Enterprise Network Virtualization - Path Isolation System Assurance Guide

Cisco Validated Design

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
http://www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 527-0883

Network Virtualization - Path Isolation Design (Campus MPLS VPN)

Preface

The aim of this document is to accelerate customer deployments of the *Network Virtualization - Path Isolation (Campus MPLS VPN).*

It presents the validation and recommendations for the deployment architectures outlined in the *Network Virtualization - Path Isolation Design Guide.*

Table 1 Modification History

Date	Comment
Feb, 2008	Initial Release

Definitions

This section defines words, acronyms, and actions which may not be readily understood.

Table 2 Definitions

Term	Definition
AP	Access Point
AS	Autonomous System
BGP	Border Gateway Protocol: Inter-domain routing protocol that exchanges reachability information with other BGP systems.
BPDU	Bridge Protocol Data Unit
CE	Customer Edge router: A router that is part of a customer network and that interfaces to a Provider Edge (PE) router. CE routers are not aware of associated VPNs.
CEF	Cisco Express Forwarding
CVD	Cisco Validated Design
DHCP	Dynamic Host Configuration Protocol
DNS	Domain Name System
ECMP	Equal Cost Multipath

Table 2 Definitions

Term	Definition
EIGRP	Enhanced Interior Gateway Routing Protocol
FHRP	First Hop Redundancy Protocol
FTP	File Transfer Protocol
GLBP	Gateway Load Balancing Protocol
GRE	Generic Routing Encapsulation
HA	High Availability
HTTP	Hyper Text Transfer Protocol
IGP	Interior Gateway Protocol
LDP	Label Distribution Protocol
MP-iBGP	Multi-Protocol internal BGP
MPLS	Multi-protocol Label Switching
NHRP	Next Hop Resolution Protocol
NSF	Nonstop Forwarding
P	Provider router: A router that is part of a service provider's network resides inside the core of the service provider and provide inter connectivity to PE routers
PE	Provider Edge router: A router that is part of a service provider's network connected to a customer edge (CE) router. All VPN processing occurs in the PE router
PIN	Place in Network
PPS	Packet Per Second
POP3	Post Office Protocol 3
PVRST+	Per-VLAN Rapid Spanning Tree (PVRST+)
QoS	Quality of Service
SIP	SPA Interface Processor
SNRD	Solution Reference Network Design
SP	Service Provider
SPA	Shared Port Adapters
SONA	Cisco Service Oriented Network Architecture
SSO	Stateful Switchover
StackWise	Stack of switches are united into a single logical unit using special stack
STP	Spanning Tree Protocol
SVI	Switch Virtual Interface
UDLD	Unidirectional Link Detect Protocol
VLAN	Virtual Local Area Network

Table 2 Definitions

Term	Definition
VPN	Virtual Private Network: A secure IP-based network that shares resources on one or more physical networks. A VPN contains geographically dispersed sites that can communicate securely over a shared backbone.
VRRP	Virtual Router Redundancy Protocol
VRF	VPN routing / forwarding instance. A VRF consists of an IP routing table, a derived forwarding table, a set of interfaces that use the forwarding table and a set of rules and routing protocols that determine what goes into the forwarding table. In general, a VRF includes the routing information that defines a customer VPN site that is attached to a PE router.
802.1D	Spanning Tree Protocol (STP, IEEE 802.1D) standard
802.1w	Rapid Spanning Tree Protocol (RSTP; IEEE 802.1w)
802.1Q	Industry-standard trunking encapsulation (IEEE 802.1Q)

CONTENTS

FIGURES

TABLES

CHAPTER 1

Executive Summary

This document describes the validation of the *Network Virtualization - Path Isolation Design Guide* in a customer representative Multilayer Campus network environment by expanding the solution test coverage of scalability, High Availability and network convergence.

The Cisco® Validated Design Program (CVD) consists of systems and solutions that are designed, tested, and documented to facilitate faster, more reliable and more predictable customer deployments. These designs incorporate a wide range of technologies and products into a broad portfolio of solutions that meet the needs of our customers. For more information on the Cisco CVD program please refer to:

http://cisco.com/en/US/partner/netsol/ns741/networking_solutions_program_home.html

This test activity supports the goals of the Cisco Validated Design program by extending coverage of CVDs, combining CVDs and exploring interactions between them, as well as developing sustaining to extend the lifecycle of Network Systems in a customer representative environment. The extended coverage of designs, combined with the sustaining capability result in recommended releases that ensure improved quality and a successful customer deployment experience.

The test program was executed by following a formal test process that ensures consistency of operation, quality of results and value for our customers.

The following are key aspects of the test process:

- All collateral is reviewed and updated for general deployment
- Solution requirements are tested and results are documented according to a formal process that includes a cross-functional team of stakeholders.
- High quality standards are met (Zero observable operationally impacting defects within the given test parameters, that is, no defects that have not been resolved either outright or through software change, redesign, or workaround (refer to reference test plan for specific details)
- A detailed record of the testing conducted is generally available to customers and field teams, which provides:
  - Design baseline that provides a foundational list of test coverage to accelerate a customer deployment
  - Software baseline recommendations that are supported by successful testing completion and product roadmap alignment
- Detailed record of the associated test activity that includes configurations, traffic profiles, memory and CPU profiling, and expected results as compared to actual testing results. Design recommendations and test results undergo detailed review by Subject Matter Experts (SMEs) within each technology area.

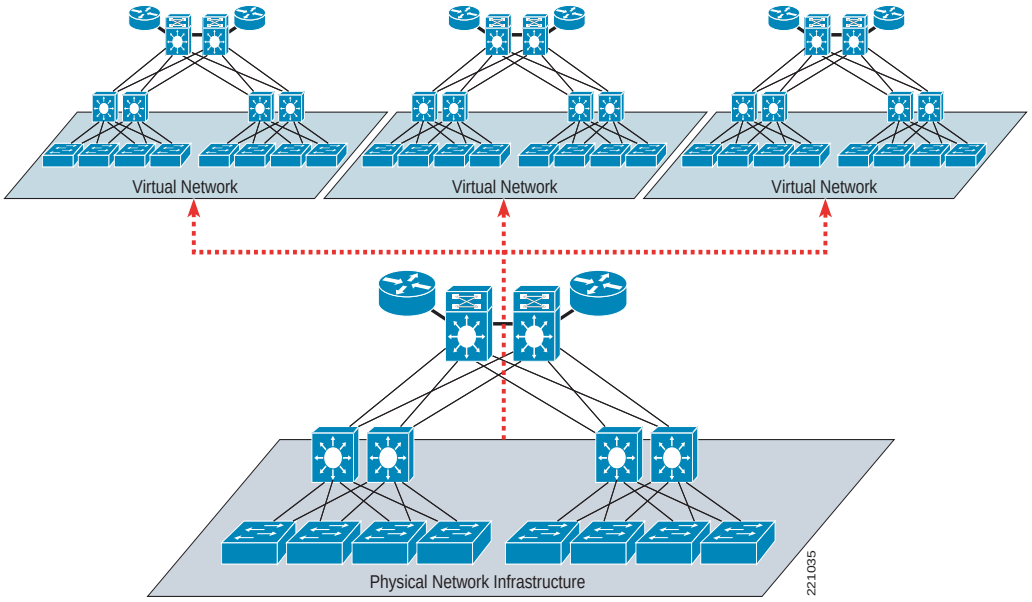
CHAPTER 2

Network Virtualization Solution Overview

2.1 Network Virtualization

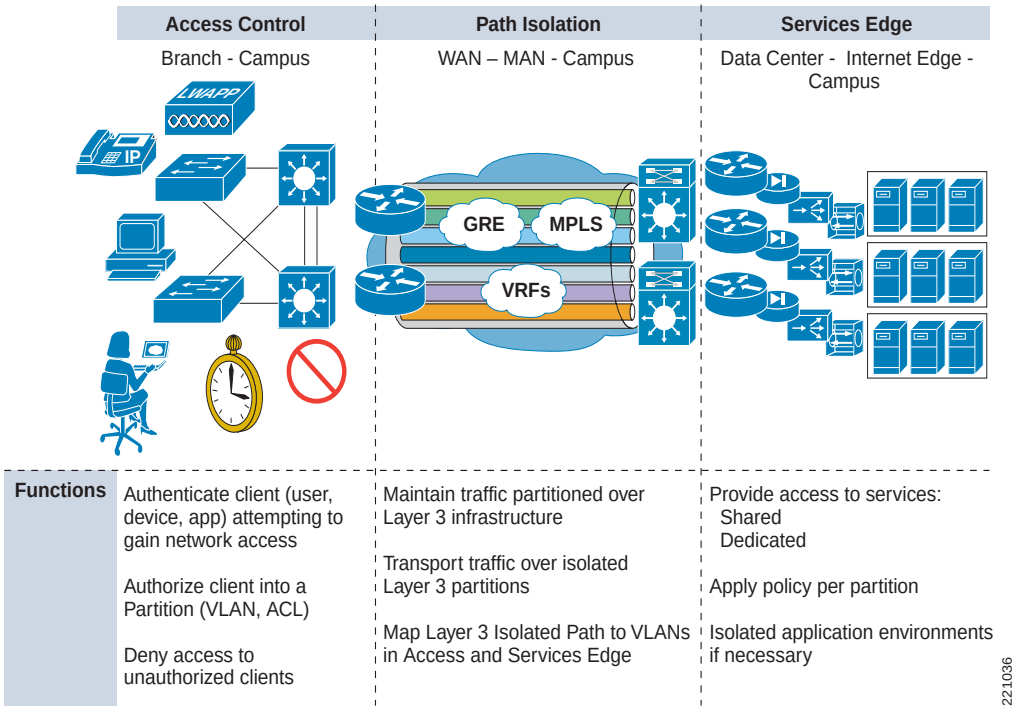
Network Virtualization is one component of the overall Cisco Service Oriented Network Architecture (SONA) that provides guidelines to accelerate applications, business processes, and profitability. Network Virtualization is a cohesive, extensible architecture that allows customers to logically partition their network infrastructure as shown in Figure 2-1. Network Virtualization simplifies network operations by enabling customers to securely share a common network infrastructure between groups of users, applications, and devices. The use of a common infrastructure places an increased emphasis on security in order to protect assets and satisfy regulatory and privacy concerns.

Figure 2-1 Network Virtualization Overview



The architecture of Network Virtualization has three main components: Access Control, Path Isolation, and Services Edge. The components highlighted in Figure 2-2 are dedicated to specific functional areas.

Figure 2-2 Network Virtualization Framework



Access Control is responsible for authenticating and authorizing devices connecting at the edge of the network. Access Control allows customers to assign devices to a specific network "segment," which usually corresponds to deploying a device in a dedicated VLAN.

Services Edge is responsible for centralizing policy enforcement points where it is possible to control and restrict communications between separate logical partitions or access to services that can be dedicated or shared between virtual networks.

Path isolation is an overlay network and refers to the creation of independent logical traffic paths to isolate traffic between users belonging to separate groups (example guest and partners) over a shared physical network infrastructure.

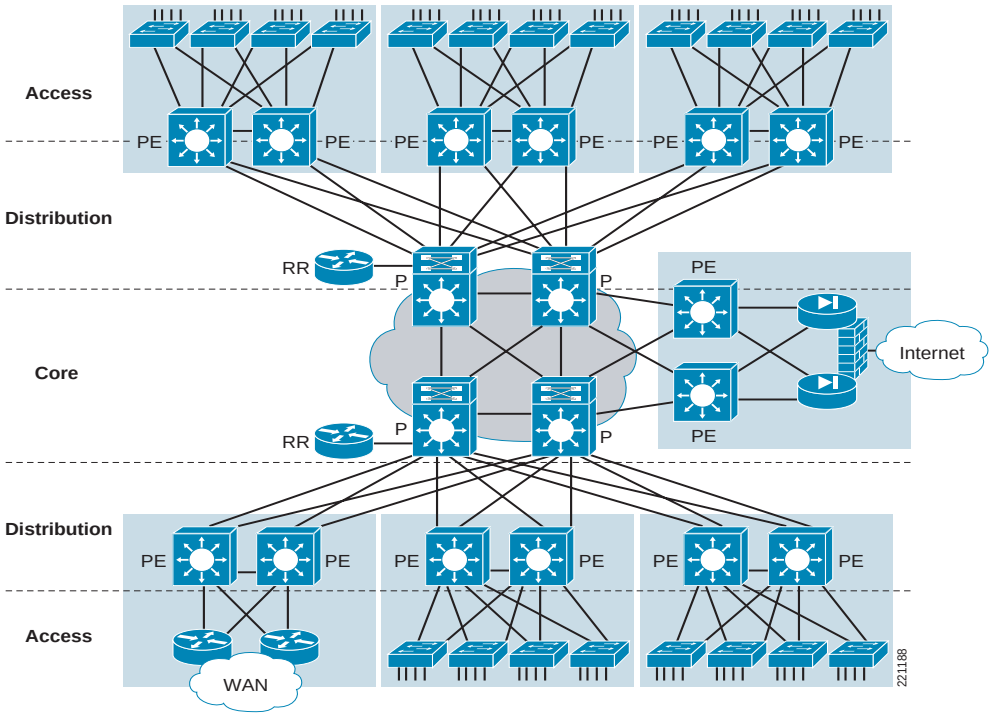
There are several mechanisms to deploy Path Isolation in the Campus network such as dynamic ACLs, VRF-Lite with GRE tunnels and MPLS VPN. The main advantage of an MPLS VPN when compared to the other Path Isolation technologies is the capability of dynamically providing any-to-any connectivity without facing the challenges of managing many point-to-point connections (for example, when using GRE tunnels). MPLS VPN provides connectivity inside each logical partition with the speed of provisioning and scalability found in no other protocol.

2.1.1 Campus MPLS VPN Path Isolation

Multi-Protocol Label Switching (MPLS) has been deployed by Service Providers to provide VPN services for customers. Service Providers have used MPLS VPN to create tunnels across their backbone networks for multiple customers.Iindividual customer traffic is carried on a common service provider network infrastructure. Using the same principle, MPLS VPN can be deployed inside the enterprise

Campus network as shown in Figure 2-3 to addresses new requirements such as network isolation, address transparency and shared services in the most scalable way while still leveraging the benefits and flexibility of IP for the existing Voice, Video and Data services.

Figure 2-3 Typical High Level Campus MPLS VPN Network



Campus MPLS VPN is an overlay network in a Multilayer Campus environment. MPLS functional roles and positioning for the network devices are shown in Figure 2-3 and as defined below:

- Provider Edge (PE): Distribution Devices
- Provider (P): Core Devices
- Route Reflector (RR): New Devices with respect to Multilayer Campus network.
  - iBGP rules require that all PEs within an autonomous system be fully meshed. For large networks, this requirement represents a severe scaling problem. Route reflectors (RRs) handle the scaled iBGP connectivity and distribute route information to the PEs. Route Reflectors avoid a need of building a fully-meshed, direct peering between PEs. As a result, the number of BGP sessions and connections is greatly reduced.
- Customer Edge (CE): There are actually no true CE devices, because the only devices connecting to the PE are Access Layer switches that perform only L2 functions.

Note Terms such as Distribution and PE and P and Core are used interchangeably.

Details about MPLS VPN technology and how to deploy Network Virtualization Path Isolation using MPLS VPN in Campus can be found in *Network Virtualization - Path Isolation Design Guide*.

CHAPTER 3

Network Virtualization Validation Strategy

The validation strategy of CVD testing is to build an End-to-End, High Availability MPLS VPN Campus network and to validate requirements associated with Campus Path Isolation services. The Network Virtualization - Path Isolation Design Guide (Campus MPLS VPN) was validated in manual and automated test environments.

3.1 Network Topology

MPLS VPN was overlaid on a High Availability multilayer Campus network and based upon the following Campus design guides:

- *Network Virtualization - Path Isolation Design Guide*
- *Designing a Campus Network for High Availability*
- *High Availability Campus Network Design Guide - Implementing Supervisor Redundancy Using NSF, SSO, and StackWise*
- *High Availability Campus Recovery Analysis*

In the Core Layer, Cat6500 platforms with dual chassis, dual supervisors (SUP-720-BXL) and ten gigabit Ethernet (10GE) links played the role of Provider (P).

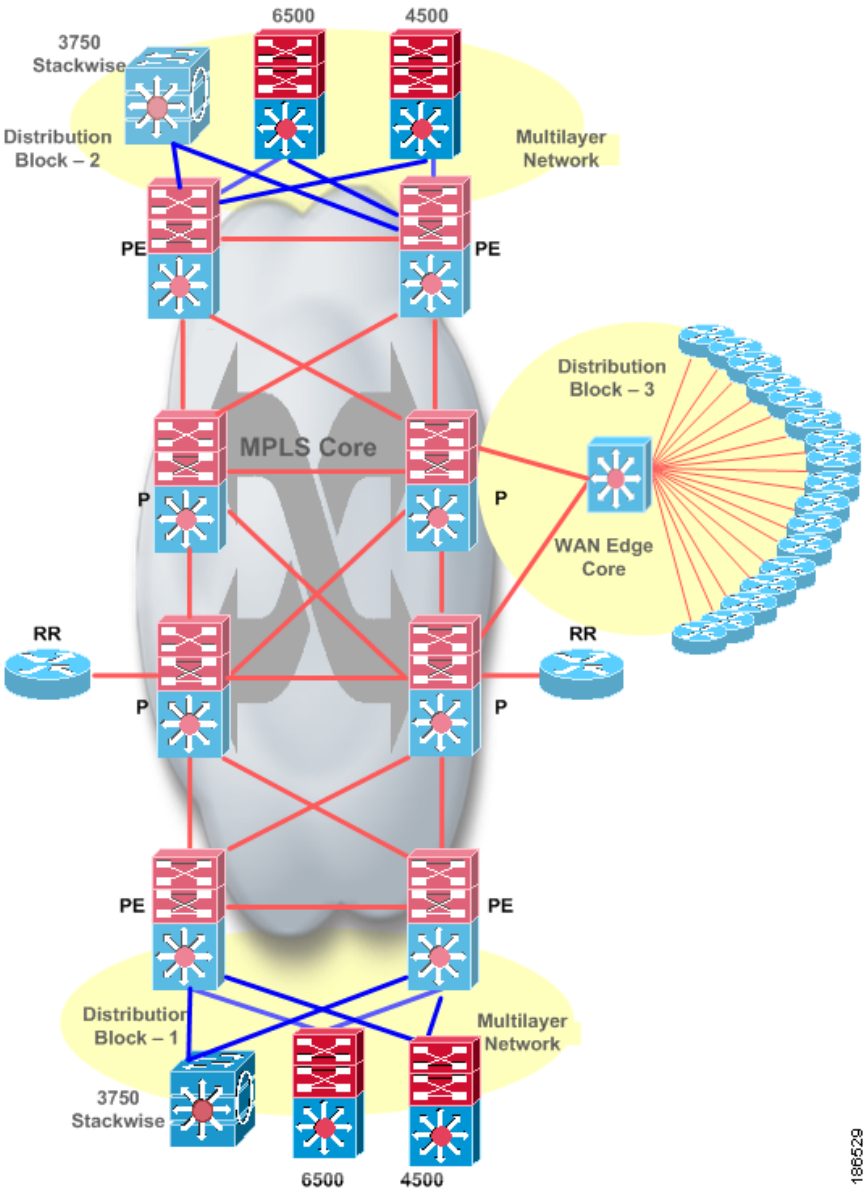
The Route Reflector role with gigabit Ethernet links to core devices was deployed on a C7200-VXR with NPE-G1 processor.

In the Distribution Layer, Cat6500 platforms with dual chassis, dual supervisors (SUP-720-BXL) and ten gigabit Ethernet (10GE) links played the role of Provider Edge (PE). Emulated distribution blocks were created using IXIA tools in order to scale EIGRP routes

In the Access Layer, Cat6500 platforms with dual supervisors (SUP-32), Cat4500 with dual supervisors (4516-10GE) and Cat3750E with (StackWise) played the role of Layer2 switches.

In order to validate the Network Virtualization – Path Isolation Design Guide, a High Availability (HA) MPLS VPN Multilayer Campus network was built as illustrated in Figure 3-1.

Figure 3-1 High Availability Campus MPLS VPN Path Isolation Topology



3.2 Traffic Profile

The traffic profiles includes stateful and stateless traffic.

