Switch Clustering Best Practices

NN48500-584
Version 8.1

December 2015

Document Use

The intention of this document is to provide a quick overview of the Avaya recommended Best Practices for implementing Switch Clustering. Please note that the recommendations may vary between designs based on the hardware platforms being used and the feature set available on each switching product. Avaya always recommends reviewing the release notes for each product before deployment. This will help to avoid any unexpected results during network operations due to software limitations or known issues with specific releases of code.

Document Version Control

- Version 2.0 published April 2009
  - Many document updates to add more information and revise format
  - Revised VLACP timer recommendations
  - Added FDB timer change
  - Added ERS 5000 Design Requirements section
- Version 2.1 published May 2009
  - Clarified FDB timer change to apply to ERS 8800 / 8600 / 8300 / 1600 only
  - Added information regarding SLPP in scaled environments
  - Added information regarding SLPP reset to clear threshold counters
  - VRRP guidelines reference ERS 5000 design requirements for additional recommendations
  - Revised recommendation on Port Rate Limiting – access ports only
  - Added additional information in ERS 5000 design requirements section – more details
- Version 2.2 published July 2009
  - Corrected timer from Hold-Down to Hold-Up for RSMLT Layer 2 Edge
- Version 3.0 published March 2012
  - Added SLPP-Guard information
- Version 4.0 published September 2012
  - Updated content
  - Added IPv6
- Version 5.0 published May 2014
  - Added VSP 7000 and VSP 8000
  - Added vIST
- Version 7.0 published February 2015
  - Updated vIST

Document Version Control

- Version 8.0 published October 2015
  - Added VSP 7200
  - CP Limit chart updates
  - SLPP update
- Version 8.1 published December 2015
  - Updated VSP 7000 SMLT behavior when SPBM is enabled

Reference Information

In addition to the guidelines provided here, the following documents are available which provide more detailed information.

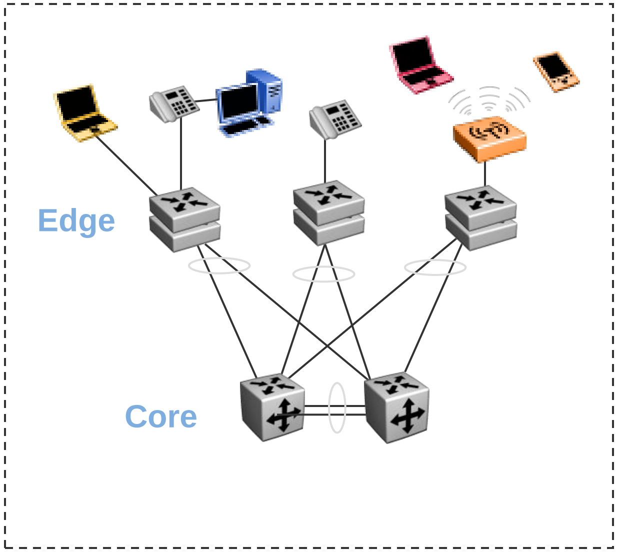
- *Switch Clustering using Split Multi-Link Trunking (SMLT) with VSP 9000, VSP 7200, VSP 4000, VSP 8000, VSP 7000, ERS 8600/8800, and 5000 Technical Configuration Guide (NN48500-518)*
- *Switch Clustering Supported Topologies and Interoperability with Virtual Services Platform 9000 & Ethernet Routing Switches(NN48500-555)*
- *Resilient Multicast Routing Using Split-Multilink Trunking for the ERS 8800/8600 Technical Configuration Guide (NN48500-544)*

Avaya's Ethernet Switching Vision

Delivering the Foundation for Enterprise Networks

Provide an always-on Ethernet infrastructure enabling uninterrupted access to Enterprise applications and services.

- Simple and Efficient Network Architectures
- Active / Active with fast failover and recovery
- High performance products and solutions
- Integrated security
- Technology innovation
- Energy Efficiency in every product and solution
- Price / Performance leader



Resiliency – Performance – Security – Efficiency

Reliability Optimization

Choosing the Right Model

	Active–Passive n + 1	Active–Active n – 1
Normal Operation	🟢 🔴	🟢 🟢
Failure Operation	🔴 🟢 ▸ Hot Standby ▸ Size to constant capacity ▸ Cost = 1/Avail x Price ▸ Transition Timing	🔴 🟢 • Distributed Use • Size to Need • Reduce Capacity in Failure • Use Business Rules

