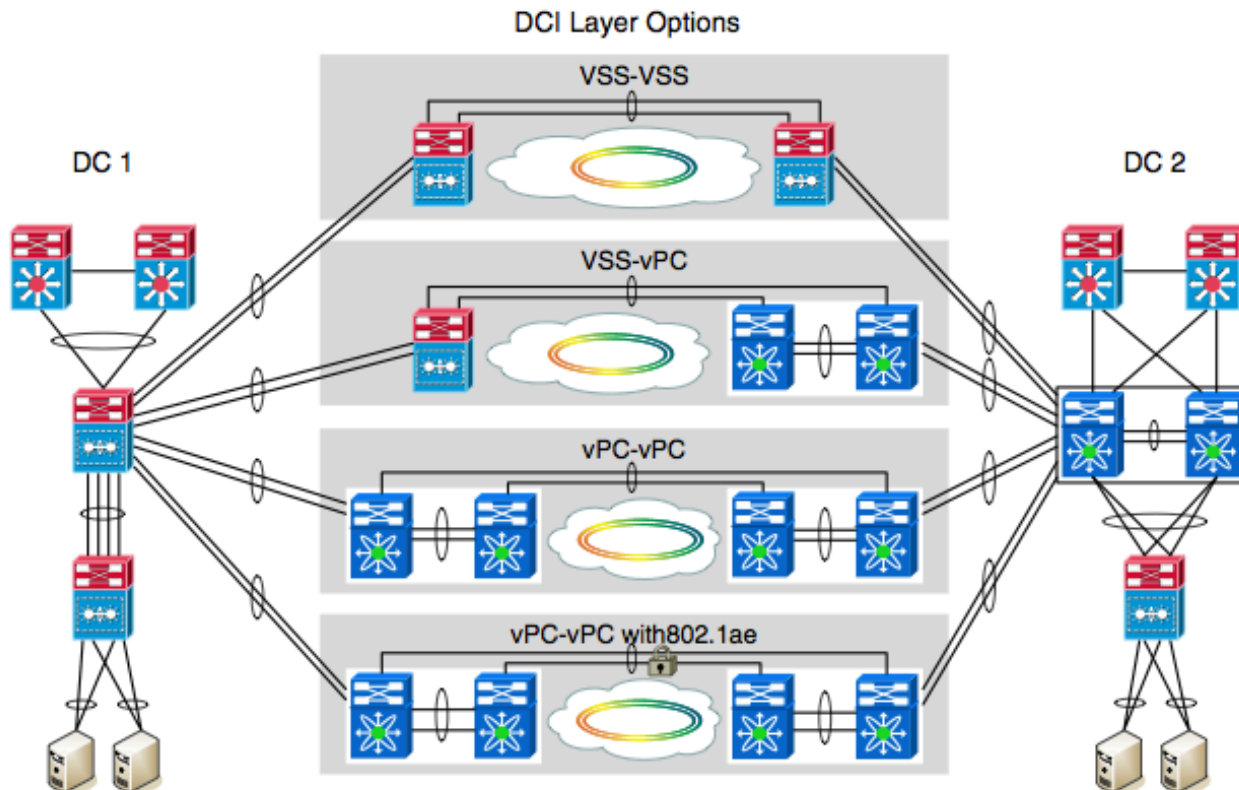
- STP Isolation (BPDU Filtering)
- Broadcast Storm Control
- FHRP Isolation

- Link utilization with MEC
- New Links for DCI
 - DCI port-channel
 - 2 with VSS
 - 4 with vPC (pre 4.2(6) release)
- Requires protected DWDM or Direct fibers



Dual Sites Use Case Summary

Cisco Validated Design on CCO



http://www.cisco.com/en/US/solutions/collateral/ns340/ns517/ns224/ns949/ns304/ns975/data_center_interconnect_design_guide.pdf

Overlay Transport Virtualization (OTV)