The stateful traffic is categorized as background traffic. Background traffic is used for simulating a live network and is included for every automated test case. The background traffic was created based on the *Enterprise QoS Solution Reference Network Design Guide*.

- 100 Stateful sessions (TELNET + FTP + HTTP + DNS + POP3)
- 100 Mbps QoS traffic includes Voice, Video, Call Control, bulk data, critical data and best effort traffic.

The stateless traffic is used for measuring network convergence in manual testing as shown below:

- Traffic flow – Upstream and Downstream
- Traffic rate – 50 pps for every flow
- Traffic type (a) – Global / IPv4 traffic - 3000 routes / flows (50 EIGRP neighbors, 60 routes/neighbor)
- Traffic type (b) – VPNv4 traffic - 5000 routes / flows (50 VRFs and 100 routes/VRF)
- Traffic state – Stateless

3.3 Hardware and Software Device Information

Table 3-1 Hardware and Software Device Information

Hardware Platform	Role	Software Version	Line Cards/Interfaces
Cisco 3750E-24TD	Access L2*	12.2(37)SE	StackWise (3750Es)
Cisco 4507R	Access L2	12.2(31)SGA2	WS-X4516-10GE (Dual) WS-X4548-GB-RJ45V
Cisco 6506	Access L2	12.2(33)SXH1	WS-SUP32-10GE-3B (Dual) RJ45 WS-X6548-GE-TX
Cisco 6506	Distribution (PE) L2 / L3	12.2(33)SXH1	WS-SUP720-3BXL (Dual) WS-F6700-DFC3BXL WS-F6K-PFC3BXL WS-SUP720 (MSFC3) WS-X6704-10GE WS-X6748-GE-TX
Cisco 6509	Core (P) L3	12.2(33)SXH1	WS-SUP720-3BXL (Dual) WS-F6700-DFC3BXL WS-F6K-PFC3BXL WS-SUP720 (MSFC3) WS-X6704-10GE WS-X6748-GE-TX
Cisco 7206VXR	Route Reflector (RR)	12.4(11)T	NPE-G1

Note ***Access Layer2**: Only the Cisco **3750E device** was used in all failures characterization scenarios. Please refer to the Network Virtualization Solution Characterization section.

3.4 Test Types

Validation tests are divided into the following categories:

- System Integration
- Scalability
- Availability
- Network Convergence
- Reliability

3.4.1 System Integration

System Integration has two major components, feature combination and feature interaction.

Feature combination focuses on testing a feature when various combinations of other features are enabled. Feature interaction tests were conducted to verify dependencies between features.

Features combination and features interaction were conducted and verified during validation tests as shown in Table 3-2.

End-to-End traffic (stateful and stateless) was validated for data, voice and video using IXIA and SmartBits tools.

Health checks were performed before and after tests to monitor for failure or system degradation. These checks included memory and CPU utilization, tracebacks, memory alignment errors, interface errors, line card status and syslog messages.

Table 3-2 Device Role and Feature Information

Role	Feature/Technology
Access L2	PVRST+ Loopguard Portfast UDLD Vlan Trunking SSO (4500/6500) StackWise (3750E)
Distribution L2	PVRST+ Loopguard Portfast UDLD Vlan Trunking
Distribution (PE) L3	HSRP EIGRP EIGRP NSF MPLS LDP MPLS LDP Graceful Restart MPLS LDP Session Protection BGP, MP-iBGP BGP Graceful Restart BGP Next-Hop Tracking ECMP IGP/BGP NSF/SSO

Table 3-2 Device Role and Feature Information

Core (P) L3	EIGRP EIGRP NSF MPLS LDP MPLS LDP Graceful Restart MPLS LDP Session Protection NSF/SSO
Route Reflector (RR)	EIGRP MP-iBGP BGP Next-Hop Tracking BGP IPv4 Unicast Multipath iBGP

3.4.2 Scalability

A key focus of CVD validation is to build a scalable MPLS VPN Campus network by simulating a large number of distribution blocks as illustrated in Figure 3-1 and to measure the network convergence in the scaled environment. For EIGRP, the network was scaled up to 50 neighbors and 3000 routes. For MPLS VPN, the network was scaled up to 5000 VPN prefixes. Details of implementing a scaled MPLS VPN network is described in the Network Virtualization Solution Characterization section.

Table 3-3 is a summary of scalability test coverage.

Table 3-3 Device Role and Network Scalability Information

Role	Network Scalability
Access L2	100 Vlans
Distribution L2	100 Vlans
Distribution (PE) L3	50 EIGRP neighbors 3000 EIGRP routes 50 VRFs 100 HSRPs 5000 VPN prefixes
Core (P) L3	50 EIGRP neighbors 60 routes/neighbors 3000 EIGRP routes
Route Reflector (RR)	5000 VPN prefixes

3.4.3 High Availability

High Availability enables network-wide protection by providing fast recovery from faults that may occur in any part of the network. In a High Availability network, hardware and software work together to enable rapid recovery from disruptions and to ensure fault transparency to users and network applications.

The key attributes of a High Availability network in the Core and Distribution Layers are supervisor redundancy with Nonstop Forwarding (NSF) and Stateful Switchover (SSO). In the Access Layer, supervisor redundancy with SSO and chassis redundancy using (StackWise) 3750E switches were implemented.

Detailed High Availability coverage is described in the Network Virtualization Solution Characterization section.

Below is a summary of High Availability test coverage.

- HSRP between Access and Distribution Layer
- NSF/SSO with dual chassis, dual supervisors, dual links in Distribution and Core Layers
- 3750E Chassis redundancy using (StackWise) in the Access Layer

3.4.4 Network Convergence

The primary focus of the network convergence testing is to measure tagged (VPN) and untagged (IP Global) traffic after network failure and restoration. Detailed convergence analysis is presented in the the Network Virtualization Solution Characterization section.

Note **Access Layer2**: Only the Cisco **3750E device** was used in all failures characterization scenarios. Please refer to the Network Virtualization Solution Characterization section.

3.4.5 Reliability

The 150-hour reliability test was executed for the entire testbed to ensure that the fully provisioned network continues to operate at the elevated utilization levels without impacting CPU utilization, memory or any operationally impacting events. Devices were monitored for tracebacks, alignment, interface errors and syslogs. End-to-End connectivity was maintained during this test.

3.5 Sustaining Coverage

On going automated regression testing provides consistent, repeatable customer representative test coverage. Automated testing also provides the ability to validate subsequent IOS software releases. Following is a summary of sustaining coverage:

- Automated test scripts for each automated test case
- Common scripts library for managing the testbed, collecting and reporting test results
- Automated procedures to capture results of the manually executed tests.

CHAPTER 4

Network Virtualization Solution Characterization

This section discusses solution characterization of Network Virtualization of Path Isolation in a Campus IP/MPLS environment.

4.1 Distribution to Access Layer Path Isolation

Network Virtualization uses 802.1Q technology to create an isolated path between the Distribution and Access Layers in Multilayer Campus network designs. Cisco Layer2 Ethernet switches can do mac-based switching within the system and establish redundant connections to Distribution Layer devices. Network Virtualization does not change widely deployed Campus hierarchical networks; it is simply overlaid on an existing Campus network.

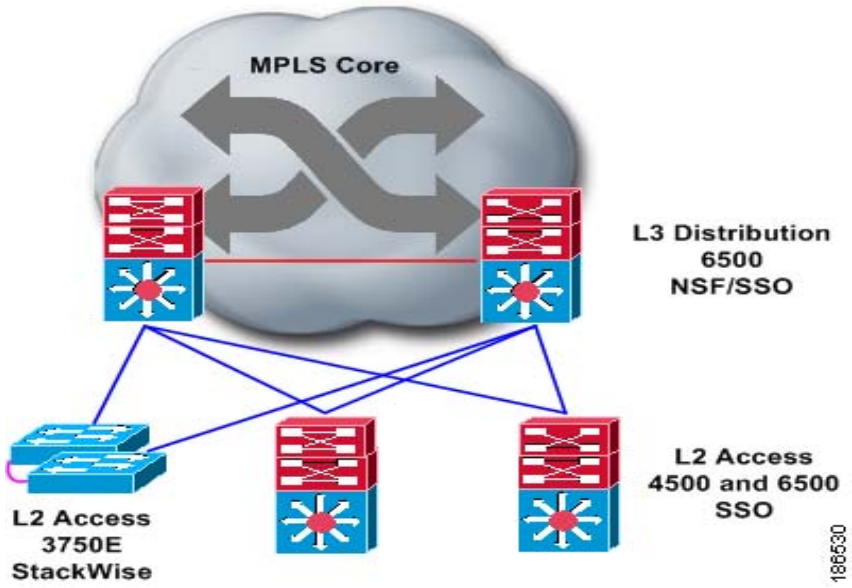
4.1.1 Building a High Availability (HA) Distribution to Access Layer Block

To implement Network Virtualization in a Multilayer Campus network, a resilient and loop-free network should be built and based on the recommended baseline configuration covered in the Multilayer Architecture Design Guides.

http://www.cisco.com/en/US/netsol/ns656/networking_solutions_design_guidances_list.html#anchor2

The 3750, 4500 and 6500 Cisco switching platforms were deployed at the Access-Layer to build a High Availability network that offers physical link, route-processor with SSO and system redundancy using (StackWise). Figure 4-1 illustrates the high availability Distribution to Access Layer deployment model.

Figure 4-1 High Availability Distribution-Access Layer Block



As network demand grows, new member switches in the same (StackWise) group can be added gracefully without impacting the existing network. StackWise technology appoints a master switch that is responsible for establishing control-plane with directly connected devices and member switches that are capable of forwarding traffic.

The Catalyst 3750E (Stackwise) must be created in a Ring topology and it is recommended that the Master switch be selected inside the ring network. Each member group uses a single physical uplink to the Distribution device.

Recommendation: The master switch should be configured statically as high priority in order for it to be selected as the Master switch. Following is a configuration example to enable (StackWise) on the Catalyst 3750-E series switch:

```
c3750(config)# switch 1 provision ws-c3750e-24td
c3750(config)# switch 2 provision ws-c3750e-24td
c3750(config)# switch 1 priority 15
Changing the Switch Priority of Switch Number 1 to 15
Do you want to continue?[confirm]
New Priority has been set successfully
```

4.1.2 Implementing Multilayer Foundation Services

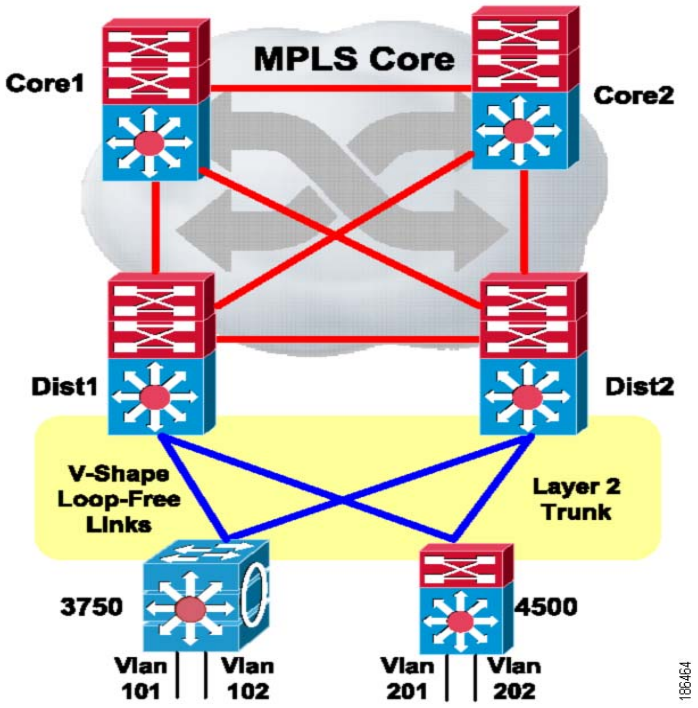
Enterprise customer planning to overlay isolated paths in a Multilayer Campus network can leverage Layer2 foundation services as recommended in the MultiLayer Design Guides to build a loop-free network and to optimize End-to-End network convergence. There are several factors that must be planned for in understanding how physical networks are to be built, how foundation services are to be implemented and function optimally to reduce convergence time during system or network failure events. The following section briefly discusses implementing foundation services on Distribution to Access Layer devices.

4.1.2.1 Building Loop-Free Spanning Tree Topology

Loop-free networks do not span VLANs across multiple Access-Layer switches in a single domain. The primary advantage of this design is that all uplinks to Distribution Layer devices are never marked "blocked" and can switch traffic across all available paths. All uplinks from the Access to the Distribution Layer are trunked and carry multiple VLANs and switch traffic based on Layer2 addresses.

Recommendation: The "V-shape" topology is the recommended topology from the Access to Distribution Layer as shown in Figure 4-2.

Figure 4-2 MultiLayer Loop-free STP topology



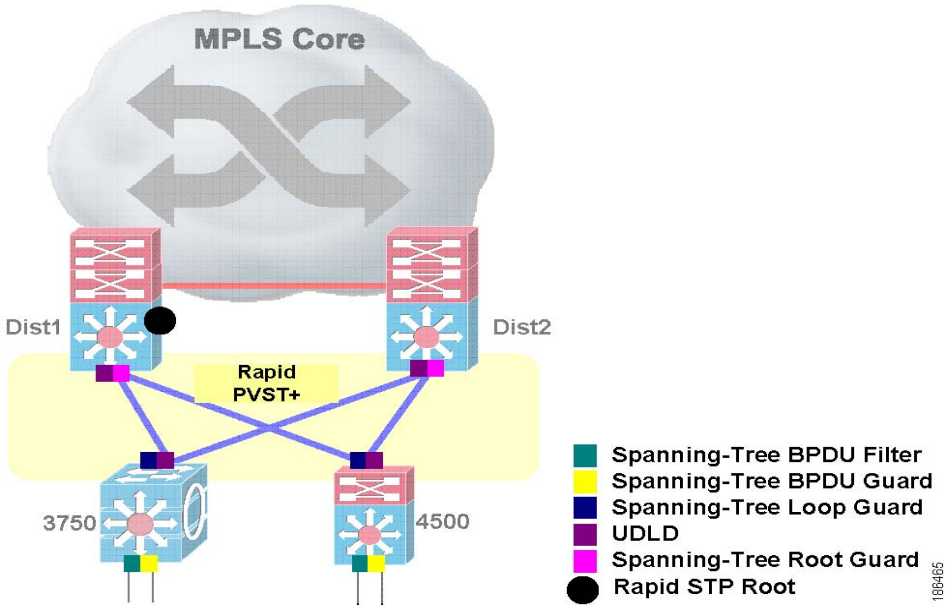
4.1.2.2 Implementing Layer2 Optimization Toolkit

The IEEE spanning-tree protocol is designed to create a loop-free Layer2 network. Cisco supports all spanning-tree protocols and has multiple solutions to secure spanning-tree topologies to reduce convergence time during failure events. The Layer2 toolkit includes sets of technologies and IOS features as illustrated in Figure 4-3 that should be implemented as a safety belt mechanism to prevent Layer2 loop in Multilayer Campus network.

The implemented Layer2 Optimization toolkit includes:

- Rapid spanning-tree – Rapid-PVST+
- Spanning-tree BPDU filter
- Spanning-tree BPDU guard
- Spanning-tree loop guard
- Spanning-tree root guard
- Rapid STP root
- Uni-directional Link Detection – UDLD

Figure 4-3 Layer2 Optimization Toolkit between Distribution-Access Layer



Cisco catalyst switching platforms support IEEE 802.1D, 802.1w Rapid PVST+ and 802.1s Multiple Spanning-tree. In building the Network-Virtualization solution, Rapid PVST+ was used because it has enhanced capability for direct and indirect link failure. Please refer to the Campus network Design Guides in the Related Documents and Links section for more information.

The following is example illustrates enabling Rapid PVST+

```
Access switch configuration
c3750(config)#spanning-tree mode rapid-pvst

Distribution switch configuration
c6506(config)#spanning-tree mode rapid-pvst
```

PortFast BPDU filtering allows administrators to prevent the system from sending or even receiving BPDUs on specified ports.

When configured globally, PortFast BPDU filtering applies to all operational PortFast ports. Ports in an operational PortFast state are supposed to be connected to hosts, that typically drop BPDUs. If an operational PortFast port receives a BPDU, it immediately loses its operational PortFast status. In that case, PortFast BPDU filtering is disabled on this port and STP resumes sending BPDUs on this port.

PortFast BPDU filtering can also be configured on a per-port basis. When PortFast BPDU filtering is explicitly configured on a port, it does not send any BPDUs and drops all BPDUs it receives.

The following is an example configuration BDPU filtering on the access port of access-layer device connected to the end device:

```
c3750(config)# interface GigabitEthernet2/0/27
c3750(config-if)spanning-tree bpdufilter enable
```

By default, Spanning-Tree builds the topology based on BPDU exchanged between bridges including the end-station configured for bridging between LAN media within a system. This may break the Multilayer Campus network and create instability in the network if the Spanning-Tree selects a root-port to the end-machine. It is recommended to disable sending and receiving BPDU on the Access ports. This technique helps to secure the Multilayer Campus network domain by preventing End devices from participating in STP domains.

The following is an example configuration on the access port of access-layer device to block BPDUs:

```
c3750(config)# interface GigabitEthernet1/0/4
c3750(config-if)spanning-tree bpduguard enable
```

The Distribution and Access Layer devices must be configured to prevent loop. Root guard must be enabled on Distribution Layer devices connected to each Access Layer device. Using this technique, Access Layer devices will never be able to function in a root device role.

It is equally important to enable loop guard on Access Layer device uplinks to the Distribution Layer to prevent the alternate/root port from being elected unless BPDU frames are detected on the link.

Following is an example configuration on the Distribution and Access-layer for the root guard and loop guard toolkit:

```
Access switch configuration:
c3750(config)#interface TenGigabitEthernet1/0/1
c6500(config-if)# spanning-tree guard loop

Distribution switch configuration
c6500(config)#interface TenGigabitEthernet7/4
c6500(config-if)# spanning-tree portfast
c6500(config-if)# spanning-tree guard root
```

When network equipment is interconnected using fiber-optic cables in a network, there is a possibility that fiber cables are connected improperly. This condition creates a uni-directional path between two devices that can cause layer2 network loops for STP and Rapid PVST+ protocols. Uni-directional Link

Detection technology resolves this problem by exchanging “hello” messages on a per link basis. When a device fails to receive UDLD “hello” messages, the port is configured incorrectly and is automatically disabled.

In a Multilayer Campus network, UDLD should be enabled on Layer2 trunk links between Distribution and Access devices.

Recommendation: The UDLD protocol functions in two different modes: slow and aggressive mode. UDLD should be enabled in aggressive mode when in global configuration mode and the UDLD “hello” message timer should be set to seven seconds.