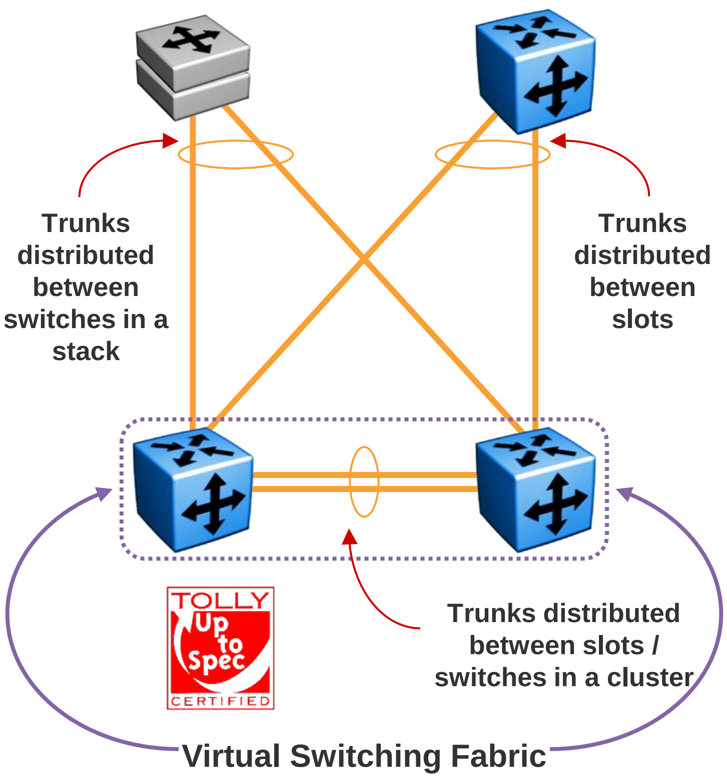
High Performance & Reliable IP Core

Active–Active

- Switch Clustering
  - Link and Nodal Redundancy (N-1)
  - Split Multilink Trunking (SMLT)
  - Routed Split Multilink Trunking (RSMLT)
- Combines Resiliency & Performance
  - All links passing traffic
  - Sub second stateful failover
  - No Spanning Tree on switch to switch links
- Virtual “Switch” Fabric
  - Centralized or distributed
  - Increased performance using full capacity
  - No single point of failure
- Scalable
  - Standalone / Stackable / Modular Chassis
  - Single pair / Square / Full Mesh
  - Mbps → Gbps → Tbps

Switch Clustering

Interoperable with ANY 3rd party edge device



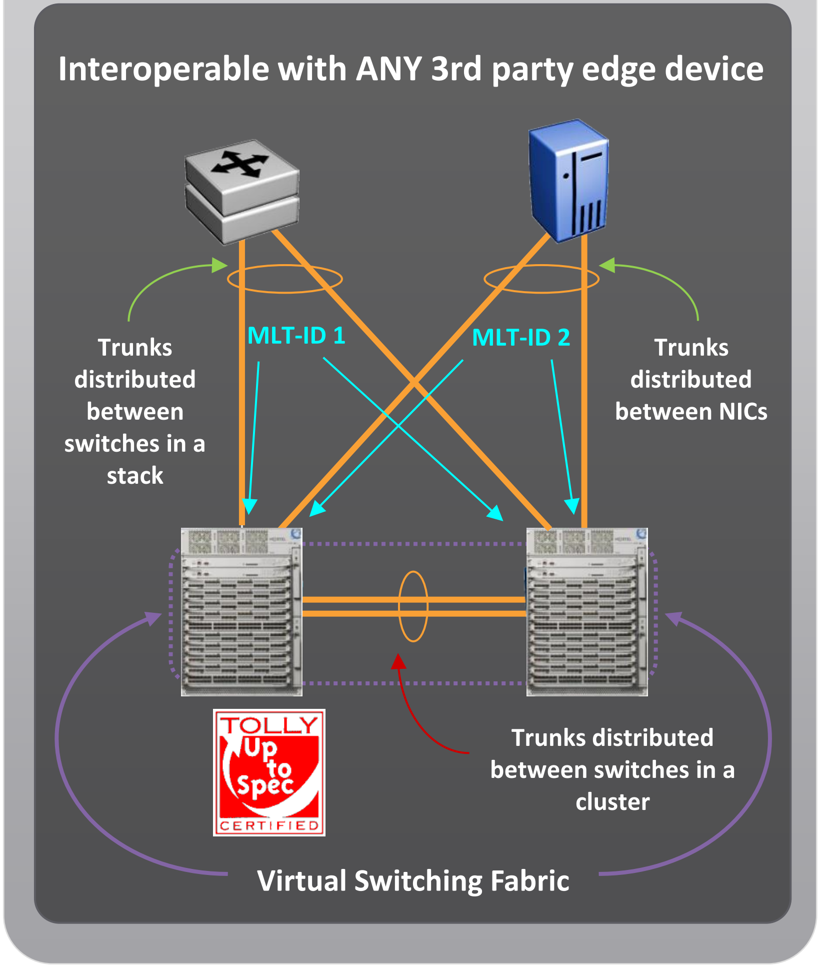
Trunks distributed between switches in a stack
Trunks distributed between slots
TOLLY Up to Spec CERTIFIED
Trunks distributed between slots / switches in a cluster
Virtual Switching Fabric