Overlay Transport Virtualization

Technology Pillars



OTV is a “MAC in IP” technique to extend Layer 2 domains
OVER ANY TRANSPORT



Dynamic Encapsulation

No Pseudo-Wire State Maintenance

Optimal Multicast Replication

Multipoint Connectivity

Point-to-Cloud Model



Nexus 7000

First platform to support OTV

Protocol Learning

Preserve Failure Boundary

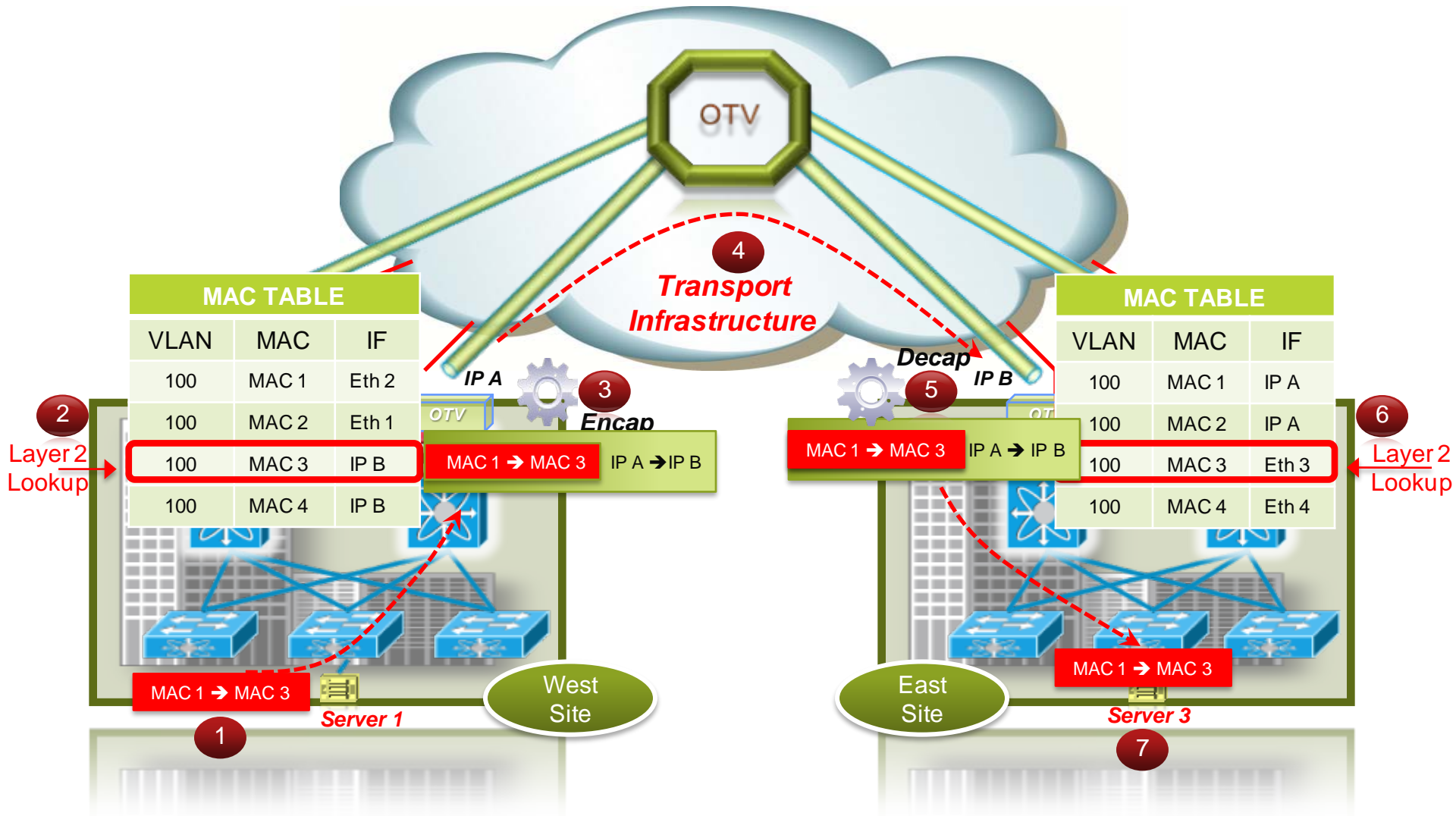
Built-in Loop Prevention

Automated Multi-homing

Site Independence

OTV Data Plane

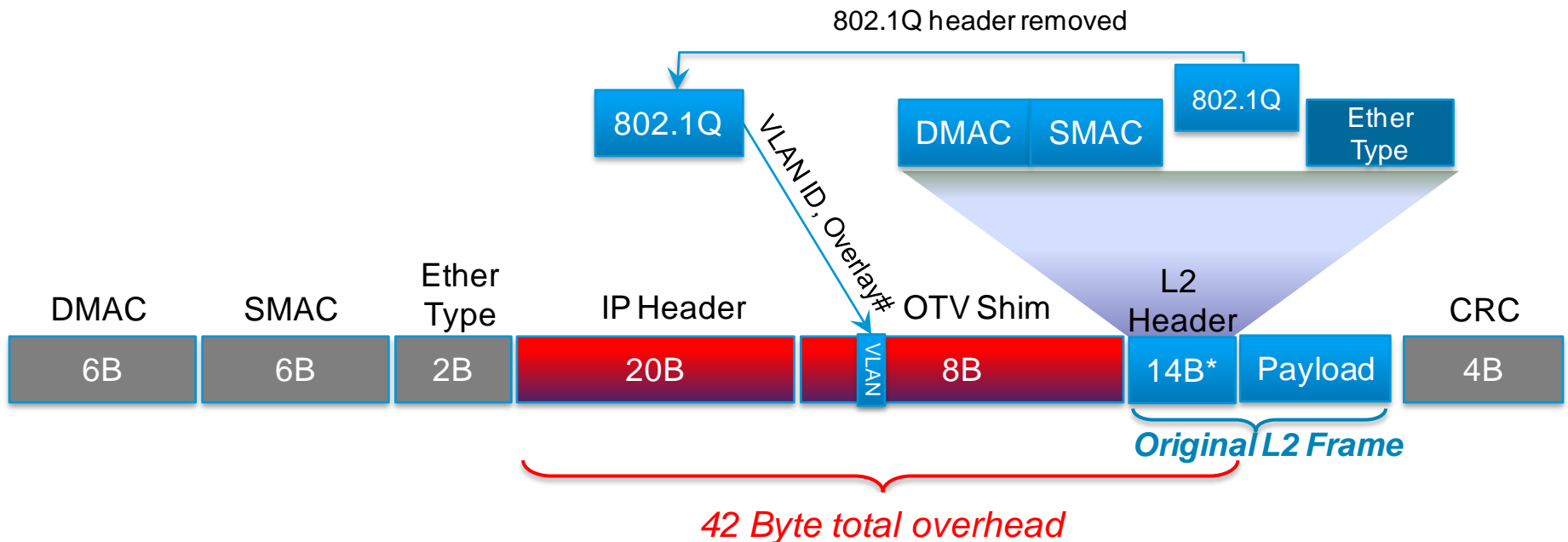
Inter-Sites Packet Flow



OTV Data Plane

Encapsulation

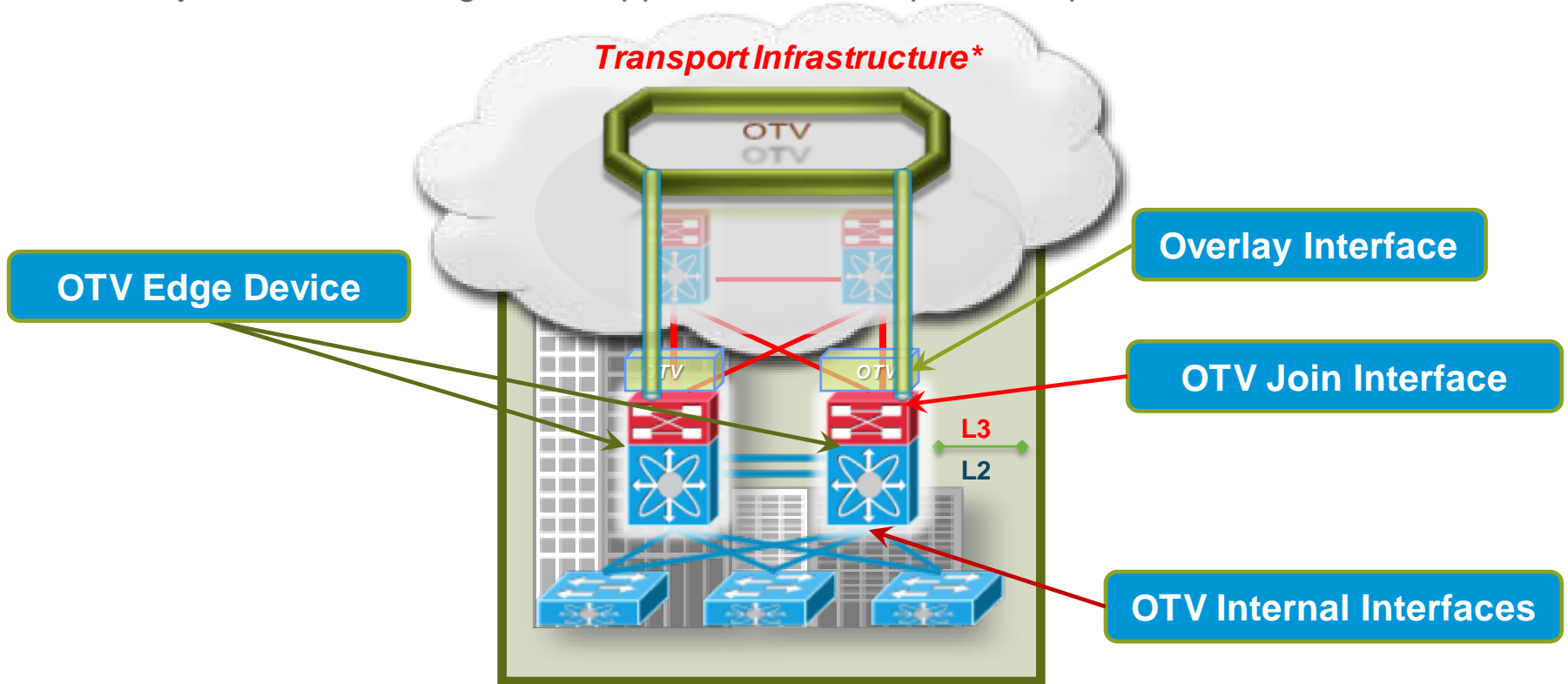
- OTV encapsulation adds 42 Bytes to the packet IP MTU size
 - Outer IP Header and OTV Shim Header in addition to original L2 Header
- The outer OTV shim header contains information about the overlay (VLAN, overlay number)
- The 802.1Q header is removed from the original frame and the VLAN field copied over into the OTV shim header