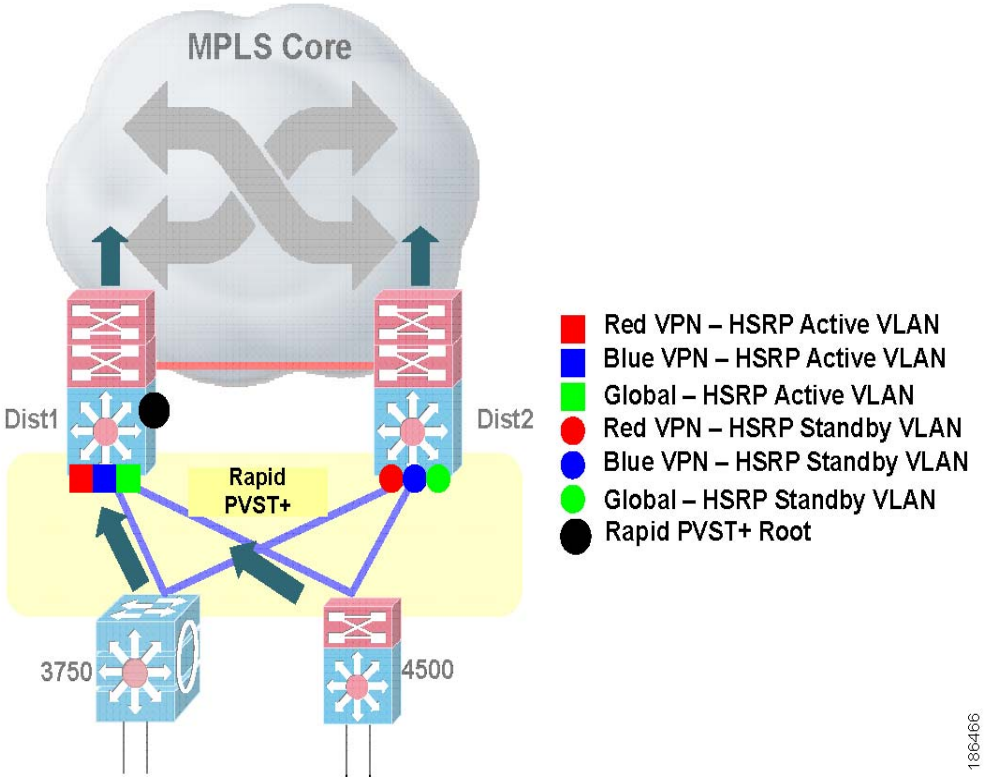
Following is an example configuration for enabling UDLD on Distribution and Access Layer devices:

```
Access switch configuration
c3750(config)#udld aggressive
c3750(config)#udld message time 7

Distribution switch configuration
c6500(config)#udld aggressive
c6500(config)#udld message time 7
```

When a HSRP is selected as the IP gateway protocol in a Multilayer Campus network, it is recommended that the same distribution device be set to function as HSRP active and in the spanning tree root role. The Rapid PVST+ protocol on the distribution block should be configured statically by increasing the PVST+ priority higher than its default setting in order for it to perform in the root device role. The Rapid PVST+ protocol can be configured as illustrated in Figure 4-4 for each VLAN that is part of the MPLS VPN or in a global forwarding table.

Figure 4-4 HSRP between Distribution-Access Layer



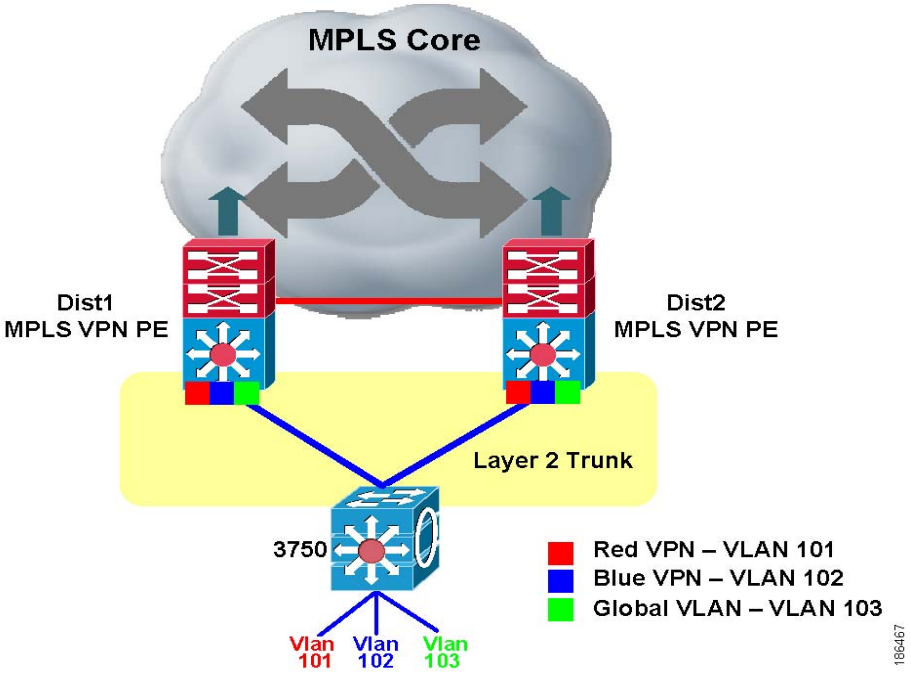
Rapid PVST+ configuration is applied in the global configuration mode of a Distribution Layer device as show in the following example:

```
c6500(config)# spanning-tree mode rapid-pvst
c6500(config)# spanning-tree vlan 1-4094 priority 24576
```

4.1.3 Implementing Path Isolation between Distribution and Access Layer

In a Multilayer Campus network, path isolation is achieved by forwarding traffic based on end-user assigned privileges. Based on privileges end-user traffic is redirected on an assigned 802.1q VLAN that is mapped to MPLS VRF at the Distribution Layer and forwarded based on virtual route forwarding (VRF) entries. Figure 4-5 shows an example of mapping end-user traffic from an Access port to a Layer2 trunk port that is assigned to a specific MPLS VRF.

Figure 4-5 Path Isolation Technique in Multilayer Campus Network



During solution characterization, the number of 802.1q VLANs was scaled to 100 at the edge of the network. Each VLAN was mapped to a unique MPLS VRF at the Distribution device.

Note Distribution and Access-layer devices are designed and fully capable to scale up to 4000 VLANs in a database, however network planners must take MPLS VRF scalability into consideration when designing Network Virtualization solutions. Currently the Cisco Catalyst 6500 is the predominant Campus positioned platform for PE and P role. PE can scale up to 512 VPNs on a single system. Just like FHRP which offers IP gateway redundancy, Cisco recommends implementing a standby IP gateway in the MPLS VPN PE role.

Creating vrf red and blue

```
ip vrf RedVPN
 rd 9001:10
 route-target export 64000:10
 route-target import 64000:10
!
ip vrf BlueVPN
 rd 9001:11
 route-target export 64000:11
 route-target import 64000:11


creating SVIs
c6500(config)#interface Vlan101

c6500(config)#interface Vlan102

attaching vrf to an SVI
c6500(config)#interface Vlan101
c6500(config-if)# ip vrf forwarding RedVPN

c6500(config)#interface Vlan102
c6500(config-if)# ip vrf forwarding BlueVPN

assigning IP to an SVI cfgs
c6500(config)#interface Vlan101
c6500(config-if)# ip vrf forwarding RedVPN
c6500(config-if)# ip address 10.0.0.1 255.255.255.0

c6500(config)#interface Vlan102
c6500(config-if)# ip vrf forwarding BlueVPN
c6500(config-if)# ip address 11.0.0.1 255.255.255.0
```

Distribution Layer devices demark the Layer2 broadcast domain and perform routing between broadcast domains and towards the MPLS Core. When there are multiple IP gateways in the same distribution block, it is important to enable FHRP protocols for an end-station to appropriately select the right gateway. Cisco supports HSRP, GLBP and VRRP protocols for gateway redundancy. HSRP redundancy protocol was implemented during validation.

Following is an example configuration that is used to characterize HSRP at the Distribution or MPLS VPN PE device:

```
c6500(config)#int Vlan102
c6500(config-if)#ip vrf forwarding blueVPN
c6500(config-if)# ip address 11.0.0.1 255.255.255.0
c6500(config-if)#standby 1 ip 11.0.0.254
c6500(config-if)# standby 1 timers msec 250 msec 750
c6500(config-if)# standby 1 priority 150
c6500(config-if)# standby 1 preempt delay minimum 180
```

Refer to following url for more detail about implementing HSRP support for MPLS VPNs:

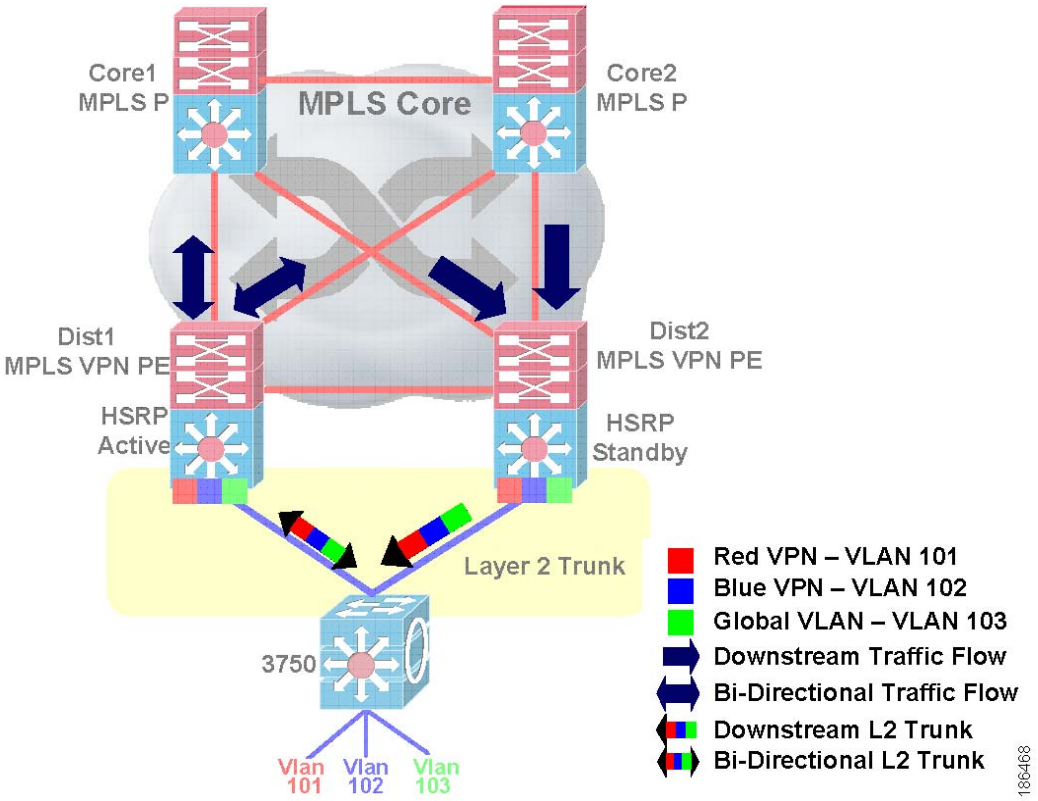
http://www.cisco.com/en/US/products/sw/iosswrel/ps1834/products_feature_guide09186a00806993c5.html

With HSRP implemented in the Multilayer Campus network, all traffic flows to a single IP gateway to forward all upstream traffic. However downstream stream traffic flows evenly load-balanced from the MPLS core and utilizes both Layer2 trunk links to send traffic to the same devices. Global traffic is

untagged should continue to switch traffic based on destination IP address lookup. This ensure that existing infrastructure like IP Telephony are not impacted when Network Virtualization is implemented in the network.

Figure 4-6 depicts traffic flow with HSRP implemented in a Multilayer Campus network and performing load-balancing in the IP/MPLS core.

Figure 4-6 Isolated Forwarding Path in Multilayer and IP/Core network



4.2 Distribution to Access Layer Failure Characterization

This section presents solution characterization data that was generated from emulating real-world failure scenarios occurring in the Access and Distribution Layers. All the Network Virtualization solution characterization reports were generated based on the traffic profile described in 3.2 Traffic Profile along with the scalability information defined in Table 3-3.

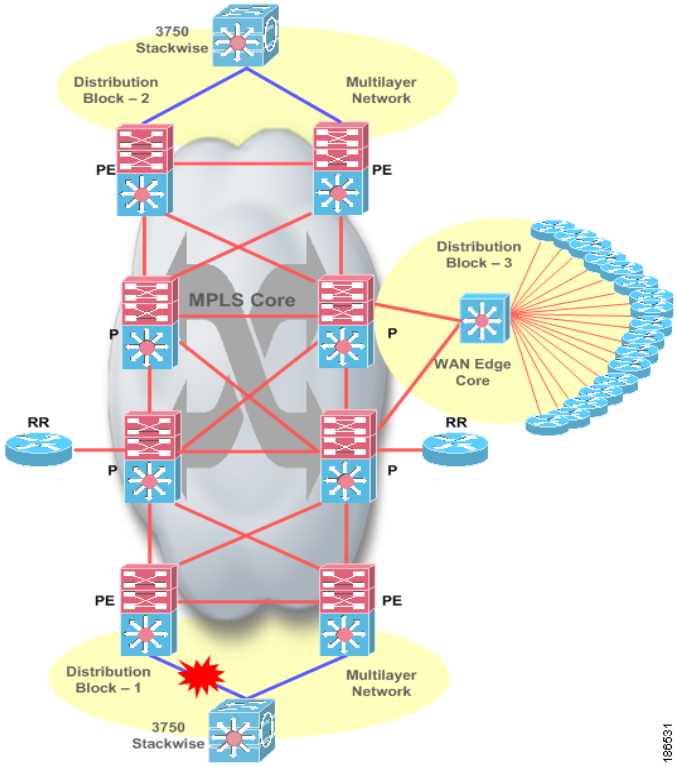
Note During Distribution to Access Layer Failure Characterization only Cisco 3750E link and device failures were characterized.

4.2.1 HSRP Link Failure

The Multilayer Campus network is deployed with HSRP as the gateway redundancy protocol. All uplink traffic is forwarded to an HSRP Active device. Both the HSRP active and standby device maintain per-VLAN and route entries to load-balance downstream traffic arriving from the IP/MPLS Core network.

When the link between the Distribution device to the Access Layer device fails, MPLS VPN and Global traffic is impacted. When a Distribution to Access link failure occurs in the network, a series of control-plane update events occur to update the forwarding paths. These include BGP update, VPN label withdrawal and forwarding table updates. Since Global traffic is untagged, removal of ECMP routes and CEF update events will change the forwarding path to reroute traffic to a single distribution device instead of load-balancing. Figure 4-7 illustrates the Active HSRP Link failure topology.

Figure 4-7 Active HSRP Link Failure Topology



Active HSRP Link Failure Analysis

Upstream traffic:

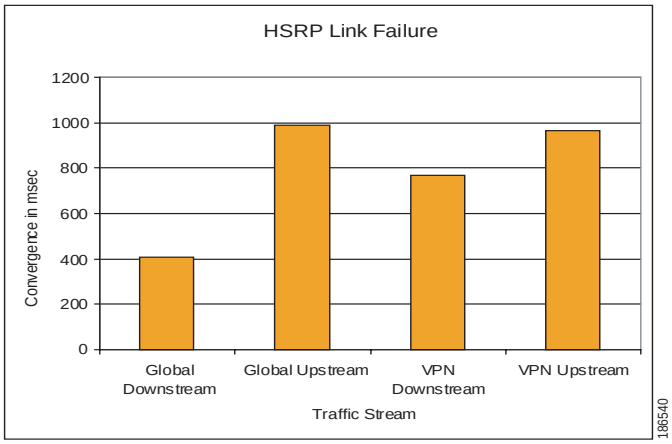
In normal state, all the upstream traffic is sent to the active HSRP device to reach the destination subnets. When a link failure occurs between the access layer device and the active HSRP distribution device, all mac address entries are flushed from the access layer device. The standby HSRP device "hold" timer begins counting down when no "hello" messages are received within 250 msec from the active HSRP device. After the dead timer expires, the standby HSRP takes over the role of active HSRP and begins forwarding traffic.

Downstream traffic:

In normal state, half of the total traffic is forwarded by the active HSRP device and the other half by the standby HSRP device. This was as a result of load-balancing that occurred at the remote distribution layer and the core layer which forwarded the traffic to the local distribution devices. When link failure occurs, the Spanning tree for the supported vlans goes down. This failure triggers the SVIs (Switched Virtual Interfaces) attached to the corresponding vlans to go down. The routes associated with the vlans go down as well. EIGRP is configured as IGP in the distribution devices. CEF entries are removed from the forwarding table for those VLANs. These routes in EIGRP table are lost. Since there is no equal cost path, EIGRP queries are sent to the rest of its neighbors to find the next available route. EIGRP receives updates for those queries from the standby HSRP device. It then installs the EIGRP routes in the routing table and updates the CEF entries in the forwarding table. Downstream traffic continues through this alternate path

Figure 4-8 indicates the convergence values for VPN and Global traffic (both upstream and downstream traffic) when the link between access to distribution is failed.

Figure 4-8 HSRP Active Link Failure Convergence



Active HSRP Link Restoration Analysis

Upstream traffic:

When the fiber link is restored on the Active HSRP device, the trunk link between the active HSRP device and the access layer device becomes active and activates all the VLANs that are configured on the trunk link. The Spanning tree becomes active and all the VLANs go into forwarding mode. The active

HSRP device becomes the primary root bridge. Since the preempted delay is configured to 180 sec, the Active HSRP device takes over its active role after 3 min (180 sec) and traffic starts forwarding through the Active HSRP device. No packet loss is observed in this case.

Downstream traffic:

In addition to link restoration (the Spanning-tree reconvergence), the connected subnets are installed back in the routing table. Once completed, traffic begins using this newly activated link. Again no packet loss is observed in this case.

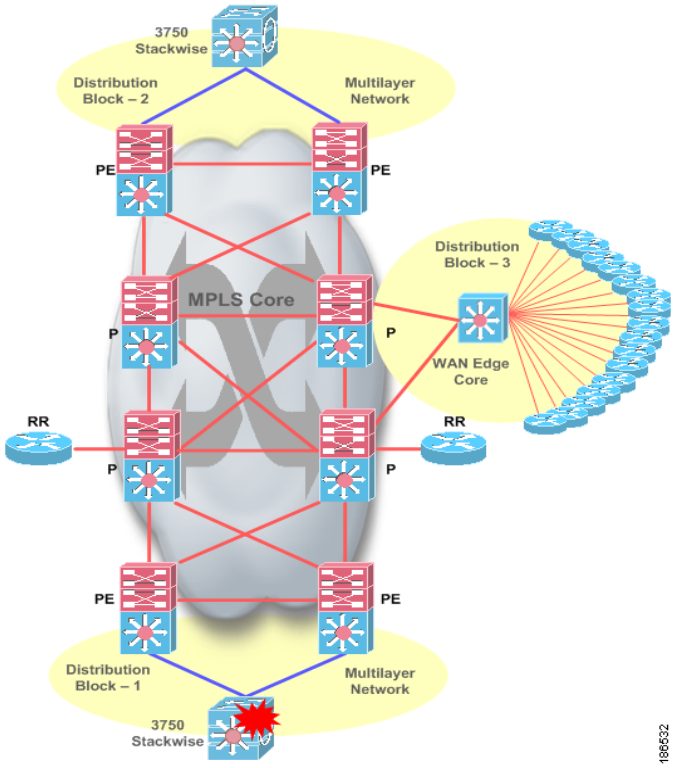
Note There was no packet loss observed in VPN and Global traffic during link restoration.

4.2.2 Access Layer Box Failure

In the Multilayer Campus test network, two Cat3750 devices act as one logical device and are configured as master and member. The master switch is responsible for exchanging control plane information to the attached PE device. The Access link connected to the test tool and the uplinks are connected to the member device.

Figure 4-9 illustrates Access Layer Box failure topology.

Figure 4-9 Access Layer Box Failure



Box failure: In this failure scenario, the master device was powered down. The member device takes the role of propagating the control plane information along with the data plane.

Note Since there was no change in the data path before and after the failure, there was no loss observed in the VPN and Global traffic.

4.3 Distribution to Core Layer Path Isolation

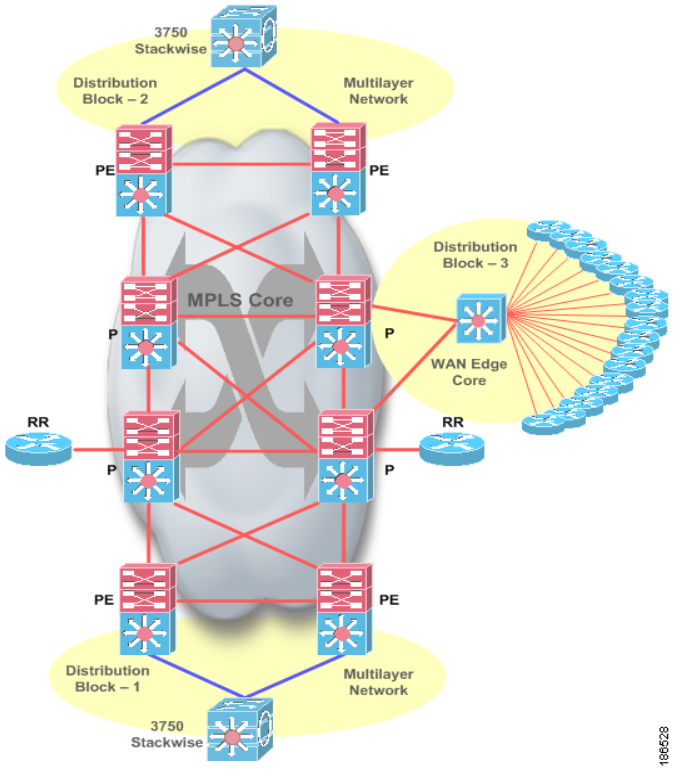
4.3.1 Building High Availability IP/MPLS Core Network

In large scale network designs, outages can be pre-planned or caused by unexpected failure events. Pre-planned network outages are generally a result of hardware and or software upgrades or new technology migration. For the most part, they do not impact end-users. Building a resilient network can reduce the impact to end-users during unexpected link, route-processor or device failures.