Switch Clustering – SMLT(v2)

VSP 9000

- Virtual “Switch” Fabric
  - New Switch Cluster HW infrastructure
  - Future extension to multiple IST’s per switch possible
- Configuration
  - Single step cluster-link configuration
  - LinkBundle ID (MLT-ID) is now cluster significant
  - MLT index needs to be chosen uniquely for a local MLT and/or Split-MLT
  - SMLT-ID has been removed
- Port Failure and recovery behavior
  - System uses sophisticated hardware support for rapid failover
    - For local MLT failover
    - For IST-peer MLT failover
  - Milliseconds stateful failover

Switch Clustering v2


Interoperable with ANY 3rd party edge device
MLT-ID 1
MLT-ID 2
Trunks distributed between switches in a stack
Trunks distributed between NICs
TOLLY Up to Spec CERTIFIED
Trunks distributed between switches in a cluster
Virtual Switching Fabric

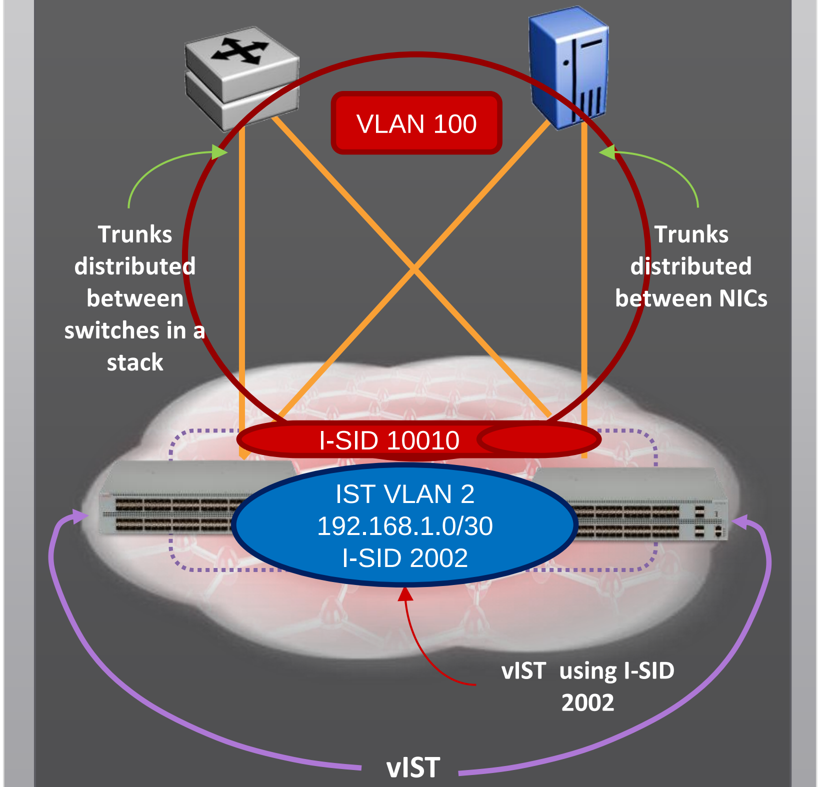
Switch Clustering – Virtual IST (vIST)

VSP 8000, VSP 7200, VSP 4000

- Virtual IST
  - Supported on the VSP 8000 in release 4.0 , the VSP 4000 in release 4.1, and VSP 7200 in 4.2.1
  - No longer requires MLT between Cluster switches
    - Removes the need for physical cabling
  - The IST is up as long as there is SPBM connectivity between the IST peers
- vIST Configuration
  - Create a L2VSN – VLAN with added IP address
  - SPB IST configuration – Peer with B-MAC of neighbor and use the same virtual B-MAC on Cluster switches
  - vIST configuration - Peering with the peer Cluster IP address & VLAN using the L2VSN VLAN ID and IP address
  - SPB IP does not need to be enabled
- vIST – VLAN Configuration (CVLAN)
  - You cannot assign edge VLANs to IST link as there are no physical ports
  - You have to assign an I-SID to the edge VLAN
  - SMLT ID has been removed