*4B of original 802.1Q header removed

OTV Terminology

- Edge Device: Responsible for performing all the OTV functionality
- Internal Interface: Regular L2 interface with typical L2 functions (local switching, STP, learning, flooding)
- Join Interface: Join OTV overlay, Discover & Adjacencies with other OTV Edge devices
- Overlay Interface: Configuration applied, OTV encapsulation performed.



OTV Control Plane

Neighbor Discovery (Multicast/ Unicast)



Mechanism

- Edge Devices (EDs) join an multicast group in the transport, as they were hosts (ASM)
- OTV hellos and updates are encapsulated in the multicast group



End Result

- Adjacencies are maintained over the multicast group
- A single update reaches all neighbors

OTV Configuration

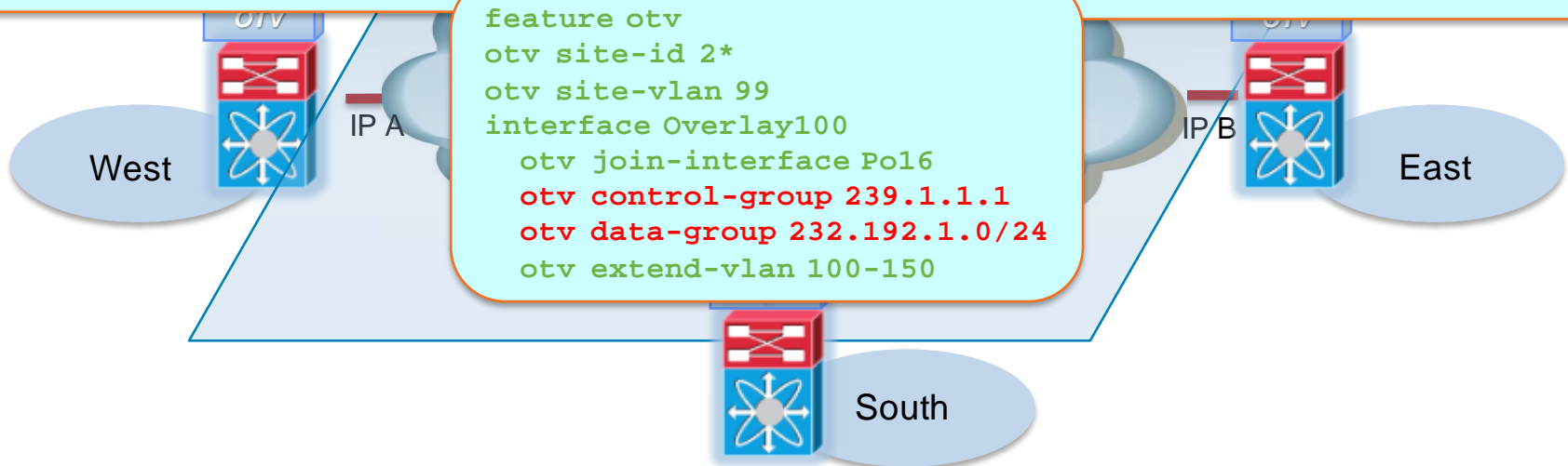
OTV over a Multicast Transport

- Minimal configuration required to get OTV up and running

```
feature otv
otv site-id 1*
otv site-vlan 99
interface Overlay100
  otv join-interface e1/1
  otv control-group 239.1.1.1
  otv data-group 232.192.1.0/24
  otv extend-vlan 100-150
```

```
feature otv
otv site-id 3*
otv site-vlan 99
interface Overlay100
  otv join-interface e1/1.10
  otv control-group 239.1.1.1
  otv data-group 232.192.1.0/24
  otv extend-vlan 100-150
```

```
feature otv
otv site-id 2*
otv site-vlan 99
interface Overlay100
  otv join-interface Po16
  otv control-group 239.1.1.1
  otv data-group 232.192.1.0/24
  otv extend-vlan 100-150
```



*Introduced from release 5.2

OTV Layer 2 Fault Isolation

- STP isolation – **No configuration required**
 - No BPDUs forwarded across the overlay
 - STP remains local to each site
 - Edge device internal interfaces behave as any other switchport
- Unknown unicast isolation – **No configuration required**
 - No unknown unicast frames flooded onto the overlay
 - Assumption is that end stations are not silent
 - Option for selective unknown unicast flooding (for certain applications)
- Proxy ARP cache for remote-site hosts – **On by default**
 - On ARP request for remote host, request forwarded through OTV and initial ARP reply generated by that host
 - OTV edge device snoops ARP replies and caches data
 - Subsequent ARP replies proxied by local OTV edge device using ARP cache