IP/MPLS Core Network designs leverages the same recommendations as defined in various Campus Multilayer and High Availability design guides. The network should be fully-meshed between edge and core devices as well as between core devices. Furthermore, a redundant supervisor should be deployed to achieve sub-second convergence during active supervisor failure using NSF/SSO technology.

Figure 4-10 illustrates a High Availability IP/MPLS Core network:

Figure 4-10 High Availability MPLS Core Network



Some attributes in the IGP protocol are automatically configured, like ECMP. Cisco recommends network summarization at the aggregation device like the Distribution Layer and WAN edge devices for global traffic. Network summarization helps avoid over use of system and network resources. Physical and logical ECMP paths and advanced software technique must be implemented to reduce end-to-end convergence during link failure or recovery events. The following components are used during solution characterization to achieve sub-second convergence.

Note When MPLS is used within the core of the network, the addresses that are used for the loopback interfaces of PE-routers must not be summarized anywhere within the network because this will cause a loss of connectivity across the backbone.

4.3.1.1 Equal Cost Multi Path

The load-balancing of Global and VPN traffic behaves differently from one another. For VPN load-balancing, a specific RD configuration together with iBGP multi-path capabilities are needed.

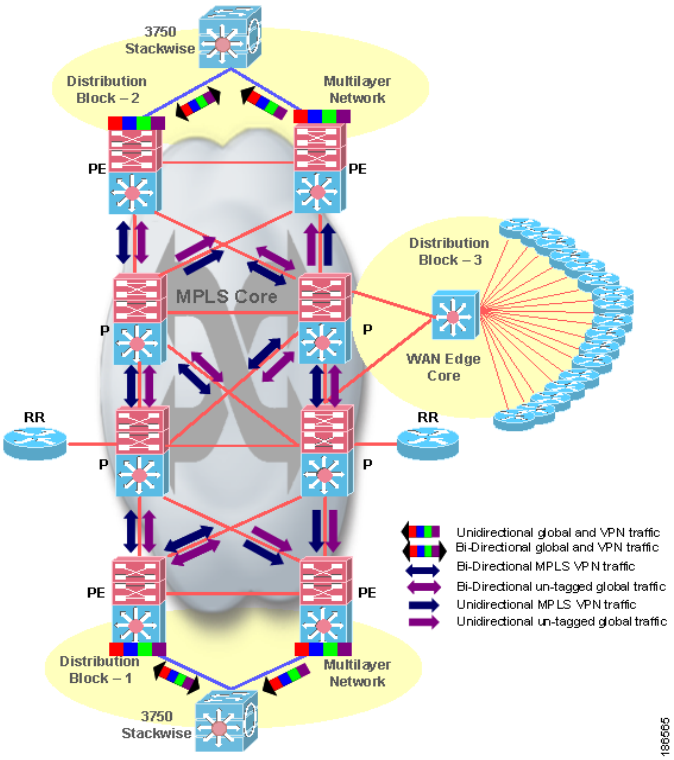
With the parallel fully-meshed uplinks in the Campus Core, by default, IGP installs up to six equal-cost paths in the routing table. Just like the IP routing table, the MPLS label forwarding table installs parallel path and performs load-balancing on a per outer label basis. When the Cisco Catalyst 6500 is deployed in a MPLS edge network, it performs the MPLS VPN PE role and the default load-balancing mechanism shares upstream traffic based on the source and destination IP address for up to three label stacks. Based on the IP/MPLS information from the ingress packet, traffic is passed to the load-balancing hashing logic to TCAM for further computation. Based on the hashing results, the MPLS VPN load-balancing rewrites the information's structure in hardware to perform hardware switching. No special configuration is required to build ECMP forwarding paths either for IP routing or the MPLS label.

However, it is recommended to configure the default CEF load balancing mechanism from the source IP and universal ID to the source and destination IP address on all the devices in the Core network using the following command in global configuration mode:

```
c6500(config)#mls ip cef load-sharing simple
```

During solution characterization, the third distribution block with a set of emulated WAN Edge devices interconnect to Core devices in the Campus for large scale IGP adjacencies and injecting networks from remote branches. All the advertised routes are considered to be in a Global network and no virtualization was extended to go beyond the Campus boundary. Solution characterization was validated and considered all scaling factors of the control and data plane that represent typical large scale Campus network deployments. Figure 4-11 depicts End-to-End load sharing of Bi-Directional traffic flow based on the Global and VPN routing table entries.

Figure 4-11 End-to-End Load-Balancing of Bi-Directional Traffic Flow



Recommendation: Deploy a fully-meshed IP/MPLS Campus network infrastructure between directly connected devices. Implement, fine-tune the control-plane and optimally use all links equally to load share the traffic.

4.3.1.2 MPLS LDP Session Protection

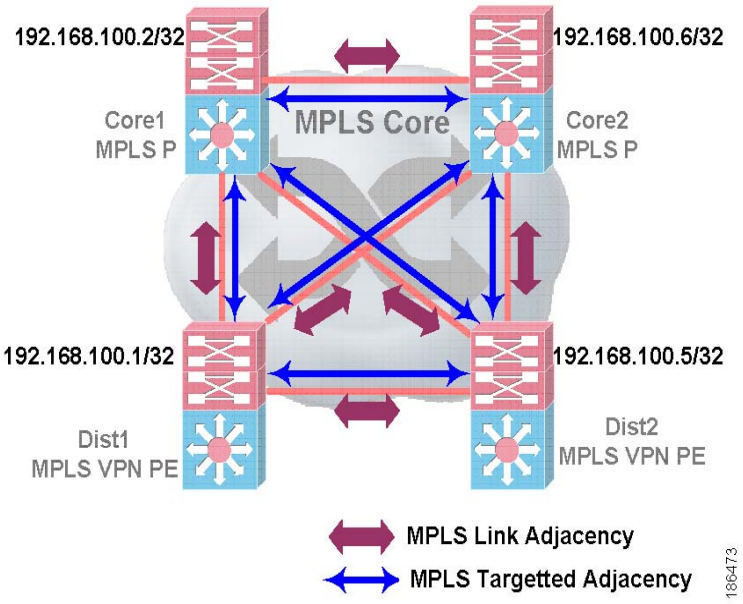
When link failures occur, the control-plane protocols reset all the forwarding entries discovered via IGP and MPLS. Since most IGP peering is done on a per-link basis, protecting the control-plane and routing entries may not be possible. However, MPLS is designed with advanced features to protect against such failures.

By default, when the MPLS protocol is enabled on a per-link basis. MPLS peers dynamically identify remote peers and establish sessions known as link-adjacency. MPLS can also establish a targeted or LDP session with peers either by using a static per neighbor LDP peer configuration or by enabling MPLS LDP session protection in the entire MPLS VPN network. At least one LDP peering remains up with the same peer over alternate paths and this allows label bindings to be retained for the same prefixes when the failed link recovers. The targeted LDP session remains up until the peer is reachable via an alternate path to the LDP source interface. This significantly reduces end-to-end network convergence as it eliminates the time it takes to re-learn the label and update the forwarding database.

In the *Network Virtualization - Path Isolation Design Guide*, the MPLS LDP Session Protection feature was not supported. After upgrading the IOS release to 12.2(33)SXH1 on the Cisco Catalyst 6500, the LDP session was automatically established by turning on the MPLS LDP Session Protection feature.

Figure 4-12 illustrates MPLS LDP Session Protection Benefits.

Figure 4-12 MPLS LDP Session Protection Benefits

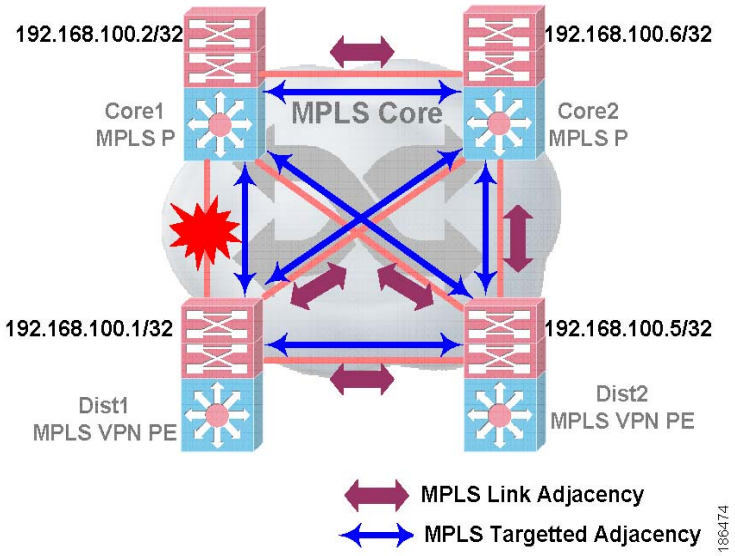


The following commands enable the MPLS and LDP session protection feature throughout the IP/MPLS Campus network:

```
C6500(config)#mpls label protocol ldp
C6500(config)#mpls ldp session protection
C6500(config)#interface TenGigabitEthernet 3/1
C6500(config-if)#mpls ip
C6500#show mpls ldp neighbor 192.168.100.2 detail
    Peer LDP Ident: 192.168.100.2:0; Local LDP Ident 192.168.100.1:0
        TCP connection: 192.168.100.2.19978 - 192.168.100.1.646
        State: Oper; Msgs sent/rcvd: 2971/2984; Downstream; Last TIB rev sent 31170
        Up time: 1d19h; UID: 4; Peer Id 0;
        LDP discovery sources:
          TenGigabitEthernet3/2; Src IP addr: 111.111.111.2
            holdtime: 15000 ms, hello interval: 5000 ms
          Targeted Hello 192.168.100.1 -> 192.168.100.2, active, passive;
            holdtime: infinite, hello interval: 10000 ms
        Addresses bound to peer LDP Ident:
     192.168.100.2   223.255.7.11   151.151.151.1   111.111.111.2
     111.111.114.2   131.131.135.2  131.131.131.1
Peer holdtime: 180000 ms; KA interval: 60000 ms; Peer state: estab
        Clients: Dir Adj Client
        LDP Session Protection enabled, state: Ready
            duration: 86400 seconds
```

As shown in the above output, once LDP Session Protection is configured, it can be verified using the show `mpls ldp neighbor <ip>` detail command. When at least one link adjacency to the peer is up, the LDP Session Protection is in the “ready” state. LDP adjacency is lost because the link between PE and P failed as shown in Figure 4-13. When this failure event occurs, an alternate LDP path remains reachable.

Figure 4-13 Links Lost Between and P



Recommendation: Configure and force the MPLS LDP protocol to source from the loopback interface and advertise it to IGP network.

Refer to following website for more details about the MPLS LDP Session Protection feature:

http://www.cisco.com/en/US/products/sw/iosswrel/ps1829/products_feature_guide09186a00802d95d9.html

4.3.1.3 Fine Tuning Protocol timers

Configuring IGP timers from the default values is another important tool for faster neighbor failure detection during link failure events. Since IGP functions on a per interface basis, it is required to adjust the heartbeat and adjacency "hold" timers on a per interface basis. Therefore, each core link must be configured with the optimal IGP timer value.

In the IP/MPLS Campus network, when a device is deployed using a single route-processor, aggressively adjusting the IGP timer may achieve the best convergence during link failure events. Cisco recommends that EIGRP "hello" timer be set to 2 and the "hold" timer be set 8 when a core Device is deployed using a dual route-processor with Non-Stop Forwarding (NSF) and Stateful Switchover (SSO) feature.

During the CVD solution validation, the EIGRP routing protocol was used as the IGP and the following configuration sample provides output from a large scale test network:

```
c6500(config)#interface Ten3/1
c6500(config-if)#ip hello-interval eigrp 1 2
c6500(config-if)#ip hold-time eigrp 1 8
```

Recommendation: Set the EIGRP "hello" timer to 2 seconds and the "hold" timer to 8 seconds in dual supervisor NSF/SSO deployment scenarios. Up to 50 EIGRP neighbor adjacency was tested during the dual supervisor deployment scenario, and no system resource issues were observed during network characterization.

When an ingress interface receives large amounts traffic, it is likely for a router to drop some of the traffic because of the default queue size on an interface. This traffic can be classified as initial IGP establishment traffic or even “hello” packets. Setting the interface queue size as shown in the configuration sample below will resolve the problem and stabilize the network:

```
c5600(config)#interface TenGigabitEthernet 3/2
c6500(config-if)#hold-queue 2000 in
c6500(config-if)#hold-queue 2000 out
```

Recommendation: Increase the inbound and outbound interface queue size to 2000 to allow the router to buffer traffic.

MPLS LDP protocol heartbeats are designed similar to IGP protocols. LDP “keep alive” and “hold” timers can be fine tuned to sub-second values for link-adjacency and directed LSP sessions. Unlike IGP, MPLS LDP “hello” and “hold” timers can be set for all or on a per-neighbor basis in the global configuration mode instead of on a per-interface basis. During solution characterization, it was observed that during core link failures, MPLS LDP sessions tear down within milliseconds after a link failure is detected. The sub-second MPLS LDP protocol timer faces similar issues as the IGP protocol timers in a dual supervisor environment using NSF/SSO. During Core failure characterization, MPLS LDP protocol timers were left in their default setting value because there was no compelling reason found to reduce the end-to-end convergence time.

Recommendation: Keep the MPLS LDP link and directed “hello” and “hold” timer to default to retain control-plane stability during link and RP switchover events.

4.3.2 Implementing Resiliency Control Plane IP/MPLS Network

Network device upgrade is a planned network outage condition. Almost all current generation Cisco devices are hot-swappable and planned outages result in minimal end-user and application impact. Supervisor modules in the Cisco Catalyst series are designed to achieve redundancy that can retain control-plane peering and perform non-stop end-to-end communication during software upgrade procedures. Furthermore, there is no observable impact during switchover due to an abnormal event, i.e. software failure on an active route-processor. Features described in the following section are designed for control-plane graceful restart.

4.3.2.1 Cisco NSF/SSO

Several Cisco IOS technologies are designed to function without control-plane interruption during stateful switchover These IOS technologies include: IGP, BGP, MPLS and LDP. During solution characterization, a Campus MPLS network was deployed with dual-supervisor with NSF/SSO. Cisco NSF/SSO technology functions independently of the Network Virtualization solution, but in a IP/MPLS Campus deployment, Cisco recommends implementing NSF/SSO for seamless global and virtualized network operation during route-processor switchover.

For more information about Configuring NSF with SSO MSFC Redundancy refer to the following guide:

http://www.cisco.com/en/US/docs/switches/lan/catalyst6500/catos/8.x/configuration/guide/nsf_sso.html

The following configuration sample provides guideline to implement SSO in Campus dual-supervisor deployments:

```
c6500(config)# redundancy
6506(config-red)# mode sso
c6500(config-red)# end
```

4.3.2.2 IGP NSF/SSO

EIGRP is fully NSF aware and capable. During stateful switchover, with the recommended “hello” and “hold” timer settings, EIGRP adjacency will gracefully synchronize the EIGRP topology table with large scale peers and continue to perform traffic switching during switchover. In a large scale solution test setup up to 50 EIGRP adjacencies were successfully tested. The following configuration is a sample output for enabling the NSF capability or the “aware” feature for EIGRP under the EIGRP process.

```
c6500(config)#router eigrp 1
c6500(config-router)#nsf
```

4.3.2.3 BGP NSF/SSO

Network virtualization leverages the MPLS VPN architecture to achieve Path Isolation. An MPLS VPN network uses MP-iBGP to exchange VPN information such as labels. SSO must be enabled on each BGP speaker. MPLS P devices do not participate in BGP AS. Distribution or PE devices should be peering internal IPv4 and VPNv4 BGP peering directly with the remote distribution device or with a dedicated Route-Reflector. Like IGP, each BGP speaker is required to turn on the SSO capability which will retain the BGP session and labels discovered from peers in the forwarding table during supervisor switchover. The following configuration is a sample output for turning on the graceful restart capability for the BGP protocol:

```
c6500(config)#router bgp 64000
c6500(config-router)#bgp graceful-restart
```

4.3.2.4 MPLS LDP NSF/SSO

The MPLS LDP protocol is the de-facto protocol for exchanging IGP or outer label information throughout the IP/MPLS Campus network. The LDP protocol uses the TCP protocol to communicate and establish peering sessions between LDP devices. In dual-supervisor NSF/SSO environments, a TCP session is established on the active supervisor and check pointing is performed to put the supervisor on standby once the LDP graceful restart capability is enabled and the global configuration mode is disabled

When the active supervisor switchover occurs, all the label bindings are retained on the standby supervisor so that MPLS traffic continues to perform label switch traffic. Unlike the IGP and BGP protocol, the MPLS LDP session will be forced by TCP to reset during supervisor switchover. During active stateful switchover events, all the MPLS labels discovered from the impacted neighbors are marked as stale for a certain period of time and continue to label switch traffic until the supervisor recovers.

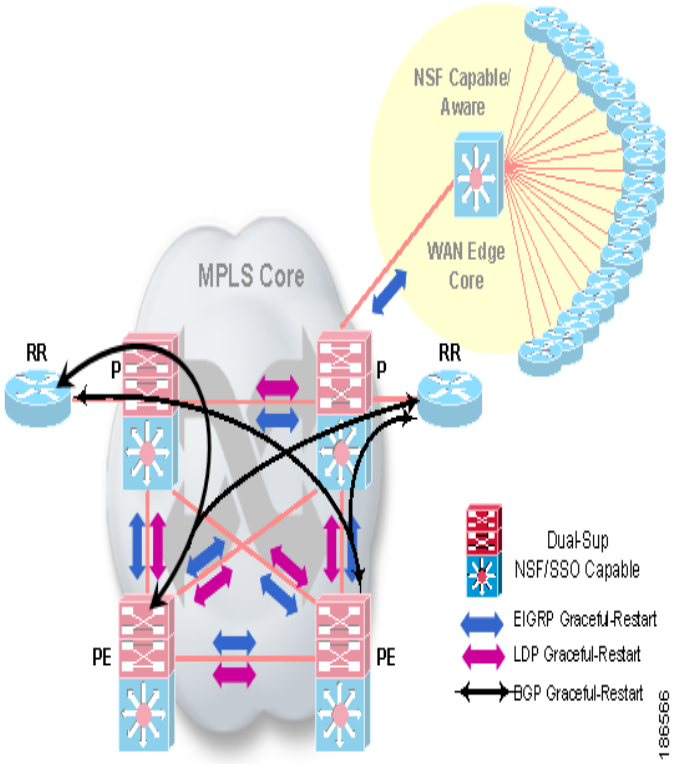
The MPLS LDP Graceful-Restart feature is supported on the Cisco Catalyst 6500 platform running the 12.2(33)SXH1 release. All solution characterization was performed based on the MPLS LDP Graceful-Restart feature. The following configuration provides sample output to enable the Graceful-Restart capability of the MPLS LDP protocol:

```
c6500(config)#mpls ldp graceful-restart
```

Recommendation: If the network deployment consists of a dual supervisor, Cisco recommends that NSF/SSO be turned on in each device to achieve non-stop traffic forwarding. Cisco Catalyst 6500 platform should be upgraded with the 12.2(33) SXH1 IOS release or later in order to enable the MPLS LDP Graceful-Restart feature.

Figure 4-14 shows a single Distribution block which depicts the control-plane resiliency.

Figure 4-14 Enabling Control Plane Resiliency Features



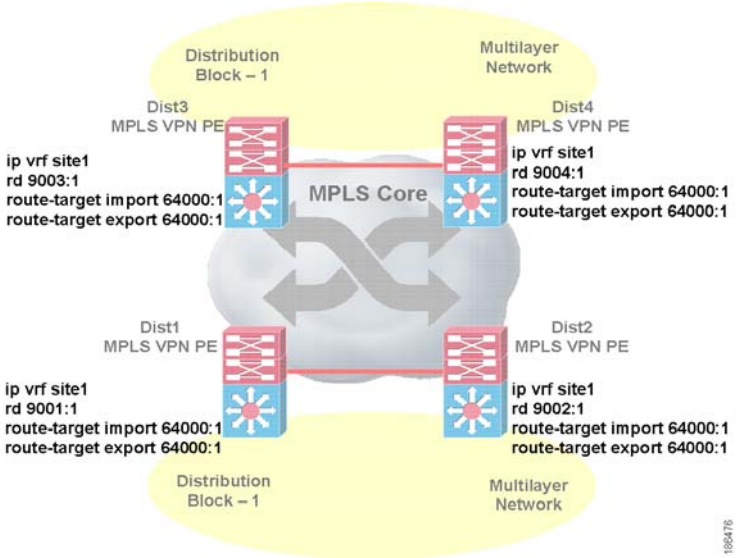
4.3.3 Additional MPLS VPN Features

Building parallel paths and redundant devices does not resolve resiliency in MPLS VPN networks. To build a redundant MPLS VPN network, the following MPLS VPN features should be implemented and fine tuned as described in the following subsections.

4.3.3.1 Load-Balancing VPN Route

Unlike IPv4, routes in an MPLS VPN are identified using a unique RT value and VPN prefix in the BGP table. Each MPLS VPN PE device in the same distribution block should be designed to announce the same network using the same RT values. With an ECMP path in IGP for the PE address, VPNv4 prefixes are considered multi-path from the perspective of the VPNv4 Route-Reflector or Remote PE. Each PE advertises a unique VPNv4 label for the same network pointing to a different PE device. Refer to the *Network Virtualization - Path Isolation Design Guide* for more details. Figure 4-15 provides sample configuration guideline to implement the RD and RT addressing scheme in a Campus MPLS VPN edge network.

Figure 4-15 Implementing RD and RT Addressing in Campus MPLS VPN Edge Networks



Once the network is deployed based on recommended configuration guideline, end-to-end MPLS VPN traffic is load-balanced as shown in Figure 4-11.

Note If Cisco Catalyst 6500 platforms are deployed in the IP/MPLS Campus network as the PE role, the maximum number of VRF supported is 1024 on the PFC3 based supervisor module. The VPN CAM size can handle up to 511 usable VRF. Scaling beyond the maximum CAM size may impact switching performance caused by per-flow packet recirculation on PFC and DFC based modules. Even during packet recirculation, non-circulated traffic will flow without any performance impact.

4.3.3.2 VPNv4 Route-Reflector

When building a large scale fully-meshed MPLS VPN network, BGP peering may impact device performance and increase complexity in managing and troubleshooting the BGP network. Implementing VPNv4 route-reflectors reduces fully-meshed iBGP peering in a Campus core. Each MPLS VPN PE device establishes VPNv4 BGP peering with dedicated Route Reflectors and each PE receives reflected routes as well as un-altered advertised VPNv4 labels and next-hop information from the route reflector.

Since each MPLS VPN PE router peers to the Route Reflector, it may become the single point of failure. A redundant VPNv4 Route Reflector should be deployed in the Campus network and each MPLS VPN device should be configured to peer with both Router Reflectors. For more details about VPNv4 BGP peering with Route Reflector please refer to Figure 4-14.

The following configuration output provides guideline on implementing a BGP VPNv4 Route-Reflector:

```
router bgp 64000
 no synchronization
 bgp router-id 192.168.100.111
 bgp cluster-id 64000
 bgp log-neighbor-changes
...
neighbor rr-clients peer-group
 neighbor rr-clients remote-as 64000
 neighbor rr-clients update-source Loopback0
```

```
 neighbor rr-clients send-community extended
 neighbor 192.168.100.1 peer-group rr-clients
 neighbor 192.168.100.4 peer-group rr-clients
 neighbor 192.168.100.5 peer-group rr-clients
 neighbor 192.168.100.8 peer-group rr-clients
 neighbor 192.168.100.112 remote-as 64000
 neighbor 192.168.100.112 update-source Loopback0
 no auto-summary
 !
 address-family vpnv4
 neighbor rr-clients send-community both
 neighbor rr-clients route-reflector-client
 neighbor 192.168.100.1 activate
 neighbor 192.168.100.4 activate
 neighbor 192.168.100.5 activate
 neighbor 192.168.100.8 activate
 neighbor 192.168.100.112 activate
 neighbor 192.168.100.112 send-community extended
exit-address-family
```

Recommendation: Deploy dedicated redundant VPNv4 Route Reflectors in the network to peer with MPLS VPN PE devices. Route Reflectors should not be deployed in any of the data paths.

4.3.3.3 MP-iBGP Multi-path

In an MPLS VPN network with multiple iBGP paths installed in a routing table, an Route Reflector advertises only a single path. If the VPNv4 prefixes are advertised with the same route distinguisher (RD), Route Reflector propagates only the best path to the iBGP neighbors. MP-iBGP load balancing can function with Route Reflectors only when VPNv4 prefixes are tagged with different RD values. Therefore, configure unique RD values for each PE device but configure common RT values on all PE devices that belong to a single distribution block

By default, MP-iBGP installs a single path in the forwarding table even if the same prefix from a different next-hop passes through the BGP best path selection process. If all the best path selection criteria match for the same prefix from a different next-hope, the highest BGP router-id installs a route in the VPN routing table. The ECMP in the MP-iBGP path is selected when all the BGP attributes like weight, local-preference, AS-path, origin code, MED and IGP metric are equal. To enable MP-iBGP multipath for each VPN site, use the following configuration to import and install all VPNv4 prefix paths in the forwarding table:

```
c6500(config)#router bgp 64000
....
c6500(config-router)# address-family ipv4 vrf  vpn_Blue
c6500(config-router-af)#maximum-paths ibgp 2 import 4
```

4.3.3.4 BGP Next-Hop Tracking

The BGP Support for Next-Hop Address Tracking feature is enabled by default when a supporting Cisco IOS software image is installed. Border Gateway Protocol (BGP) next-hop address tracking is event driven. BGP prefixes are automatically tracked as peering sessions are established. Next-hop changes are rapidly reported to the BGP routing process as they are updated in the Routing Information Base (RIB). This optimization improves overall BGP convergence by reducing the response time to next-hop changes for routes installed in the RIB. When a best-path calculation is run between BGP scanner cycles, only next-hop changes are tracked and processed.

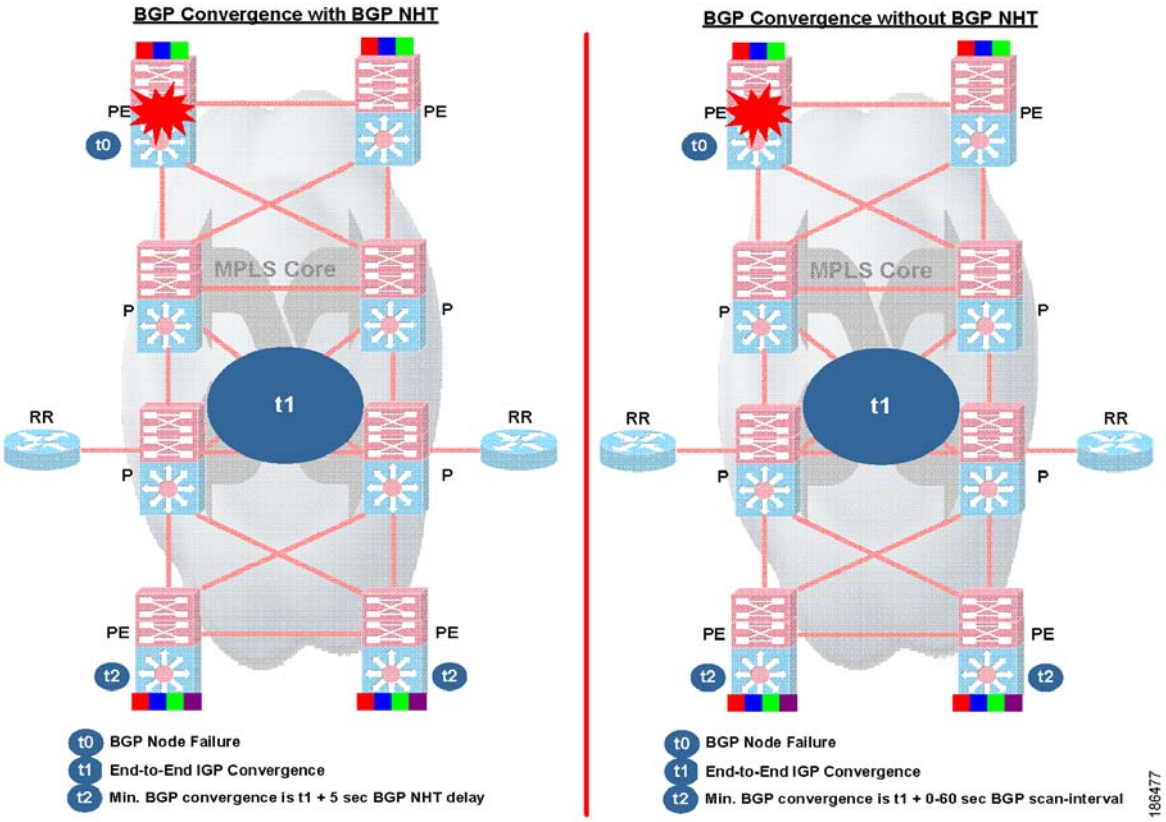
Beginning with the 12.2(33)SXH1 software release on the Cisco Catalyst 6500 platforms, the BGP Next-Hop tracking feature is supported. BGP Next-Hop Tracking (NHT) was enabled on each MPLS PN PE and VPNv4 Route Reflector devices during solution validation.

The BGP Next-Hop tracking feature is event driven instead of on a scan-time basis. The BGP table is updated instead of waiting for the scan-interval timer to improve end-to-end network convergence when remote PE link or node failures occur in the network. The BGP Next-Hop tracking feature operates on per address-family basis. The BGP scanner continues to function independently on every address-family even if BGP Next-Hop Tracking is enabled.

Whenever a change in the next-hop is detected in RIB, the BGP Next Hop Hopping timer starts a scheduler to update the BGP table. The default BGP Next Hop Hopping scheduler or delay value is 5 seconds and can be modified depending on how fast the IGP network is and the expected end-to-end convergence. The recommended value for BGP Next Hop Track delay timer value is “0” seconds for fast convergence.

Figure 4-16 illustrated the end-to-end convergence with and without using BGP Next-Hop tracking feature in the Campus network:

Figure 4-16 Convergence with and without BGP Next-Hop Tracking Enabled



Refer to following web site for more information about BGP Next-Hop Tracking feature:

http://www.cisco.com/en/US/products/ps6566/products_feature_guide09186a0080796ead.html

4.3.4 Optimizing IP/MPLS Network

The IP/MPLS Campus network can be further optimized to improve end-to-end convergence. The features described in the following section can be implemented in the Campus core and edge devices to address specific network optimization needs:

4.3.4.1 Tuning TCP Protocol

Protocols like MPLS LDP, and BGP leverage all the benefits of the TCP protocol to establish adjacency prior to exchanging network information. When scalability grows, it becomes necessary to configure TCP protocol attributes throughout the network to improve network device stability, and network convergence.

During peak hours, the output interfaces on each Campus core or edge device may queue high and low priority traffic. This may include MPLS LDP “keepalive” packets. The MPLS LDP “keepalive” packets are TCP based and sessions may timeout due to delays in switching the TCP heartbeat in output interface queue. MPLS LDP adjacency resets due to loss of the LDP “hello” message. This may destabilize all isolated path traffic. Each Campus core device should be set to prioritize LDP packets during peak times. Implement LDP “pak-priority” on each Campus core as show below:

```
c6500(config)#mpls ldp tcp pak-priority
```

Handling large scale TCP traffic such as MPLS LDP or BGP on a router can degrade performance that may result in network destabilization. The TCP window-size should be fine tuned. Fine tuning the TCP window-size may help in handling an increased number of TCP bytes and reduce delays in processing. The default TCP window-size in Cisco IOS releases must be fine tuned from 4198 bytes to 65535 bytes. Fine tuning the TCP window-size will improve the BGP input and output process that result into faster end-to-end convergence.

The following configuration sample provides output of fine tuning TCP window size in global configuration mode:

```
c6500(config)#ip tcp window-size 65535
```

Another TCP attribute like the “Path MTU Discovery” method can be implemented to maximize the use of available core bandwidth in the network between the endpoints of a TCP connection. Predetermining the MTU size can help optimization during the MPLS LDP and BGP adjacency setup.

Enable TCP path the mtu discovery setting on all Campus core devices as defined below:

```
c6500(config)#ip tcp path-mtu-discovery
```

4.3.4.2 Fine Tuning MPLS Edge

When an MPLS VPN PE device forwards untagged IP traffic over the MPLS core, by default it will reduce the TTL value as is done in an IP based network. Default settings can be changed on MPLS VPN PE devices to forward traffic received from the Multilayer Campus network without decrementing the TTL value.

The following configuration must be implemented on each distribution or MPLS VPN PE device to disable TTL propagation:

```
c6500(config)#no mpls ip propagate-ttl
```

As recommended in the *Network Virtualization - Path Isolation Design Guide*, crucial applications like unified communications may not require Path Isolation and should not be impacted in any way when overlaying virtualization. All applications that are required to function based on the global forwarding table will continue to operate reliably by disabling MPLS label advertising for the entire global network on each Campus device. MPLS label advertising should only match the MPLS LDP source loopback interface.

The following configuration illustrated disabling MPLS label advertising on each Campus core device and will only allow the loopback interface:

```
c6500(config)#no mpls ldp advertise-labels

c6500(config)# mpls ldp advertise-labels for loopbacks

c6500(config)# ip access-list standard loopbacks
 c6500(config-std-nacl)#permit 192.168.100.8
 c6500(config-std-nacl)#permit 192.168.100.4
 c6500(config-std-nacl)#permit 192.168.100.5
 c6500(config-std-nacl)#permit 192.168.100.6
 c6500(config-std-nacl)#permit 192.168.100.7
 ....
```

4.3.4.3 Fine Tuning Core Interface

If a link fails, by default there is a two-second timer that must expire before an interface and the associated routes are declared as being down. If a link goes down and comes back up before the carrier delay timer expires, the down state is effectively filtered, and the rest of the software on the switch is not aware that a link-down event occurred. You can configure the carrier-delay seconds command in interface configuration mode to extend the timer up to 60 seconds.

Recommendation: Configure carrier-delay timer on the interface to a value of zero (0) to ensure no additional delay in the notification that a link is down. It is recommended as a best practice to hard code the carrier-delay value on critical interfaces with a value of 0 msec to ensure the desired behavior:

```
C6500(config)#interface tenGigabitEthernet 3/2
C6500(config-if)#carrier-delay msec 0
```

The IP Event Dampening feature introduces a configurable exponential decay mechanism to suppress the effects of excessive interface flapping events on routing protocols and routing tables in the network. This feature allows the network operator to configure a router to automatically identify and selectively dampen a local interface that is flapping.

When the link failures occur, the notification to the upper layer control-plane should be instantly performed for better convergence to the alternate path. However, when the link comes back up, it may not be fully stable or reliable to start forwarding traffic. This condition may further de-stabilize the network and may require “IP dampening” to add a delay between the actual link up detection until the upper layer is notified. IP dampening should be implemented on each MPLS core interface to control link stability.

Configure IP dampening as described below on each Campus core device and on each IP/MPLS core interface:

```
C6500(config)#interface tenGigabitEthernet 3/2
C6500(config-if)#dampening
```

The IP routing protocol purge interface command enables routing protocols that can respond to interface failures to delete dependent routes from the RIB when a link on a router goes down and the interface is removed from the routing table.

In the IP/MPLS Campus networks, when link failures occur, the traffic moves to an alternate interface and this can impact end-to-end convergence time.

Convergence time in the Campus core is optimized by using the “ip routing protocol purge interface” command as follows:

```
C6500(config)# ip routing protocol purge interface
```

4.4 Distribution to Core Layer Failure Characterization

This section presents solution characterization data that was generated from emulating real-world failure scenarios occurring in the L3 distribution and Core layer. All Network Virtualization solution characterization reports were generated based on the traffic profile described in section 3.2 Traffic Profile along with the scalability defined in 3.4.2 Scalability.

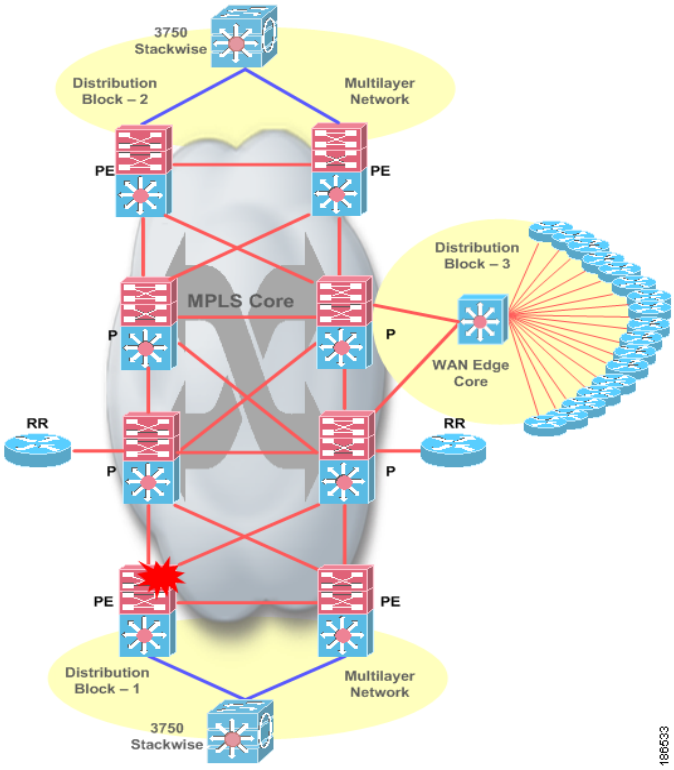
4.4.1 PE Supervisor Failure (NSF/SSO Switchover)

In Multilayer Campus test network, the L3 distribution router acts as an MPLS PE router. The router has dual supervisor running in NSF/SSO mode.

Supervisor Failure: Under normal conditions, the router running in NSF/SSO mode will have one active and one standby supervisor. Issuing a redundant force-switchover command on the active supervisor forces the standby to take on its active role. Since the standby supervisor had its IGP, BGP and MPLS states synced with the active supervisor, there is no change in the traffic path for both BGP and MPLS.

The following figure illustrates the failure introduced in the topology.

Figure 4-17 MPLS PE Supervisor Failure (NSF/SSO switchover)



PE Supervisor Failure (NSF/SSO switchover) Analysis

Upstream traffic:

In the validated IOS image 12.2(33)SXH1, local distribution switches have dual Supervisors and are configured to use new feature such as NSF or GR for EIGRP, MP-iBGP and MPLS. IOS 12.2(33)SXH1 also supports SSO for HSRP. Standby supervisor will be “hot” sync. This means that all the control plane information is synced to the standby supervisor. During Supervisor Switchover (a) HSRP will not lose its adjacency (b) EIGRP neighbor relationship will not be lost (c) MP-iBGP will continue to hold its routing table and (d) MPLS/LDP sessions are not torn down. There is no change in the data path when the switchover occurs, however, a brief amount of traffic loss occurs for the following reasons:

1. The type of Chassis used in the testbed.
2. The type of Linecards used.
3. The firmware version in the Linecards.
4. The ingress and egress port connections for the test traffic.

Note Please refer to Table 3-1. for further information about hardware and software used.

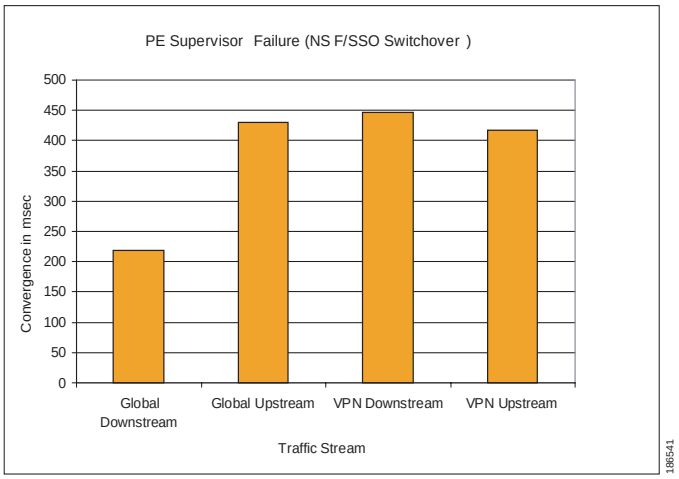
Downstream traffic:

For the downstream traffic, the effects are exactly the same as described in the previous Upstream traffic section.