OTV Scalability

Current and Future Supported Values

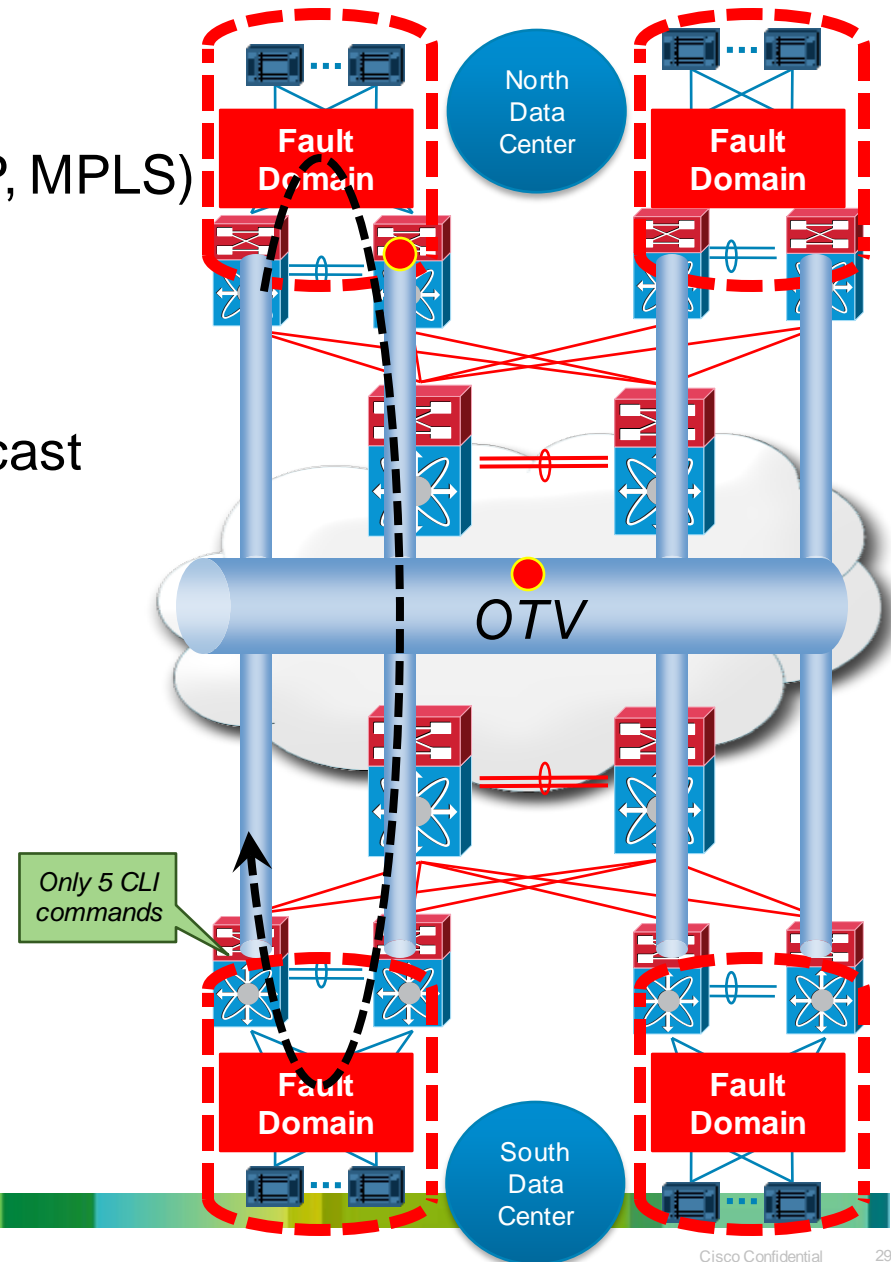
The following values have been tested and verified:

- 6 Sites
- 256 OTV extended VLANs
- 16K MAC Addresses across all the extended VLANs
- 3000 Sites' Multicast Data Groups

OTV

Summary

- Extensions over any transport (IP, MPLS)
- Failure boundary preservation
- Site independence
- Optimal BW utilization with multicast enabled transport infrastructure (no head-end replication)
- Automated Built-in Multihoming
- End-to-End loop prevention
- Scalability
 - Sites, VLANs, MACs
- Operations simplicity