The graph below indicates the convergence values for VPN and Global traffic (upstream and downstream) when the Supervisor switchover occurs in the PE node.

The following graph depict convergence values for VPN and global traffic.

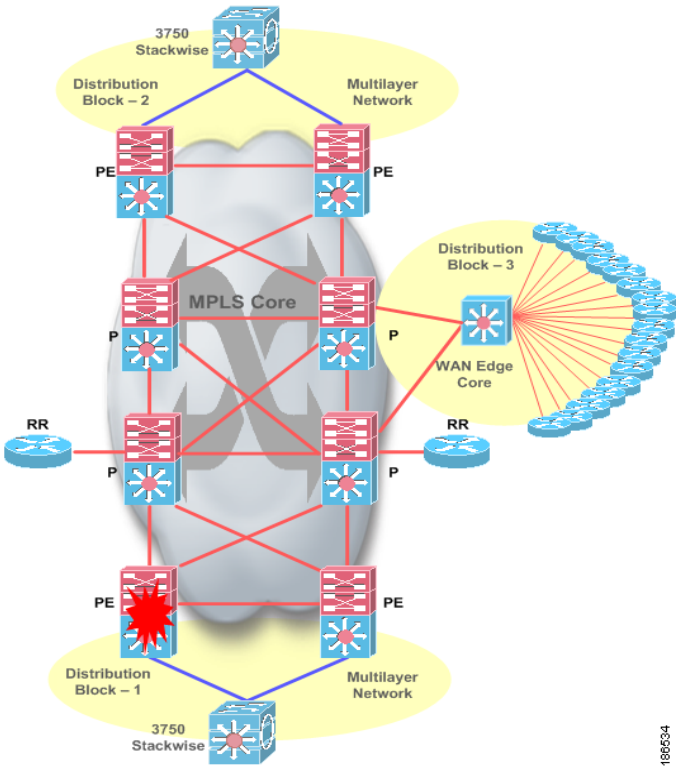
Figure 4-18 Convergence Values for VPN and Global Traffic (PE Supervisor Switchover)



4.4.2 PE Node Failure

The Multilayer Campus test network is deployed with HSRP as the gateway redundancy protocol. The Distribution router (PE1) is an L2/L3 router maintaining L2 states towards the Access Layer device and acts as an L3 device by maintaining IGP/BGP and MPLS neighbors. All uplink traffic is forwarded to an HSRP Active device (PE1 router in this case). Both the HSRP active and standby devices maintain per-VLAN and route entries to load-balance downstream traffic arriving from the IP/MPLS core network. Figure 4-19 illustrates MPLS PE (Distribution) Node Failure topology.

Figure 4-19 MPLS PE (Distribution) Node Failure



PE Node Failure Analysis

Upstream traffic:

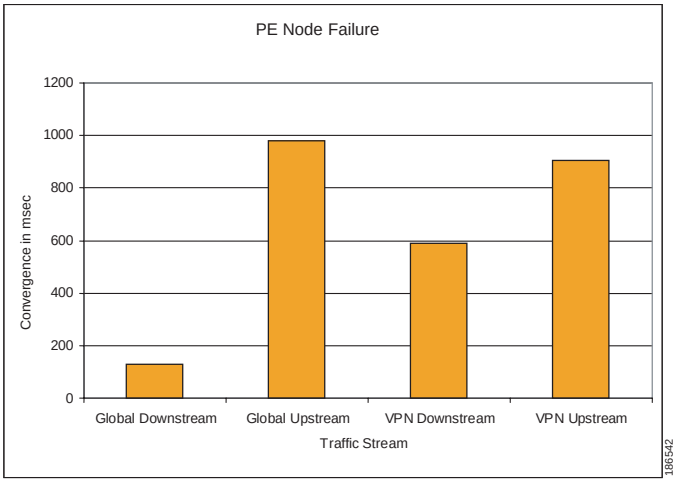
The failure mechanism for the PE Node for upstream traffic is very similar to the failure mechanism described in 4.2.1 HSRP Link Failure, page 4-11.

Downstream traffic:

The failure mechanism for the PE Node for downstream traffic is affected by the equal cost multipath mechanism (ECMP) in the core switches. The core switches have dual paths to the local distribution switches. When the PE1 fails, the core switch (C1 in this case) installed in the routing table, needs to update its routing table (for the subnets that were forwarded to the PE1) to point to PE2 as its next-hop.

Figure 4-20 describes the convergence values for VPN and Global (Upstream and Downstream traffic) when the distribution router (PE1) fails.

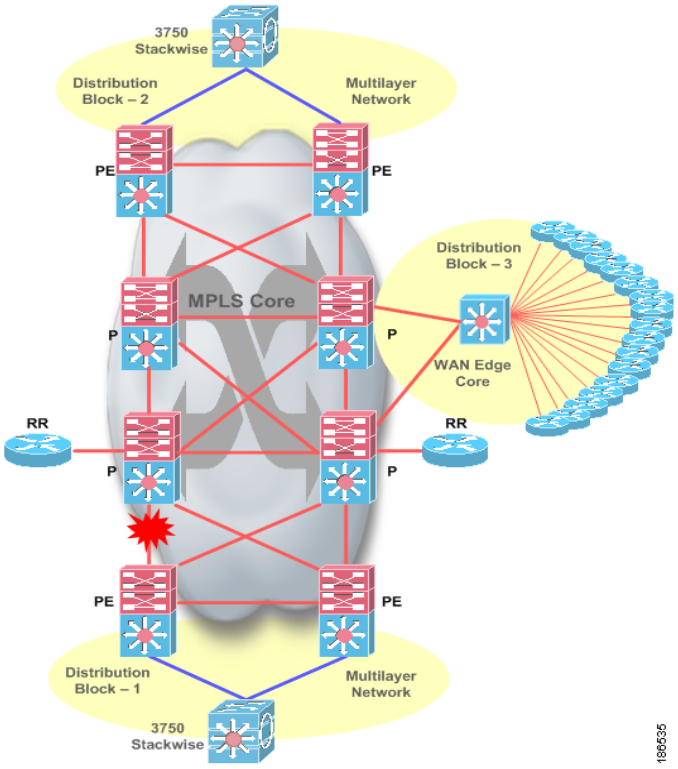
Figure 4-20 VPN and Global traffic Convergence Values (PE1 Node Failure)



4.4.3 Distribution to Core Link Failure

The following section highlights Distribution to Core link failure as illustrated in Figure 4-21.

Figure 4-21 Distribution to Core Link Failure



Distribution - Core link Failure Analysis

Upstream traffic:

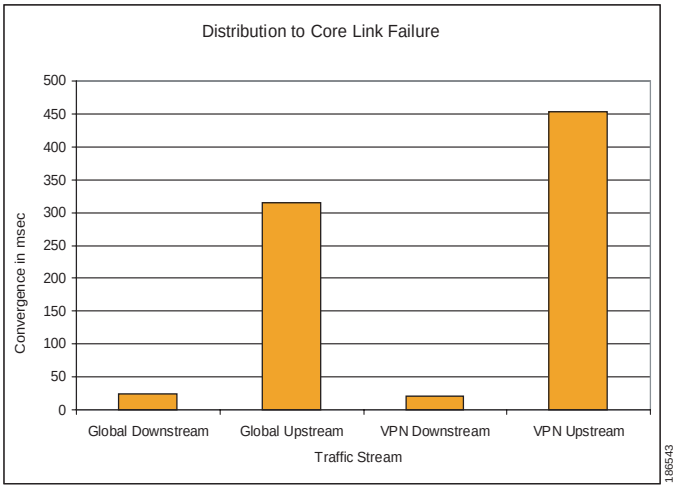
In the Distribution to core failure, the traffic that was going through the link that was failed, is re-routed to an alternate path. Since there are already two paths in the routing table, there is no need for IGP reconvergence. Once the route information is updated, then the hardware will update the CEF entries in the forwarding table and begin forwarding traffic.

Downstream Traffic:

The effect of downstream traffic is exactly the same as upstream traffic. Since the entire network is symmetrical on both sides only the direction of the traffic is reversed. During this failure event, only ECMP is involved for the traffic to take a new path.

Figure 4-22 describes the convergence values for VPN and Global traffic (Upstream and Downstream traffic) when the link between Distribution and Core routers fail.

Figure 4-22 VPN and Global traffic Convergence Values (Distribution - Core Link Failure)



Distribution to Core Link Restoration Analysis

Upstream Traffic:

Link restoration is dependent on the routing protocol to establish adjacency and install new routes. Once the routes are installed, the neighbor switch installs the second equal cost entry in its routing and forwarding table. This does not invalidate the entries for the traffic that was already sent. Once the routing and forwarding table are updated, traffic continues to forward on both available links based on the load balancing mechanism.

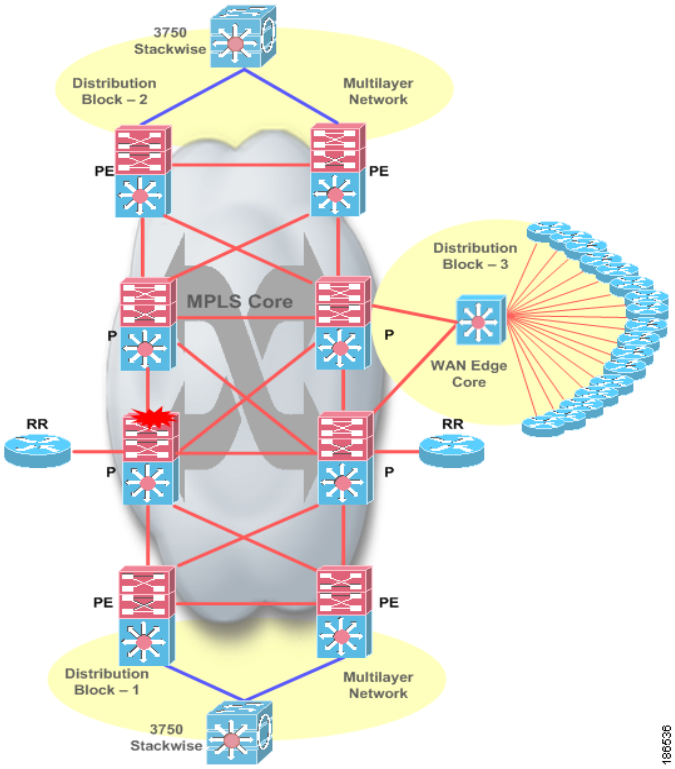
Downstream Traffic:

The effect of downstream traffic is the same as described in the section above. There was no loss observed during the distribution to core link restoration for downstream traffic.

4.4.4 Core Supervisor Failure (NSF/SSO Switchover)

Under normal conditions, the router running in NSF/SSO mode will have one active and one standby supervisor. Issuing a redundant force-switchover command on the active supervisor forces the standby to take on its active role. Since the standby supervisor IGP and MPLS states are synced with the active supervisor, there is no change in the traffic path.

Figure 4-23 Core Supervisor Failure (NSF/SSO Switchover)



Core Supervisor Switchover Analysis

Upstream traffic:

Upstream traffic:

In the validated IOS image 12.2(33)SXH1, local core switches have dual Supervisors and are configured to use new feature such as NSF or GR for EIGRP, MPLS. Standby supervisor will be “hot” sync. This means that all the control plane information is synced to the standby supervisor. During Supervisor Switchover, EIGRP neighbor relationship will not be lost and MPLS/LDP sessions are not torn down. There is no change in the data path when the switchover occurs, however, a brief amount of traffic loss occurs for the following reasons:

1. The type of Chassis that is used in the testbed.
2. The type of Linecards that are used.
3. The firmware version in those Linecards.

4. The ingress and egress port connections for the test traffic.

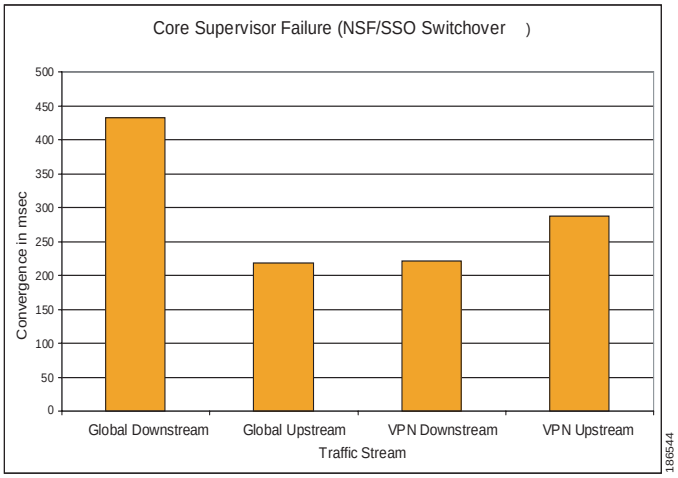
Note Please refer to Table 3-1. for further information about hardware and software used.

Downstream traffic:

For the downstream traffic, the effect is exactly the same as described in the previous section titled - Upstream traffic.

Figure 4-24 illustrates VPN and Global traffic Convergence Values (Core Supervisor Failure)

Figure 4-24 VPN and Global traffic Convergence Values (Core Supervisor Switchover)

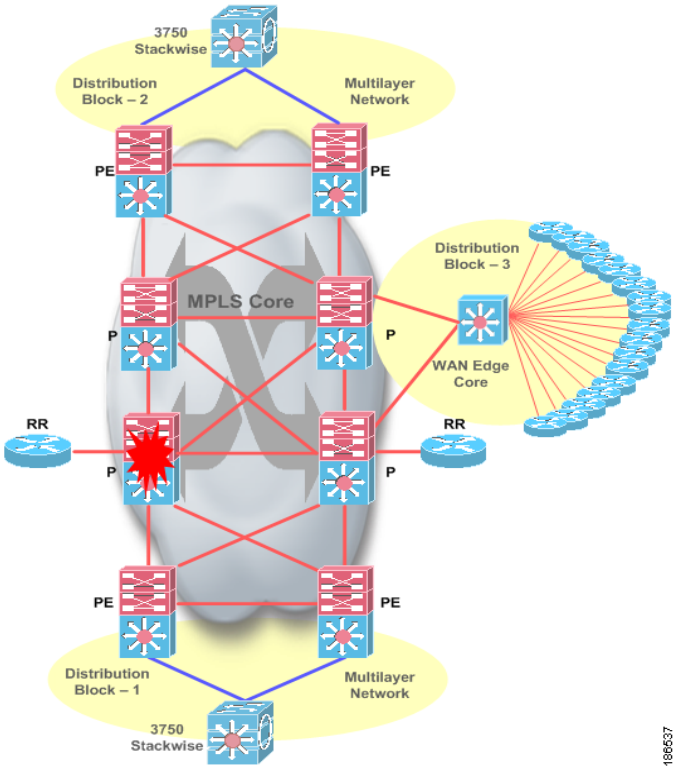


4.4.5 Core Node Failure

In the Multilayer Campus network, the core layer is primarily responsible for switching traffic.

Figure 4-25 illustrates the failure introduced in the topology.

Figure 4-25 Core Node Failure



Core Node Failure Analysis

Upstream Traffic:

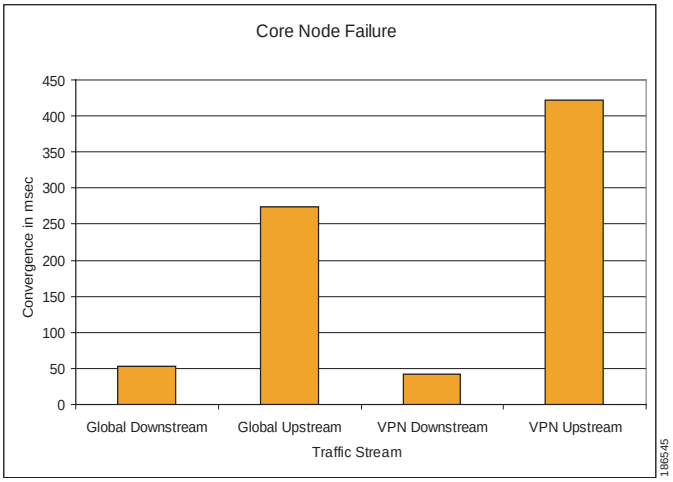
The campus core network is fully meshed. All core routers have redundant equal cost links to all the other core routers. The core device failure does not depend on the routing protocol convergence to redirect the traffic to a new path. The distribution switch (PE1) has two routes and two associated hardware CEF entries in its forwarding table. When the core device goes down (C1), the PE1 device detects the loss of link to the core device. It then re-routes all the traffic to the other available path. The time to reroute the traffic is dependent on the time required to detect the physical link failure and to update the software and corresponding hardware forwarding entries.

Downstream Traffic:

Downstream traffic is affected in the same way as described in the upstream traffic mentioned above.

Figure 4-26 describes the convergence values for VPN and Global traffic (Upstream and Downstream traffic) when the core router (C1) fails.

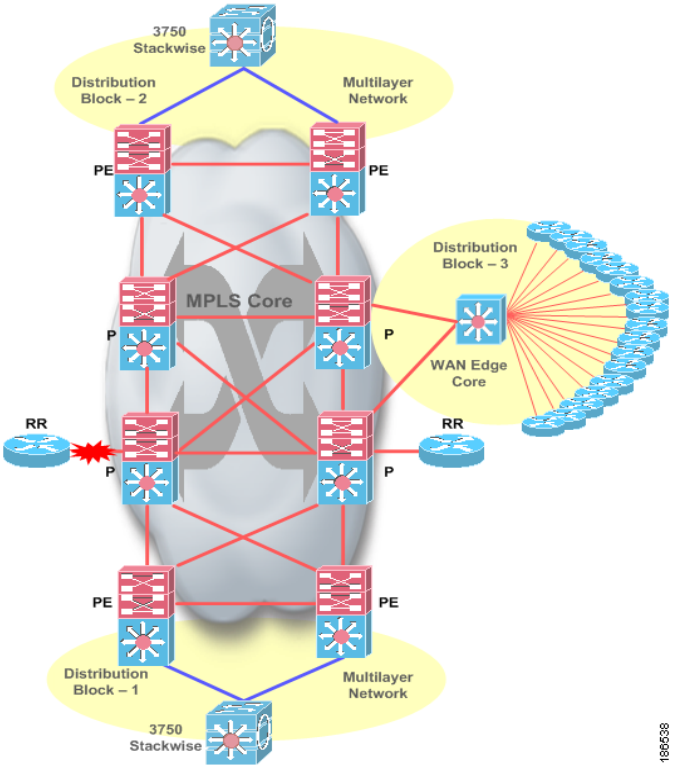
Figure 4-26 VPN and Global traffic Convergence Values (Core Node Failure)



4.4.6 Core to Route-Reflector Link Failure

In the Multilayer Campus network, a failure is introduced in the link between the Core router and Route Reflector. The Route-Reflector maintains fully-meshed connectivity. In this network, PE1, PE2, PE3 and PE4 are Route Reflector clients and RR1 and RR2 are redundant route reflectors in the same BGP cluster.

Figure 4-27 MPLS VPN Core to Route-Reflector Link Failure



Core to Route Reflector Link Failure Analysis

Upstream traffic:

The distribution switches in the campus network are not iBGP fully meshed. Route-reflectors are used to establish and reflect the BGP route information to other distribution switches. Two route-reflectors are used to provide redundancy in case of a RR (Route Reflector) node failure. Route Reflectors only propagate BGP updates within the distribution switches and are not used for forwarding traffic. As a result, a link failure introduced between the Core devices and the Route Reflector device will not affect data traffic.

Downstream traffic:

Downstream traffic behavior is exactly the same as the upstream traffic behavior. Since the Route Reflector (Route-reflector) are not in the data path, traffic is not affected by the link failure introduced between the core device and the Route Reflector device.

Note Since there was no change in the data path before and after the failure, there was no loss observed in the VPN and Global traffic.

Core to Route Reflector Link Restoration Analysis

Upstream traffic:

The distribution switches in the campus network are not iBGP fully meshed. Route-reflectors are used to establish and reflect the BGP route information to other distribution switches. Two route-reflectors are used to provide redundancy in case of a Route Reflector (Route-Reflector) node failure. Route Reflectors only propagate BGP updates within the distribution switches and are not used for forwarding traffic. As a result, a link restoration introduced between the Core devices and the Route Reflector device will not affect data traffic.

Downstream traffic:

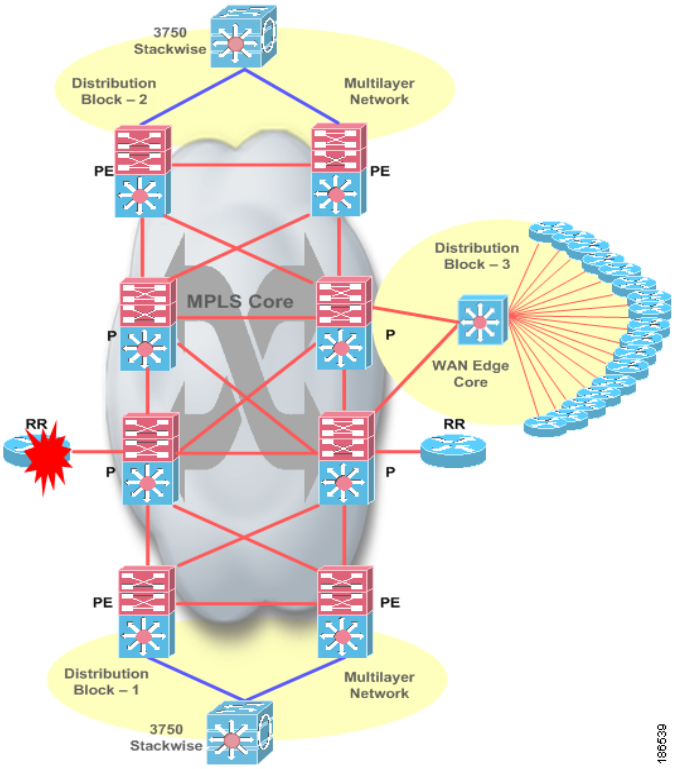
Downstream traffic behavior is exactly the same as the upstream traffic behavior. Since the Route Reflector (Route-reflector) are not in the data path, traffic is not affected by the link restoration introduced between the core device and the Route Reflector device.

Note Since there was no change in the data path before and after the failure, there was no loss observed in the VPN and Global traffic.

4.4.7 Route-Reflector Node Failure

In the Multilayer Campus network, a Route Reflector node failure is introduced. The Route-Reflector maintains fully-meshed connectivity. In this network, PE1, PE2, PE3 and PE4 are Route Reflector clients and RR1 and RR2 are redundant route reflectors in the same BGP cluster.

Figure 4-28 MPLS VPN Route-Reflector Node Failure



Route Reflector Node Failure Analysis

Upstream traffic:

The distribution switches in the campus network are not iBGP fully meshed. Route-reflectors are used to establish and reflect the BGP route information to other distribution switches. Two route-reflectors are used to provide redundancy in case of a Route Reflector node failure. Route Reflectors only propagate BGP updates within the distribution switches and are not used for forwarding traffic. As a result node failure of the Route Reflector device, there is no loss in the data traffic

Downstream traffic:

For the downstream traffic, the effect is exactly the same as described in the previous section titled upstream traffic. Since there is a secondary Route Reflector is available to reflect routes, traffic is not affected by the node failure introduced in primary Route Reflector.

Note Since there was no change in the data path before and after the failure, there was no loss observed in the VPN and Global traffic.

CHAPTER 5

Related Documents and Links

Cisco CVD program:

http://cisco.com/en/US/partner/netsol/ns741/networking_solutions_program_home.html

Network Virtualization-Path Isolation Design Guide

http://www.cisco.com/application/pdf/en/us/guest/netsol/ns171/c649/ccmigration_09186a0080851cc6.pdf

High Availability Campus Design Guide

http://www.cisco.com/application/pdf/en/us/guest/netsol/ns431/c649/ccmigration_09186a008093b876.pdf

High Availability Recovery Analysis

http://www.cisco.com/application/pdf/en/us/guest/netsol/ns431/c649/ccmigration_09186a008093d708.pdf

High Availability Campus Network Design Guide -Implementing Supervisor Redundancy Using NSF, SSO and (StackWise)

http://wwwin-eng.cisco.com/Eng/ESE/Campus/Design_Guides/Campus_Supervisor_Redudancy.pdf@

Note: this is an internal document. Please contact your Cisco account team to have access to this document.

Cisco-recommended Campus Network Design Guides

http://www.cisco.com/en/US/netsol/ns656/networking_solutions_design_guidances_list.html#anchor2

Enterprise QoS Solution Reference Network Design Guide

http://www.cisco.com/application/pdf/en/us/guest/netsol/ns432/c649/ccmigration_09186a008049b062.pdf

HSRP Support for MPLS VPNs:

http://www.cisco.com/en/US/products/sw/iosswrel/ps1834/products_feature_guide09186a00806993c5.html

BGP Support for Next-Hop Address Tracking

http://www.cisco.com/en/US/products/ps6566/products_feature_guide09186a0080796ead.html

MPLS LDP Session Protection:

http://www.cisco.com/en/US/products/sw/iosswrel/ps1829/products_feature_guide09186a00802d95d9.html

Configuring NSF with SSO MSFC Redundancy:

http://www.cisco.com/en/US/docs/switches/lan/catalyst6500/catos/8.x/configuration/guide/nsf_sso.html

APPENDIX **A**

Referenced Role Device Configuration

Table A-1 Access Switch Configuration

Access Switch Configuration	Configuration Comment
vtp domain sjc	Common VTP domain name across the Campus Network
vtp mode transparent	Every Device in the Network configured in VTP Transparent Mode
!	!
switch 1 provision ws-c3750e-24td	Provisioning Switch 1 in Stackwise
switch 2 provision ws-c3750e-24td	Provisioning Switch 2 in Stackwise
switch 1 priority 15	Manually setting Stackwise Master priority from 1 to 15
spanning-tree mode rapid-pvst	STP in RPVST+ Mode
udld message time 7	Change UDLD message timer from default 15 to 7 seconds
vlan 2-51, 151-200	VLAN Range created per Access Layer device
!	!
interface GigabitEthernet1/0/2	Interface facing HSRP Primary device (PE1)
description link to PE1 (int g4/47)	
switchport trunk encapsulation dot1q	802.1Q encap
switchport trunk native vlan 200	Change native from 1 to Global traffic VLAN to prevent VLAN Hopping Attack
switchport trunk allowed vlan	Allowing range of VLAN traffic over the trunk interface
2-51,151-200	
switchport mode trunk	Statically assigning port in Trunk mode
switchport nonegotiate	Disabling DTP packets
load-interval 30	Change default interface stats updates timer from 300 seconds to 30 seconds
carrier-delay msec 0	Change delta timer between Alarm and link deactivation default 2 seconds to 0 msec
udld port aggressive	Enabling UDLD in aggressive
spanning-tree guard loop	Enabling STP Loop guard
!	
interface GigabitEthernet1/0/3	!
description link to PE2 (int g2/48)	Interface facing HSRP Secondary device (PE2)
switchport trunk encapsulation dot1q	
switchport trunk native vlan 200	802.1Q encapsulation
switchport trunk allowed vlan	Change native from 1 to Global traffic VLAN to prevent VLAN Hopping Attack
2-51,151-200	Allowing range of VLAN traffic over the trunk interface
switchport mode trunk	Statically assigning port in Trunk mode
switchport nonegotiate	Disabling DTP packets
load-interval 30	Change default interface stats updates timer from 300 seconds to 30 seconds
carrier-delay msec 0	Change delta timer between Alarm and link deactivation default 2 seconds to 0 msec
udld port aggressive	Enabling UDLD in aggressive
spanning-tree guard loop	Enabling STP Loop guard
!	!
interface GigabitEthernet1/0/14	Interface facing IXIA Test Tool (PE2)
description conn IXIA - 9/1	
switchport trunk encapsulation dot1q	802.1Q encapsulation
switchport trunk native vlan 200	Change native from 1 to Global traffic VLAN to prevent VLAN Hopping Attack
switchport trunk allowed vlan	Allowing range of VLAN traffic over the trunk interface
2-51,151-200	
switchport mode trunk	Statically assigning port in Trunk mode
switchport nonegotiate	Disabling DTP packets
load-interval 30	Change default interface stats updates timer from 300 seconds to 30 seconds
no keepalive	Disabling default keep alive account as data packet to impact convergence result
no cdp enable	Disabling default CDP packets account as data packet to impact convergence result
spanning-tree bpdufilter enable	Disabling STP packets sending out to this port
spanning-tree bpduguard enable	Blocking STP packets to receive from this port
!	!

Table A-2, Part 1 Distribution (PE1) Configuration

Distribution (PE1) Configuration	Configuration Comment
ip routing protocol purge interface	IGP and purge all routes that point to the intf that down to improve convergence time
!	!
!	!
ip tcp window-size 65535	Improves BGPIO by increasing TCP window size
ip tcp path-mtu-discovery	Improves TCP performance to pre-determine MTU size
!	!
vtp domain sjc	Common VTP domain name across the Campus Network
vtp mode transparent	Every Device in the Network configured in VTP Transparent Mode
mls ip cef load-sharing simple	Enable MLS CEF Load Balancing at Layer 3 level
!	!
mpls ldp graceful-restart	Enable GR capability for MPLS LDP protocol
mpls ldp tcp pak-priority	Change the default behavior to switch traffic in special queue
!	!
mpls ldp session protection	Auto-detect and builds Link + Targeted LSP between MPLS device to avoid relearning labels when Link adjacency gets reset for any reason
!	!
mpls ip default-route	No label assigned to default route hence all traffic gets CEF switched. This command enables default route traffic getting label switched
!	!
no mpls ip propagate-ttl	Disabling IP TTL altering by imposition PE
no mpls ldp advertise-labels	Disable advertising MPLS labels for global network
mpls ldp advertise-labels for loopbacks	Advertise MPLS labels only Loopback interface
mpls label protocol ldp	Globally enable LDP protocol
!	!
redundancy	All default sso configuration
keepalive-enable	!
mode sso	!
main-cpu	!
auto-sync running-config	!
!	!
spanning-tree mode rapid-pvst	Enabling RPVST+
spanning-tree vlan 1-4094 priority 24576	Making PE1 / HSRP Primary device as STP Root for all the VLANs
udld message time 7	Change UDLD message timer from default 15 to 7 seconds
process-max-time 50	Prevents any process to take CPU time more than 50 msec, from default 200 msec.
!	!
vlan 2-51,111,151-200,1001	Creating VLANs
!	!
ip vrf site1	Local MPLS VPN VRF
rd 9001:1	Unique RD across network even for same VPN
route-target export 64000:1	Exporting local RD
route-target import 64000:1	Importing PE3 and PE4 RD for same VPN
!	!
interface TenGigabitEthernet1/1	Interface facing Distribution / PE in same Distribution Block (PE2)
dampening	Enabled to add stability for any type Interface flap
ip address 111.111.113.1 255.255.255.0	
ip hello-interval eigrp 1 2	Fine tune hello to 2 seconds. Redundant if BFD is enabled for EIGRP protocol.
ip hold-time eigrp 1 8	Fine tune hold timer to 8 seconds. Redundant if BFD is enabled for EIGRP protocol
load-interval 30	Change default interface stats updates timer from 300 seconds to 30 seconds
carrier-delay msec 0	Change delta timer between Alarm and link deactivation default 2 seconds to 0 msec
mpls ip	Enabling LDP on Core Interface
!	!

Table A-2, Part 2 Distribution (PE1) Configuration

Distribution (PE1) Configuration (Continue)	Configuration Comment
interface TenGigabitEthernet1/3	Interface facing Core2 in same Distribution Block
dampening	All remaining configuration is same as TenGigabitEthernet1/2
ip address 111.111.112.1 255.255.255.0	
ip hello-interval eigrp 1 2	
ip hold-time eigrp 1 8	
load-interval 30	
carrier-delay msec 0	
mpls ip	
!	!
interface TenGigabitEthernet1/4	Interface facing Core1 in same Distribution Block
dampening	All remaining configuration is same as TenGigabitEthernet1/2
ip address 111.111.111.1 255.255.255.0	
ip hello-interval eigrp 1 2	
ip hold-time eigrp 1 8	
load-interval 30	
carrier-delay msec 0	
mpls ip	
!	!
interface GigabitEthernet4/47	Interface facing cat3750 (CE1)
description link to sjc-3750-b1a3	
switchport	
switchport trunk encapsulation dot1q	802.1Q encapsulation
switchport trunk native vlan 200	Change native from 1 to Global traffic VLAN to prevent VLAN Hopping Attack
switchport trunk allowed vlan 2-51,200	Allowing range of VLAN traffic over the trunk interface
switchport mode trunk	Statically assigning port in Trunk mode
switchport nonegotiate	Disabling DTP packets
load-interval 30	Change default interface stats updates timer from 300 seconds to 30 seconds
carrier-delay msec 0	Change timer between Alarm and link deactivation default 2 seconds to 0 msec
udld port aggressive	Enabling UDLD in aggressive
spanning-tree guard root	Enabling STP Loop guard
!	!
interface Vlan2	SVI facing to cat3750
ip vrf forwarding site1	site1 VPN interface. 1 SVI or VLAN per site.
ip address 2.0.0.1 255.255.255.0	VPNv4 connected route
load-interval 30	Change default interface stats updates timer from 300 seconds to 30 seconds
standby 1 ip 2.0.0.254	HSRP Virtual IP address
standby 1 timers msec 250 msec 750	HSRP Timer fine tune
standby 1 priority 150	Priority increased from 100 to 150 on PE1 Primary HSRP/ STP Root device
standby 1 preempt delay minimum 180	Increase preemption delay to minimum of 180 seconds
!	

Table A-3 Distribution (PE1) Configuration (Continue)

Distribution (PE1) Configuration (Continue)	Configuration Comment
router eigrp 1	Single EIGRP AS in complete network
passive-interface Loopback0	Stops outgoing and incoming routing updates on Loopback0 interface
network 1.1.0.0 0.0.255.255	!
network 2.2.2.1 0.0.0.0	Enabling EIGRP on Core Interface
network 111.111.111.0 0.0.0.255	Enabling EIGRP on Core Interface
network 111.111.112.0 0.0.0.255	Enabling EIGRP on Core Interface
network 111.111.113.0 0.0.0.255	Enabling EIGRP on Core Interface
network 192.168.100.1 0.0.0.0	Enabling EIGRP on Loopback 0 Interface
network 200.200.200.0	Enabling EIGRP on SVI 200 for Global Traffic
no auto-summary	Disable auto route summarization
eigrp router-id 192.168.100.1	Statically define EIGRP RID
nsf	Enable NSF capability for EIGRP Routing Protocol
!	!
router bgp 64000	Single BGP AS in complete network
no synchronization	Disable RIB synchronization
bgp router-id 192.168.100.1	Statically define BGP RID
bgp log-neighbor-changes	
bgp graceful-restart restart-time 120	Default BGP GR Restart time
bgp graceful-restart stalepath-time 360	Default BGP GR Stalepath time
bgp graceful-restart	Enable BGP GR capability to all their neighbors
bgp nexthop trigger delay 0	Enable BGP NHT immediately after receiving NH change notification from RIB
neighbor 192.168.100.111 remote-as 64000	Peering with RR1
neighbor 192.168.100.111 update-source Loopback0	
neighbor 192.168.100.112 remote-as 64000	Peering with RR2
neighbor 192.168.100.112 update-source Loopback0	
neighbor 192.168.100.112 fall-over	Enable Fast Deactivation of BGP relationship
maximum-paths ibgp 2	Enable IPv4 Unicast Multipath iBGP
no auto-summary	
!	!
address-family vpnv4	VPNv4 AF
neighbor 192.168.100.111 activate	Activate RR1 as VPNv4 peer
neighbor 192.168.100.111 send-community both	Send/Receive standard and extended BGP Community
neighbor 192.168.100.112 activate	Activate RR2 as VPNv4 peer
neighbor 192.168.100.112 send-community both	Send/Receive standard and extended BGP Community
bgp nexthop trigger delay 0	Enable VPNv4 BGP NHT after receiving NH change notification from RIB
exit-address-family	
!	!
address-family ipv4 vrf site1	IPv4 VRF AF
redistribute connected	Redistribute connected SVI network
redistribute static	Redistribute per VRF 49 static route into VPN
maximum-paths ibgp 2 import 4	Enable BGP Multipath for VPNv4 prefix
no synchronization	Disable RIB synchronization
exit-address-family	
!	!
mpls ldp router-id Loopback0 force	Force MPLS LDP to consider Loopback0 as RID
!	!

Table A-4, Part 1 Core 1 Configuration

Core 1 Configuration	Configuration Comment
ip routing protocol purge interface	IGP and purge all routes point to intf that went down to improve convergence time
!	!
ip tcp window-size 65535	Improve BGPIO by increasing TCP window size
ip tcp path-mtu-discovery	Improve TCP performance to pre-determine MTU size
!	!
vtp domain sjc	Common VTP domain name across the Campus Network
vtp mode transparent	Every Device in the Network configured in VTP Transparent Mode
mls ip cef load-sharing simple	Enable MLS CEF Load Balancing at Layer 3 level
!	!
mpls ldp graceful-restart	Enable GR capability for MPLS LDP protocol
mpls ldp tcp pak-priority	change the default behavior to switch traffic in special queue
!	!
mpls ldp session protection	Autodetects and builds Link + Targetted LSP between MPLS device to avoid relearning labels when Link adjacency gets reset for any reason
!	!
mpls ip default-route	No label assigned to default route hence all traffic gets CEF switched. This command enables default route traffic getting label switched
!	!
mpls label protocol ldp	Globally enabling LDP protocol
!	!
redundancy keepalive-enable mode sso main-cpu auto-sync running-config	All default sso configuration
!	!
spanning-tree mode rapid-pvst	Enabling RPVST+
!	!
process-max-time 50	Prevent any process to CPU to take more than 50 msec
!	!
interface TenGigabitEthernet1/1 dampening ip address 111.111.114.2 255.255.255.0	Interface facing Distribution / PE2 in same Distribution Block (PE2) Enabled to add stability for any Interface flap
ip hello-interval eigrp 1 2	Fine tune hello to 2 seconds.
ip hold-time eigrp 1 8	Fine tune hold timer to 8 seconds.
!	!
load-interval 30	Change default interface stats updates timer from 300 to 30 seconds
carrier-delay msec 0	Change timer between Alarm and link deactivation from 2000 to 0 msec
mpls ip	Enabling LDP on Core Interface
!	!
interface TenGigabitEthernet1/3 dampening ip address 111.111.112.1 255.255.255.0 ip hello-interval eigrp 1 2 ip hold-time eigrp 1 8 load-interval 30 carrier-delay msec 0 mpls ip	Interface facing Core2 in same Distribution Block All remaining configuration is same as TenGigabitEthernet1/1
!	

Table A-4, Part 2 Core 1 Configuration

Core 1 Configuration (Continue)	Configuration Comment
!	!
interface TenGigabitEthernet1/4	Interface facing PE1 in same Distribution Block
dampening	All remaining configuration is same as TenGigabitEthernet1/1
ip address 111.111.111.1 255.255.255.0	
ip hello-interval eigrp 1 2	
ip hold-time eigrp 1 8	
load-interval 30	
carrier-delay msec 0	
mpls ip	
!	!
interface GigabitEthernet3/48	Interface facing RR1 in same Distribution Block
description link to sjc-7206-rr1 (link f2/1)	All remaining configuration is same as TenGigabitEthernet1/2
dampening	
ip address 151.151.151.1 255.255.255.0	
load-interval 30	
carrier-delay msec 0	
mpls ip	
!	!
router eigrp 1	Single EIGRP AS in complete network
passive-interface Loopback0	Stops outgoing and incoming routing updates on Loopback0 interface
network 111.111.111.0 0.0.0.255	Enabling EIGRP on Core Interface
network 111.111.114.0 0.0.0.255	Enabling EIGRP on Core Interface
network 131.131.131.0 0.0.0.255	Enabling EIGRP on Core Interface
network 131.131.132.0 0.0.0.255	Enabling EIGRP on Core Interface
network 131.131.135.0 0.0.0.255	Enabling EIGRP on Core Interface
network 151.151.151.0 0.0.0.255	Enabling EIGRP on Core Interface
network 192.168.100.2 0.0.0.0	Enabling EIGRP on Loopback 0 Interface
no auto-summary	Disable auto route summarization
eigrp router-id 192.168.100.2	Statically define EIGRP RID
nsf	Enable NSF capability for EIGRP Routing Protocol
!	
mpls ldp router-id Loopback0 force	Force MPLS LDP to consider Loopback0 as RID
!	!