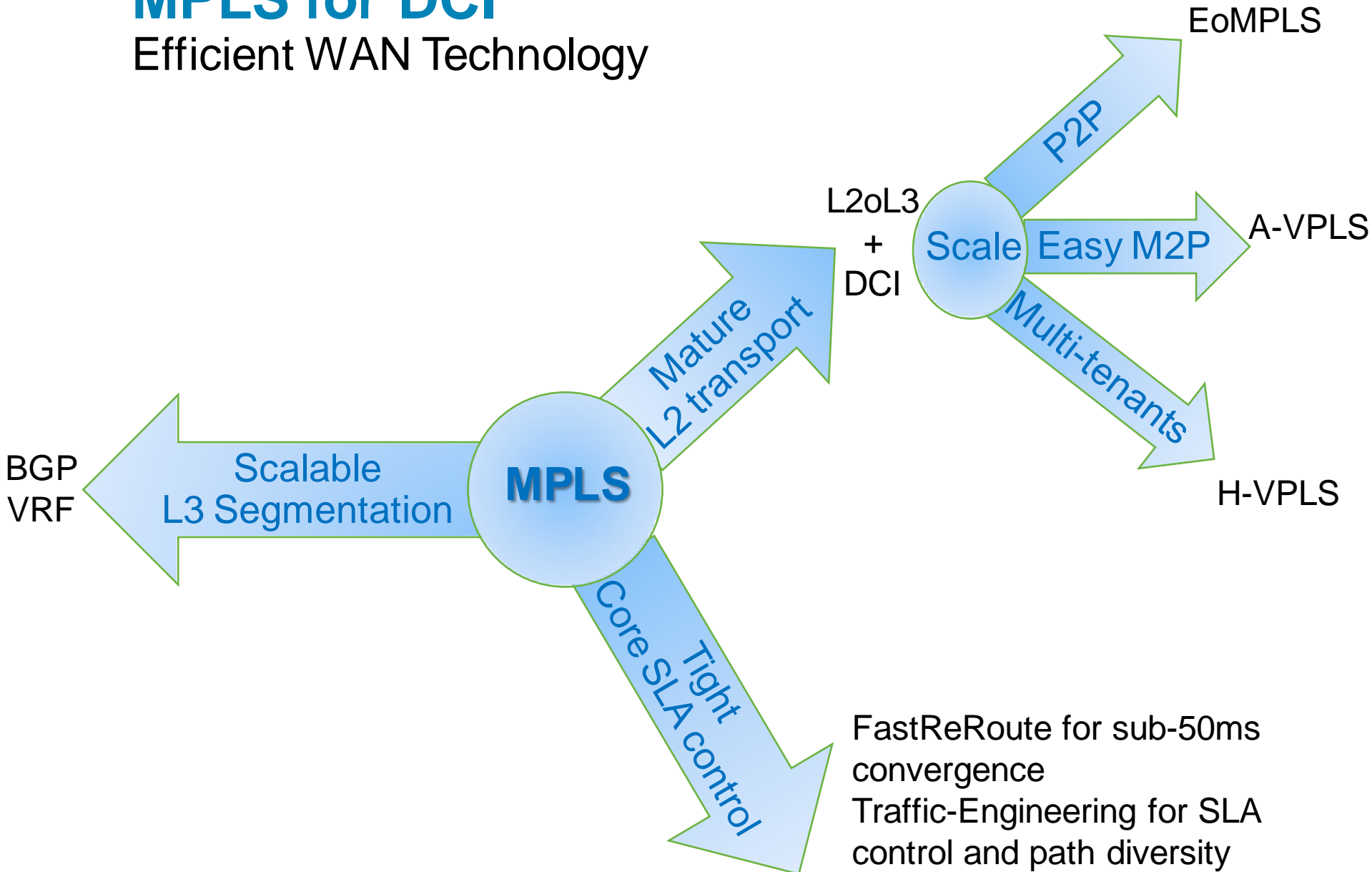
MPLS

EoMPLS, A-VPLS, VPLS



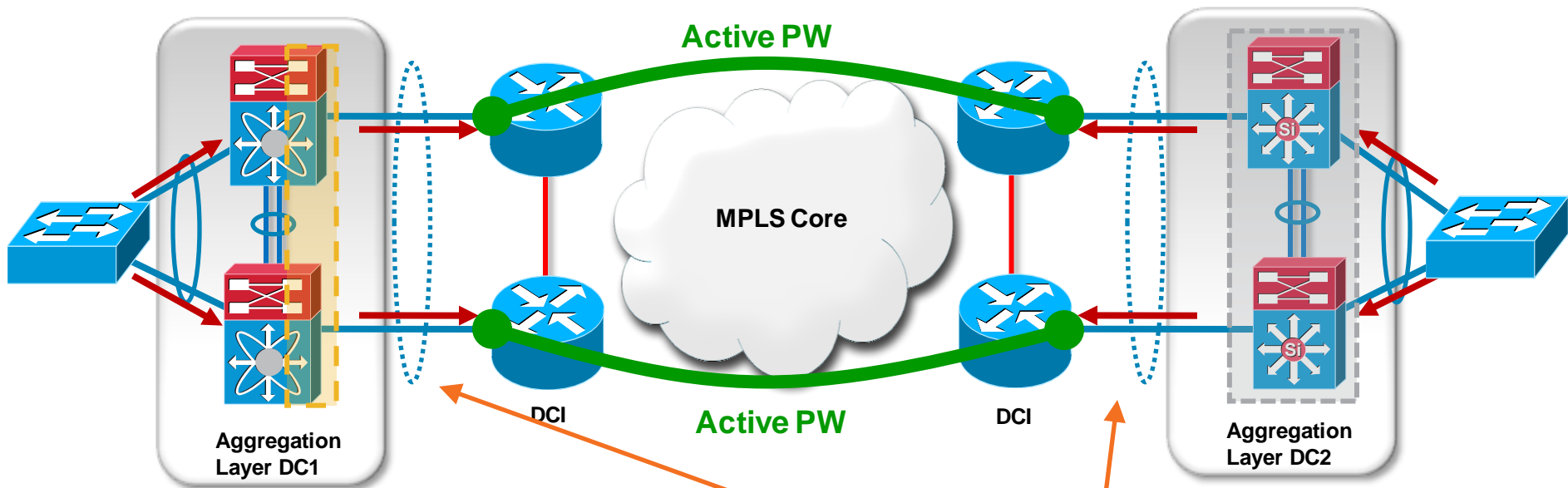
MPLS for DCI

Efficient WAN Technology



EoMPLS Usage with DCI

End-to-End Loop Avoidance Using Edge to Edge LACP



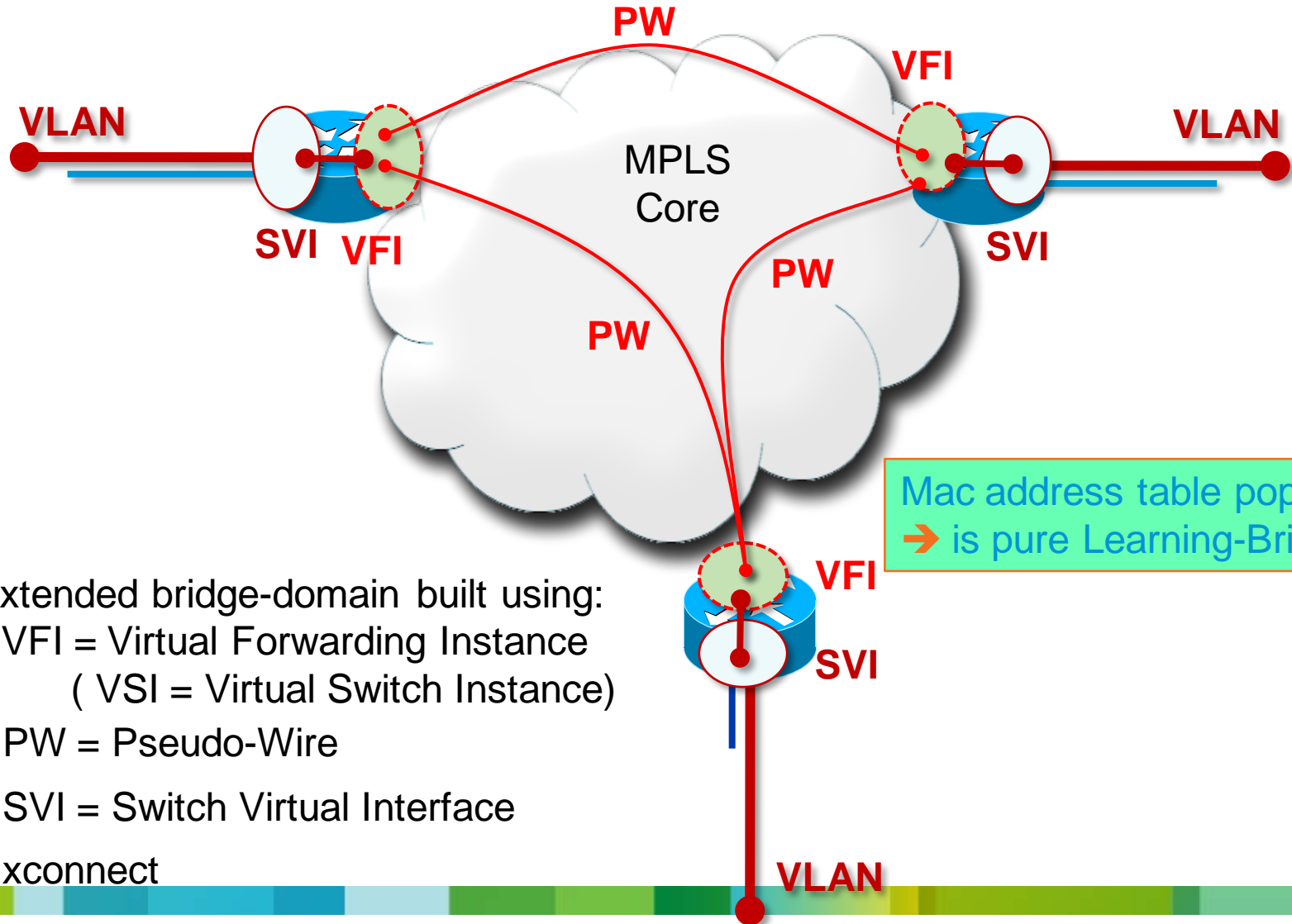
```
interface port-channel170
description L2 PortChannel to DC 2
spanning-tree port type edge trunk
spanning-tree bpdufilter enable
storm-control broadcast level 1*
storm-control multicast level x
```

*Value to be tuned, min is 0.3

- BPDU Filtering to maintain STP domains isolation
- Storm-control for data-plane protection
- Configuration applied at aggregation layer on the logical port-channel interface

Multi-Point Topologies

What Is VPLS?



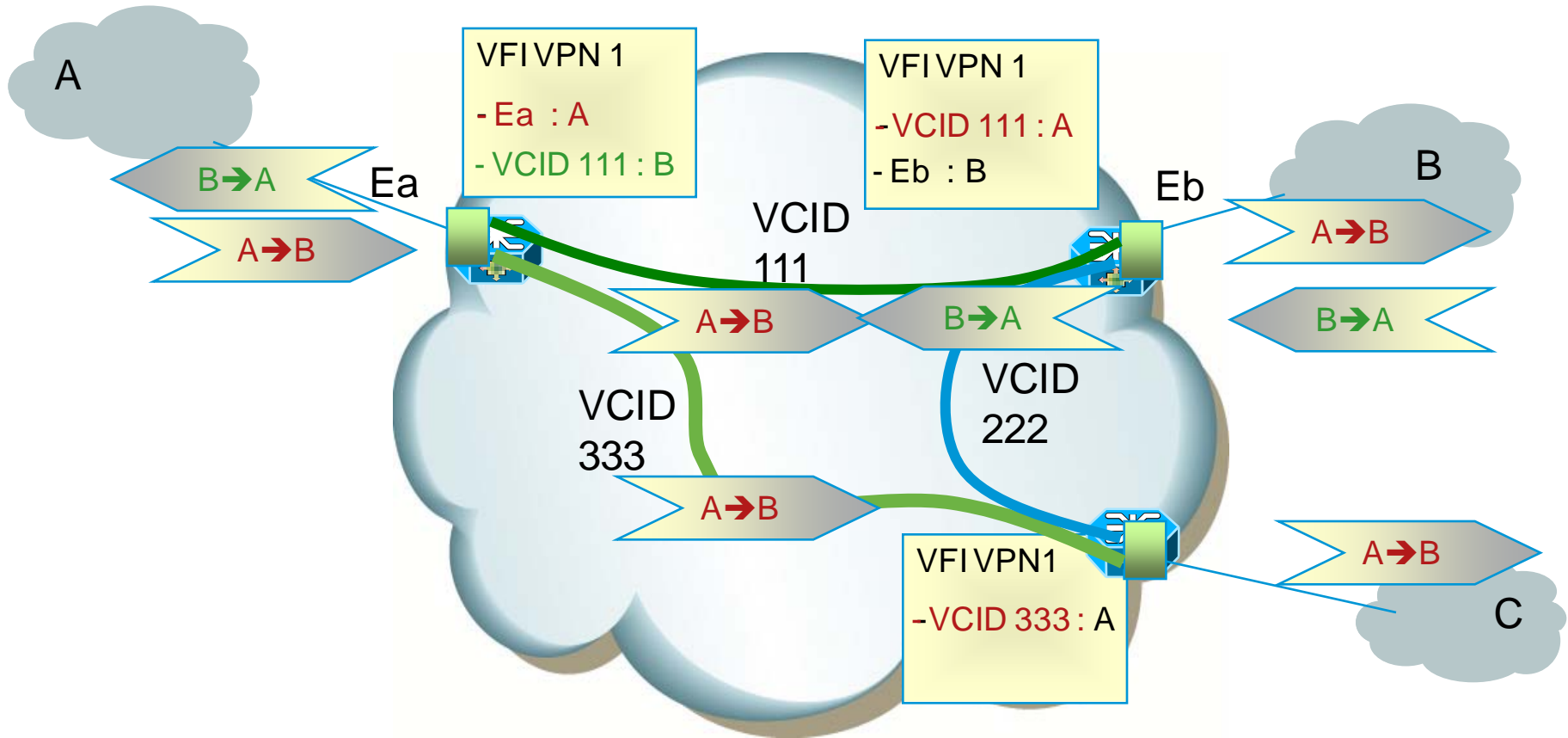
One extended bridge-domain built using:

- VFI = Virtual Forwarding Instance (VSI = Virtual Switch Instance)
- PW = Pseudo-Wire
- SVI = Switch Virtual Interface
- xconnect

VPLS

L2 Signalling and Forwarding (aka Transparent-Bridging)

A VSI/VFI operates like a conventional L2 switch!



VPLS

Split-Horizon for Loop Avoidance

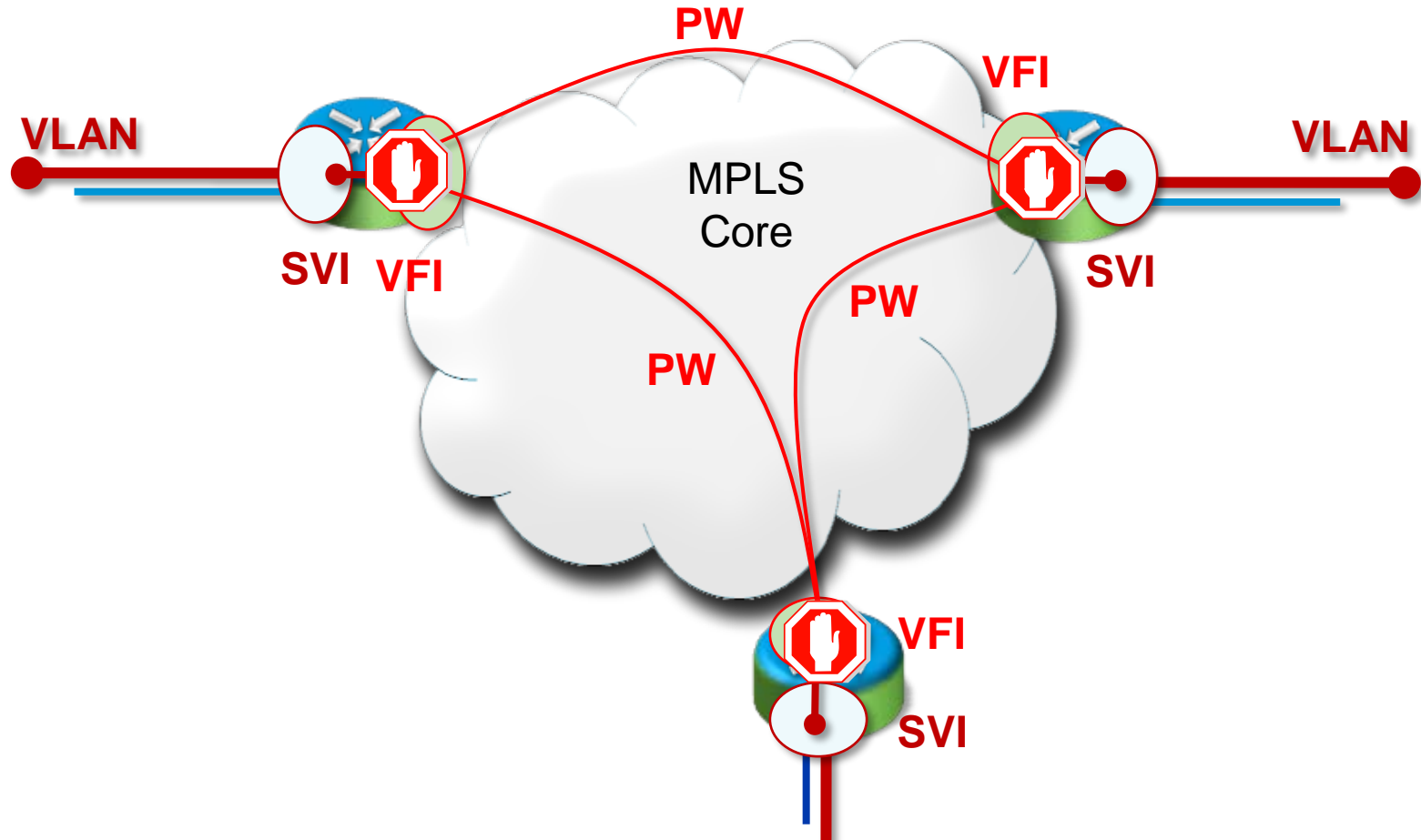
- A packet will never be bridged from a PW to an other PW in the VFI
- Assuming PW full-mesh in a VFI:
 - Full reachability
 - Core link back-up
 - No core L2 loop
 - No need for a loop prevention core STP