Table A-5 Router Reflector 1 Configuration

Route Reflector 1 Configuration	Configuration Comment
ip cef	Enable cef
ip tcp window-size 65535	Improve BGPIO by increasing TCP window size
ip tcp path-mtu-discovery	Improve TCP performance to pre-determine MTU size
!	!
interface Loopback0 ip address 192.168.100.111 255.255.255.255	
!	!
interface GigabitEthernet0/2	Interface facing Core
dampening ip address 151.151.151.2 255.255.255.0	Enable to add stability for any Interface flap
ip hello-interval eigrp 1 2 ip hold-time eigrp 1 8 load-interval 30 carrier-delay msec 0 duplex full speed 1000	Fine tune hello to 2 seconds. Redundant if BFD is enabled for EIGRP protocol. Fine tune EIGRP hold timer to 8 seconds Change default interface stats updates timer from 300 seconds to 30 seconds Change timer between Alarm and link deactivation default 2 seconds to 0 msec
!	
router eigrp 1	Single EIGRP AS in complete network
passive-interface Loopback0	stops outgoing and incoming routing updates on Loopback0 interface
network 151.151.151.0 0.0.0.255	Enable EIGRP on connected Core Interface
network 192.168.100.111 0.0.0.0	Enabling EIGRP on Loopback 0 Interface
no auto-summary	Disable auto route summarization
eigrp router-id 192.168.100.111	Statically define EIGRP RID
!	
router bgp 64000	Single BGP AS in complete network
no synchronization	Disable RIB synchronization
bgp router-id 192.168.100.111	Statically define BGP RID
bgp cluster-id 64000	
bgp log-neighbor-changes	
bgp graceful-restart restart-time 120	Default BGP GR Restart time
bgp graceful-restart stalepath-time 360	Default BGP GR Stalepath time
bgp graceful-restart	Enable BGP GR capability to all their neighbors
bgp nexthop trigger delay 0	Enable BGP NHT immediately after receiving NH change notification from RIB
neighbor rr-clients peer-group	Define rr-client peer-group
neighbor rr-clients remote-as 64000	
neighbor rr-clients update-source Loopback0	
neighbor rr-clients send-community extended	Peering with rr-client PE1
neighbor 192.168.100.1 peer-group rr-clients	Peering with rr-client PE2
neighbor 192.168.100.4 peer-group rr-clients	Peering with rr-client PE3
neighbor 192.168.100.5 peer-group rr-clients	Peering with rr-client PE4
neighbor 192.168.100.8 peer-group rr-clients	Peering with RR2
neighbor 192.168.100.112 remote-as 64000	
neighbor 192.168.100.112 update-source Loopback0	
maximum-paths ibgp 2	
no auto-summary	
!	
address-family vpnv4	Enable IPv4 Unicast Multipath iBGP VPNv4 AF
neighbor rr-clients send-community both	Send/Receive standard and extended BGP Community
neighbor rr-clients route-reflector-client	Activate rr-clients VPNv4 peers
neighbor 192.168.100.1 activate	Activate VPNv4 PE1 peer
neighbor 192.168.100.4 activate	Activate VPNv4 PE2 peer
neighbor 192.168.100.5 activate	Activate VPNv4 PE3 peer
neighbor 192.168.100.8 activate	Activate VPNv4 PE4 peer
neighbor 192.168.100.112 activate	Activate RR2 as VPNv4 peer
neighbor 192.168.100.112 send-community extended	Send/Receive standard and extended BGP Community
bgp nexthop trigger delay 0	Enable VPNv4 BGP NHT to receive NH change notification from RIB
exit-address-family	

APPENDIX **B**

Test Case Descriptions and Results

Revised: Janaury, 2008

B.1 System Integration Test Suite

Test	Manual Test Case	Defects	Automation Test Case	Defects
System Integration Test Suite: This test suite has a combination of various features configured in the NV-PI testbed. The features include VLANs, Rapid-PVST+, Trunks, UDLD (in access, Distribution Layer switches), VRFs, HSRP, EIGRP(as IGP), MP-iBGP, MPLS (Distribution Layer switches), EIGRP and MPLS (in core switches), Multicast, QoS.				
IP Baseline The IP Baseline provides the initial configuration and background traffic for Network Virtualization Path Isolation. Configuration includes EIGRP, Multicast, Voice and QoS. The baseline configuration serves as an initial configuration of Network Virtualization services. Background traffic (stateful / stateless) using SmartBit and Avalanche test tools . The test setup includes: 5000 mroutes 3000 EIGRP routes 100 Stateful sessions (TELNET + FTP + HTTP + DNS + POP3) 100 Mbps QoS traffic (includes Voice, Multicast Video, Call Control, bulk data, critical data and best effort traffic) based on *Enterprise QoS Solution Reference Network Design Guide*	—	—	Passed	—
NV – Path Isolation Infrastructure The integrated features provisioning is listed below				

HSRP Peers Configure and verify HSRP are forming peers between the distribution devices. There are two distribution routers in each distribution block. The first router (PE1/PE3) will act as active HSRP device and the second router (PE2/PE4) will act as the standby HSRP router. The HSRP priorities are adjusted to achieve this objective. The standby HSRP router is configured with preempt delay for 180 sec. The “hello” and “hold” timers for HSRP peers are adjusted to 250 and 750 msec to achieve fast convergence. - Verify the HSRP functionality and the roles by issuing “sh standby brief” in both active and standby HSRP devices.	Passed	—	—	—
VlAN/VRF Mapping Configure and verify VLAN to VRF mapping on the distribution routers (PE1, PE2, PE3 and PE4). The Layer3 interface for that particular VLAN has an extra configuration to bind the VRF to the corresponding VLAN. - Verify the configuration with the CLI – “sh vrf <vrfname> for proper mapping.	Passed	—	—	—
EIGRP neighbors Configure and verify the EIGRP neighbors on Distribution routers (PE1,PE2,PE3,PE4), and Core routers (C1, C2, C3, C4) and RR routers (RR1,RR2). EIGRP routing protocol is configured in AS 1. All the loopbacks networks and direct connected links are advertised into EIGRP. EIGRP’s “hello” and “hold” timers are adjusted to 2 and 8 seconds respectively to achieve fast convergence. EIGRP is also configured to operate in NSF mode (EIGRP NSF command inside the EIGRP process). - Verify the EIGRP neighbors using “sh ip EIGRP neighbors” command in all the routers mentioned in this sub-test case. 2. Core routers (C2 and C4) have trunks (carrying 100 different vlans on each link) to pagent router for simulating 50 EIGRP neighbors on each links. The pagent router is configured to send 60 routes per neighbor, which total 3000 EIGRP routes. - Verify in the distribution routers (PE1/PE2/PE3 and PE4) – sh ip route summary for 3000 EIGRP routes.	Passed	—	—	—

BGP neighbors Configure and verify the BGP neighbor relationship between distribution routers (PE1/PE2/PE3 and PE4) and Route Reflectors (RR1 and RR2). BGP is configured in AS 64000 peering from PE1,PE2,PE3 and PE4 and two Route reflectors. VRF address family is configured under BGP to import all the static routes configured in the remote PE devices to the respective VRFs. VPN address family under BGP is configured to export all the locally originated routes to the Errs. Maximum number of redundant paths is configured to 2. BGP is also configured NSF capable. - Verify the bgp neighbor relationship and the number of routes using 'sh ip bgp vpnv4 all summary" command.	Passed	—	—	—
MPLS neighbors Configure and verify MPLS neighbor relationship on distribution routers (PE1,PE2, PE3, and PE4), core routers (C1, C2, C3, C4) and Route reflectors (RR1, RR2). MPLS is configured on these routers for data switching. LDP is used as the label switching protocol. MPLS LDP session protection is configured on all these routers to achieve fast convergence. Verify the following CLI on these routers, for MPLS neighbors – "sh mpls ldp neighbor" and "sh mps ldp discovery" as well.	Passed	—	—	—

B.2 Scalability Test Suite

Test	Manual Test Case	Defects	Automation Test Case	Defects
NV - Path Isolation - Scalability Test Suite This test suite builds on top of the previous – NV – Path Isolation System integration test suite where we - Configure the distribution routers for 100 vlans, 50 VRFs.				
Vlan scalability Configure and verify scaling VLANs on Access Layer switches and Distribution devices (PE1,PE2,PE3 and PE4). - Configure the devices (mentioned in this test case) for 100 vlans. - Verify the configuration, by issuing “sh vlan summary” command on the devices configured.	Passed	—	—	—
VRF scalability Configure and verify the scaled configuration of VRFs in the Distribution devices (PE1,PE2, PE3 and PE4). - Configure the following on the Distribution devices: Creation of 50 VRFs Route-distinguisher, Route target (export and import for this VRF). Mapping of VRF to its corresponding VLANs. - Verify the configurations by issuing “sh vrf” command on the devices configured.	Passed	—	—	—
MP-iBGP – VRF address family scaled configuration. Configure and verify the scaled configuration on MP-iBGP VRF address family on the distribution routers (PE1,PE2,PE3 and PE4). - Configure the following on the distribution routers: Create a VRF address family under BGP routing process for each VRFs (- Configured above) Redistribute all the connected and static routes. - Verify the configuration by issuing “sh ip bgp vpnv4 vrf <vrfname>.	Passed	—	—	—

B.3 High Availability Test Suite

Test	Manual Test Case	Defects	Automation Test Case	Defects
This test suite builds on top of the previous NV – Path Isolation Scalability test suite. - Focus on Supervisor switchover with MPLS HA features (NSF/SSO - MPLS LDP and LDP Graceful Restart and NSF/SSO – MPLS VPN) in Core and Distribution Layer. NSF/SSO - MPLS LDP and LDP Graceful Restart - Multiprotocol Label Switching (MPLS) Label Distribution Protocol (LDP) uses SSO, NSF, and graceful restart to allow a Route Processor to recover from disruption in control plane service (specifically, the LDP component) without losing its MPLS forwarding state. LDP NSF works with LDP sessions between directly connected peers and with peers that are not directly connected (targeted sessions). NSF/SSO – MPLS VPN - This feature allows a provider edge (PE) router with redundant Route Processors to preserve data forwarding information in a Multiprotocol Label Switching (MPLS) Virtual Private Network (VPN) when the primary Route Processor restarts. - It also focus on Supervisor switchover with SSO in Access Layer - Network health is monitored to make sure End-to-End Traffic is not disrupted during NSF/SSO switchover				
NSF/SSO switchover in Core Layer Active route processor is forced to switchover and this results in the standby RP to taking over. During this switchover period, data traffic is still actively forwarding. - Use the 'show redundancy states' command to check if both supervisors are up. One should be active and other "Standby Hot." -Verify that redundancy mode is SSO. - Use the command 'redundancy force-switchover' - Use the command 'show logging' and grep for traceback if any. - Verify data traffic is continue forwarding and did not get disrupted.	Passed	—	Passed	—
HA: NSF/SSO switchover in Distribution Layer Test Active route processor is forced to switchover and this results in the standby RP to taking over. During this switchover period, data traffic is still actively forwarding. - Use the 'show redundancy states' command to check if both supervisors are up. One should be active and other "Standby Hot." -Verify that redundancy mode is SSO. - Use the command 'redundancy force-switchover' - Use the command 'show logging' and grep for traceback if any. - Verify data traffic is continue forwarding and did not get disrupted.	Passed	—	Passed	—

4k access SSO switchover This type of switchover ensures that Layer2 traffic is not interrupted - Create the SSO switchover on the Access router 'c4500-a1' using the command 'redundancy force-switchover' - Use the 'show redundancy states' command to check if both supervisors are up. One should be active and other "Standby Hot." -Verify that redundancy mode is SSO. - Verify that Layer2 traffic is not interrupted during the SSO switchover. - Use the command 'show logging' and grep for traceback if any.	Passed		Passed	
6k access SSO switchover This type of switchover ensures that Layer2 traffic is not interrupted - Create the SSO switchover on the Access router 'c6500-a1' using the command 'redundancy force-switchover' - Use the 'show redundancy states' command to check if both supervisors are up. One should be active and other "Standby Hot." -Verify that redundancy mode is SSO. - Verify that Layer2 traffic is not interrupted during the SSO switchover. - Use the command 'show logging' and grep for traceback if any.	Passed		Passed	

B.4 Performance (IP/MPLS Convergence) Test Suite

Test	Manual Test Case	Defects	Automation Test Case	Defects
NV - Path Isolation – Convergence Test Suite: This test suite builds on top of the previous – NV – Path Isolation System integration test suite where HA Link failure/restoration and node failure are done in access layer device, Distribution devices, core nodes and RRs. All the test scenarios below were executed with VPN and Global traffic in Upstream and Downstream direction. The network convergence is measured and characterized with following failure scenarios: Uplink fiber to active HSRP failure Cat3750 StackWise box failure PE Supervisor Switchover PE node failure - PE - Core link failure Core Supervisor Switchover Core Node Failure Core to RR link failure RR node failure				
Uplink fiber to active HSRP failure Measure the network convergence numbers for the link failure introduced between the active HSRP (PE1) and the access layer switch (3k-a1) - Start sending the traffic into the network under stable condition. Traffic sent must be equal to traffic received. - Wait for few minutes under this condition. - Do a Link failure between(PE1 and 3k-a1) on the network. Do a “shut” command on the upstream DUT’s (PE1) link towards 3k-a1 to simulate the physical link pull out. (Traffic starts taking the alternate path to the destination) - Wait for few minutes for the counters to be updated in the test tool. - Record the frames lost and calculate the convergence number (ms) using the formula = (tx-rx)*pps/1000. The convergence numbers are calculated and recorded for VPN and Global traffic (upstream and downstream).	Passed	--	—	—

Cat3750 (StackWise) master box failure Measure the network convergence numbers for the Cat3750 master box failure in the network. - Start sending the traffic into the network under stable condition. Traffic sent must be equal to traffic received. - Wait for few minutes under this condition. - Do a node failure of the master Cat3750 in the network. (Since the master box only participates in the control plane and is not on the data path, traffic is not affected because of this failure). - Wait for few minutes for the counters to be updated in the test tool. - Record the frames lost and calculate the convergence number (ms) using the formula = (tx-rx)*pps/1000. The convergence numbers are calculated and recorded for VPN and Global traffic (upstream and downstream).	Passed	---	--	--
PE Supervisor Switchover (NSF/SSO) This sub-test will test and record the convergence numbers for the NSF/SSO switchover on the active HSRP (PE1). -Start sending the traffic into the network under stable condition. Traffic sent must be equal to traffic received. -Wait for few minutes under this condition. -Initate a NSF/SSO switchover on the active HSRP node (PE1) on the network. This is done by executing the CLI - “redundancy force-switchover” command in PE1. (Traffic starts taking the alternate path to the destination) -Wait for few minutes for the counters to be updated in the test tool. -Record the frames lost and calculate the convergence number (ms) using the formula = (tx-rx)*pps/1000.	Passed	---	Passed	---

PE node failure Measure the network convergence numbers for the node failure on the active HSRP (PE1). - Start sending the traffic into the network under stable condition. Traffic sent must be equal to traffic received. - Wait for few minutes under this condition. - Power down the active HSRP node (PE1) on the network. (Traffic starts taking the alternate path to the destination) - Wait for few minutes for the counters to be updated in the test tool. - Record the frames lost and calculate the convergence number (ms) using the formula = (tx-rx)*pps/1000. The convergence numbers are calculated and recorded for VPN and Global traffic (upstream and downstream).	Passed	--	--	--
PE to Core Link failure This sub-test will test and record the convergence numbers for the link failure introduced between the distribution switch (PE1) and the core switch (C1). -Start sending the upstream traffic into the network under stable conditions. Traffic sent must be equal to traffic received. -Wait for few minutes under this condition. -Do a Link failure between(PE1 and C1) on the network. Remove Cable between PE1 and C1. Traffic starts taking the alternate path to the destination -Wait for few minutes for the counters to be updated in the test tool. -Record the frames lost and calculate the convergence number (ms) using the formula = (tx-rx)*pps/1000 .	Passed	---	--	--

Core Supervisor Switchover (NSF / SSO) This sub-test will test and record the convergence numbers for the NSF/SSO switchover on the core node(C1). -Start sending into the network under stable condition. Traffic sent must be equal to traffic received. -Wait for few minutes under this condition. -Initate a NSF/SSO switchover on the core node (C1) on the network. This is done by executing the CLI - “redundancy force-switchover” command in C1. (Traffic starts taking the alternate path to the destination) -Wait for few minutes for the counters to be updated in the test tool. -Record the frames lost and calculate the convergence number (ms) using the formula = (tx-rx)*pps/1000.				
Core Node failure Measure the network convergence numbers for the Core node failure (C1). - Start sending the traffic into the network under stable condition. Traffic sent must be equal to traffic received. - Wait for few minutes under this condition. - Power down the remote Core node (C1) on the network. - Wait for few minutes for the counters to be updated in the test tool. - Record the frames lost and calculate the convergence number (ms) using the formula = (tx-rx)*pps/1000. The convergence numbers are calculated and recorded for VPN and Global traffic (upstream and downstream).	Passed	--	--	--

Core to RR link failure Measure the network convergence numbers for the link failure introduced between the RR node (RR1) and the core switch (C1). - Start sending the traffic into the network under stable condition. Traffic sent must be equal to traffic received. - Wait for few minutes under this condition. - Do a Link failure between(RR1 and C1) on the network. Do a “shut” command on the upstream DUT (RR1) link towards C1 to simulate the physical link pull out. (Since this link is not on the data traffic path, traffic is not affected because of the link failure). - Wait for few minutes for the counters to be updated in the test tool. - Record the frames lost and calculate the convergence number (ms) using the formula = (tx-rx)*pps/1000. The convergence numbers are calculated and recorded for VPN and Global traffic (upstream and downstream).	Passed	--	--	--
RR node failure Measure the network convergence numbers for the node failure introduced in RR (RR1). - Start sending the traffic into the network under stable condition. Traffic sent must be equal to traffic received. - Wait for few minutes under this condition. - Do a node failure in RR1 on the network. (Since this node is not on the data traffic path, traffic is not affected because of the node failure). - Wait for few minutes for the counters to be updated in the test tool. - Record the frames lost and calculate the convergence number (ms) using the formula = (tx-rx)*pps/1000. The convergence numbers are calculated and recorded for VPN and Global traffic (upstream and downstream).	Passed	---	--	--

B.5 Reliability Test Suite

Test	Manual Test Case	Defects	Automation Test Case	Defects
NV - Path Isolation - Reliability Test Case: This test case relies on the previous test case where the testbed is run continuously for an extended period of time. During this time, a link failure event or NSF/SSO switchover is done (simulating a real life environment). Network health is monitored and make sure there is no CPU hog, memory leak and End-to-End traffic is not disrupted during 150 hours period.				
System Reliability -Verify End-to-End traffic on the testbed and introducing failures (link failures and NSF/SSO switchover). -Start with a stable network. End to End traffic is verified for no loss. The setup is run for an extended period of time (usually 150 hours). During this time, random failure scenarios are introduced in the network (link failures/ NSF/SSO switchover etc.). - Verify the end-end traffic at the end of the test duration for traffic loss (The traffic loss should account only for the failures introduced in the network)	Passed	—	—	—

APPENDIX **C**

Defects and Technical Notes

C.1 Defects

Note There were no operationally impacting defects encountered in the Network Virtualization-Path Isolation system assurance validation testing.

C.2 Technical Note

Description: EIGRP “hello” timer times out upon RP switchover. (hello/hold - 1 sec and 4 sec configuration)
Root cause: This is a very aggressive timer when used in SSO mode.
Solution: Use the recommended timers (2/8 for “hello” and “hold” timers) to avoid this situation.