Important remark:

Split-Horizon protects against core loop, but Split-Horizon does not protect against global loops due to dual homing of edge devices to PE

VPLS

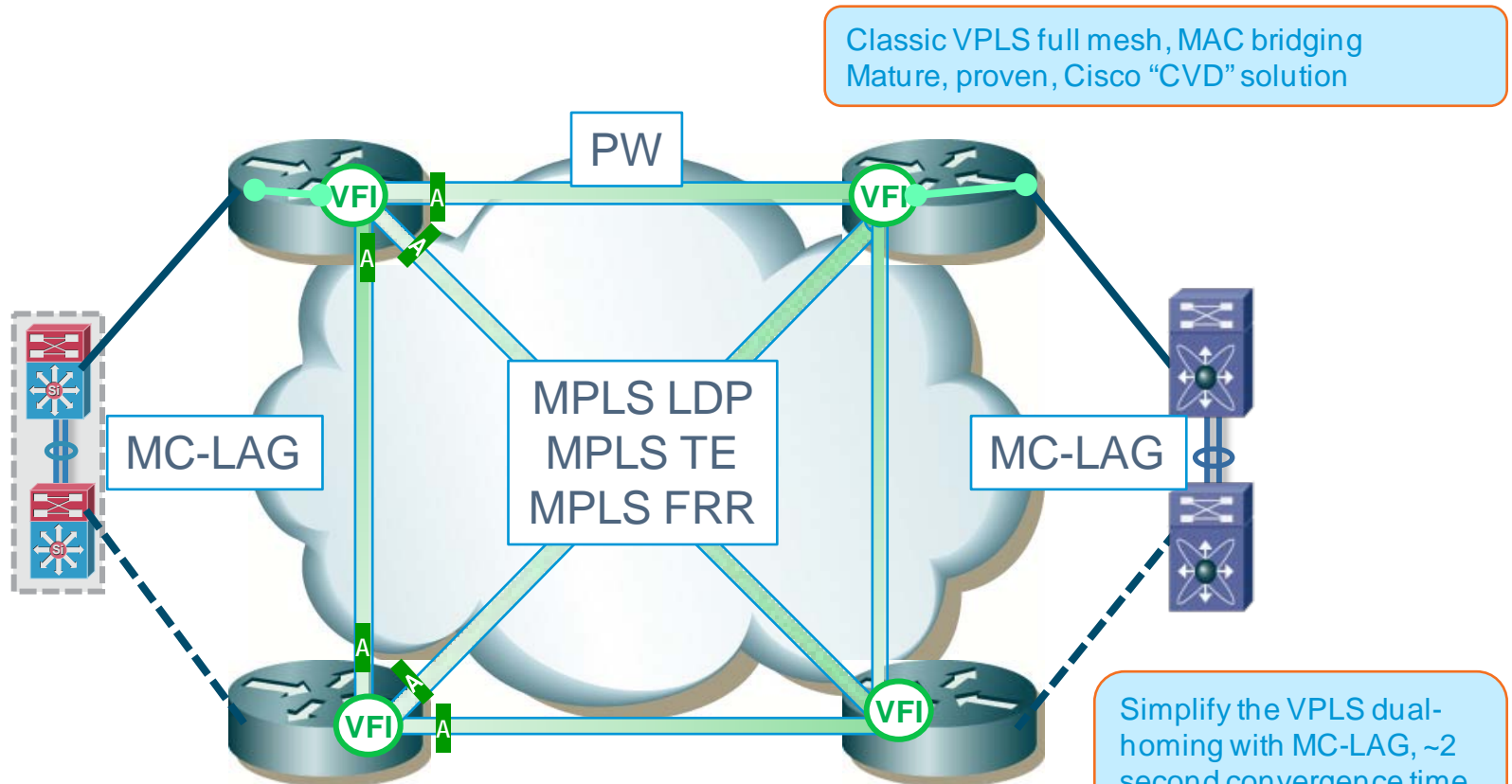
Applying Key DCI Functions with VPLS



- BPDUs are not transmitted by default
- Storm-control is on ingress link
- FHRP isolation to allow active/active default gateway + localization

High Scale VPLS DCI

The Main Model



Main functional elements:

- Standard MPLS core
- Standard VPLS
- Multi-Chassis Etherchannel access circuit

High-Scale DCI Key Requirements

Building the Data Center Cloud

Resilient

- ✓ <50ms~1sec service convergence
- ✓ Multi-Homing
- ✓ Load balancing: per VLAN/per flow
- ✓ L2 domain isolation, Storm control

Scale

- ✓ VLAN, MAC, VPLS PW Scale
- ✓ FIB, ARP Scale

SLA

- ✓ Guaranteed QoS
- ✓ Security: OAM, etc.

Perf

- ✓ Dense 10G Line Rate Forwarding
- ✓ 100GE roadmap
- ✓ Multicast replication efficiency



Cloud Dispersed Data Centers to Empower:

- Distributed applications
- Pool and maximize global resources
- Business continuity

Advanced VPLS (A-VPLS)



Easy-to-use CLI for VLAN extension

Reduced complexity!

Easy Configuration

Inter-operates with Standard VPLS
Without flow based LB

VSS single chassis redundancy

Fast sub-second failover

Simplified Redundancy

A-VPLS

L2/L3/L4 Flow Based Balancing

- DC edge to aggregation
- DC edge to WAN
- WAN core

Optimal Loadbalancing

Ethernet LAN extension over MPLS or IP:

- "Any flow Any Link" load-balancing
- Multipoint loop-free connectivity

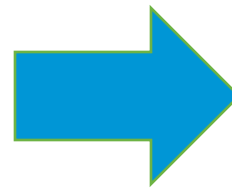
A-VPLS Simplified Configuration

Traditional VPLS Config (Repeat per VLAN)

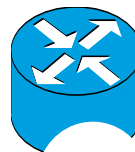
```
12 vfi for_10 manual
   vpn id 10
   neighbor 2.2.2.2 encap mpls
   neighbor 3.3.3.3 encap mpls
!
12 vfi for_20 manual
   vpn id 20
   neighbor 2.2.2.2 encap mpls
   neighbor 3.3.3.3 encap mpls
!
.....
!
interface Vlan10
  xconnect vfi for_10
!
interface Vlan20
  xconnect vfi for_20
.....
```

A-VPLS Simplified Config (Configure once)

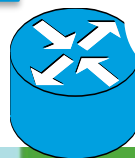
```
pseudowire-class c11
  encap mpls
  ! enable ML PW (ECMP LB)
  load-balance flow
  ! enable FAT PW
  flow-label enable
!
interface virtual-ethernet 1
  ! transport configuration
  transport vpls mesh
  neighbor 2.2.2.2 pw-class c11
  neighbor 3.3.3.3 pw-class c11
!
! service configuration
switchport
switchport mode trunk
switchport trunk allowed vlan
range 10 to 2000
```



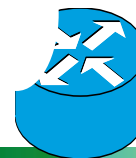
PE1 (1.1.1.1)



IP/MPLS

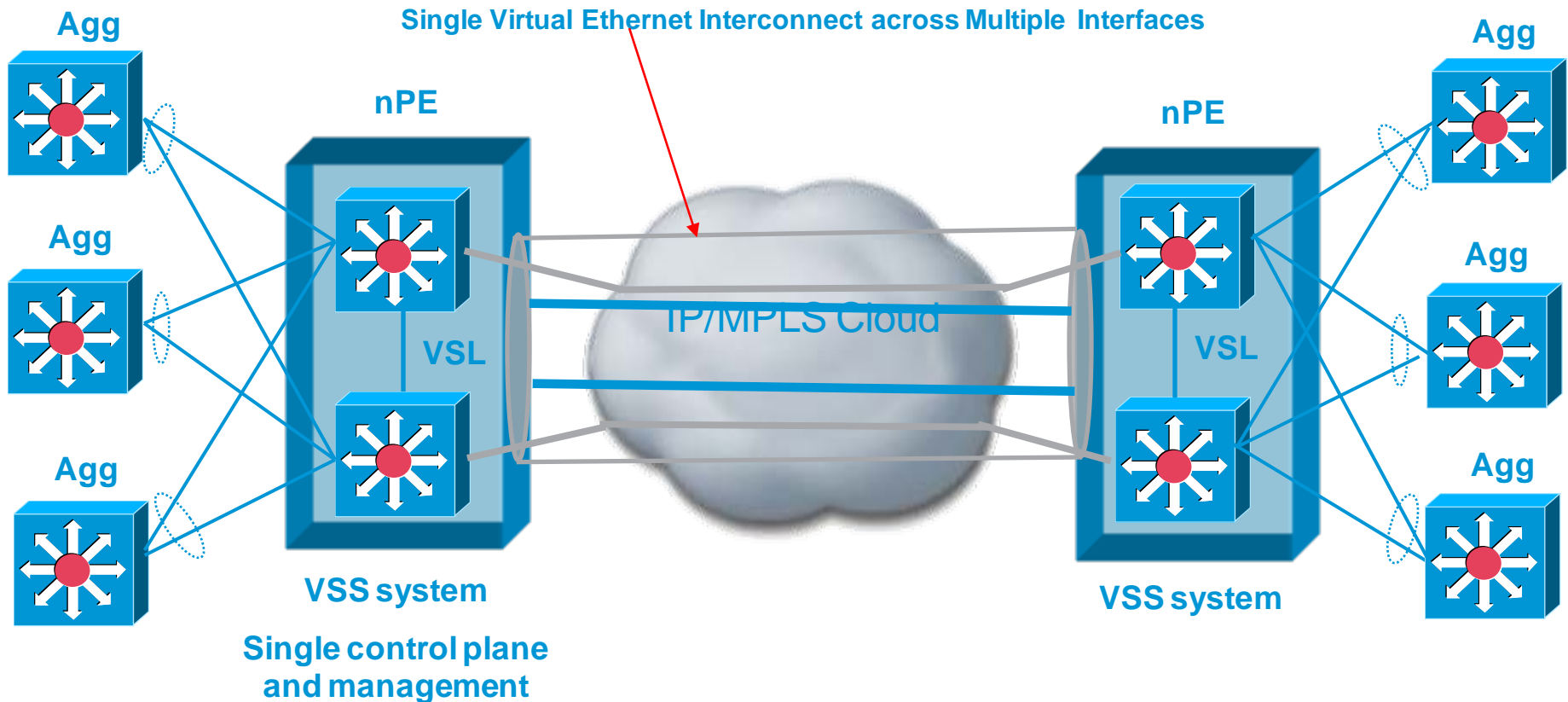


PE2 (2.2.2.2)



PE3 (3.3.3.3)

Simplified Redundancy



Agenda

- Data Center Interconnect (DCI):
 - Business Drivers
 - DCI Solution Components
 - SAN Extension
 - LAN Extension
 - Path Optimization
- LAN Extension Technologies
 - Overview
 - Ethernet Based
 - Overlay Transport Virtualization
 - A-VPLS
 - VPLS
- Conclusion and Q&A

Choosing an Appropriate DCI Design

- Number of VLANs that need to be extended across data centers
- Number of data centers that need to be interconnected
- Total amount of intra-data center bandwidth required
- Convergence and recovery times in failure scenarios
- Number of servers (MAC address scalability)
- Platform capable of providing existing data center features
- Possibility of leveraging existing network equipment

LAN Extension

Solution to Product Portfolio Table

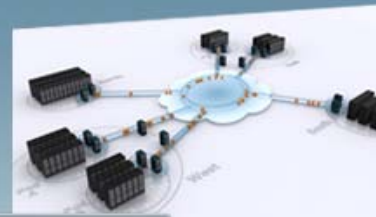
	ASR 1000	Cat 6500	Nexus 7000	ASR 9000
Ethernet Based				
vPC			✓	
VSS		✓	✓	
MPLS Based				
EoMPLS	✓	✓	Q4 2012	✓
EoMPLSoGRE	✓	✓ (requires SIP card support)		
VPLS		✓	Q4 2012	✓
IP Based				
OTV	Q4 2011		✓	

Data Center Interconnect

Where to Go for More Information

Extend Layer 2 Applications

Overlay Transport Protocol (OTV) technology supports transparent workload mobility within and across dispersed data centers. (3:54 min)



Watch Video 

Overview

In-Depth

Products

Resources

Get Instant Workload Mobility Among Data Centers

Cisco Data Center Interconnect (DCI) solutions can help your IT organization meet business continuity and corporate compliance objectives.

- Reduce the business impact of disaster events and help ensure business continuity
- Improve productivity through enhanced application and data availability
- Meet corporate and regulatory compliance needs and improve data security

These solutions transparently extend LAN and SAN connectivity and provide accelerated, highly secure data replication, server clustering, and workload mobility between geographically dispersed data centers. This enhances business resilience, and helps enable application and data mobility between data centers, while maintaining operational consistency.

Featured Products



Cisco Nexus 7000 Series LAN Extension

Simplify Layer 2 applications across distributed data centers.



Cisco Catalyst 6500 Series LAN Extension

Deliver high-performance, scalable Layer 2 extension with subsecond convergence.



Cisco MDS 9500 Series SAN Extension Over IP

Gain an integrated, cost-effective, reliable business continuance solution.

<http://www.cisco.com/go/dci>

Thank you.