Virtual IST

Interoperable with ANY 3rd party edge device



Switch Clustering –Simplified vIST

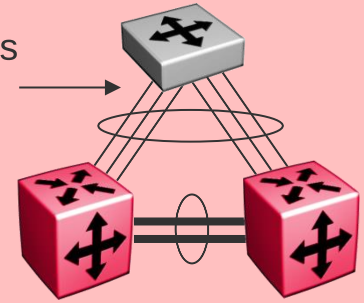
VSP 8000, VSP 7200, VSP 4000

- Please note that vIST is only supported over a Shortest Path Bridge (SPB) network
  - For non-SPB environments, you will need to enable SPB on the two VSP cluster switches
- Beginning in the 4.0.1 release for the VSP 8000, 4.1 for the VSP 4000, and 4.2.1 for the VSP 7200, a simplified vIST option is available that auto-configures SPB
  - No explicit SPB parameters are required to be entered by the end user
  - The simplified vIST script will auto-provision all the necessary SPBM parameters without any user intervention

Switch Clustering

Terminology – SMLT & SLT

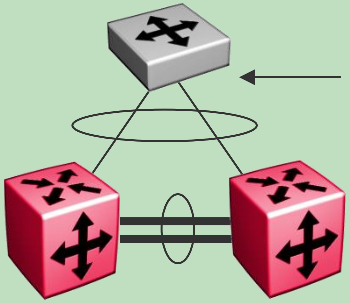
SMLT – Split MultiLink Trunking
“Standard” layer 2 design using
MLT-based connections



Scaling Uplinks
Between the
Edge and the
Core

Maximum Number of SMLT’s supported per Cluster is number of MLT groups supported less one required for IST

SLT – Single Link Trunking
“Standard” layer 2 design using
port-based connections



Maximum Two
Uplinks per
Edge
Connection

Maximum Number of SLT’s supported per Cluster is number of ports on one core switch less two required for IST

Note: VSP 9000, VSP 4000, VSP 7200 & VSP 8000 only supports SMLT, scales to avoid need for SLT

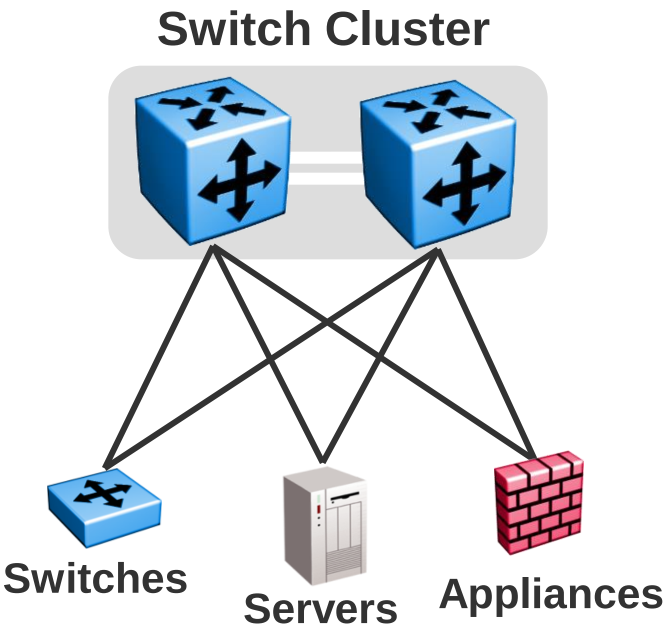
Can Configure both SMLT and SLT on a single Switch Cluster

Note: ERS 8000, ERS 5000 and VSP 7000 both support SLT configuration to increase SMLT scaling

How Do I Build a Switch Cluster ?

Physical Design

- **Create a Switch Cluster core with two like Ethernet Routing Switches**
  - ○ ERS 8800/8600
  - ○ ERS 8300
  - ○ ERS 5000
  - ○ VSP 9000
  - ○ VSP 7000
  - ○ VSP 4000/7200/8000
- **Create the Inter-Switch Trunk (IST) between core switches**
  - Multilink Trunk (MLT) for resiliency
  - Responsible for forwarding/control synchronization
  - Must be the same speed: 10Mbps to 10Gbps
- **Connect edge devices**
  - SMLT, SLT, RSMLT on the Switch Cluster
  - Link aggregation configuration (802.1AX, MLT, etc.)
  - Disable Spanning Tree on link aggregation group
  - Autonegotiation on 100FX & GbE interfaces for Remote Fault Indication (RFI) support
    - There is no Autonegotiation on 10 GbE which has integrated RFI support



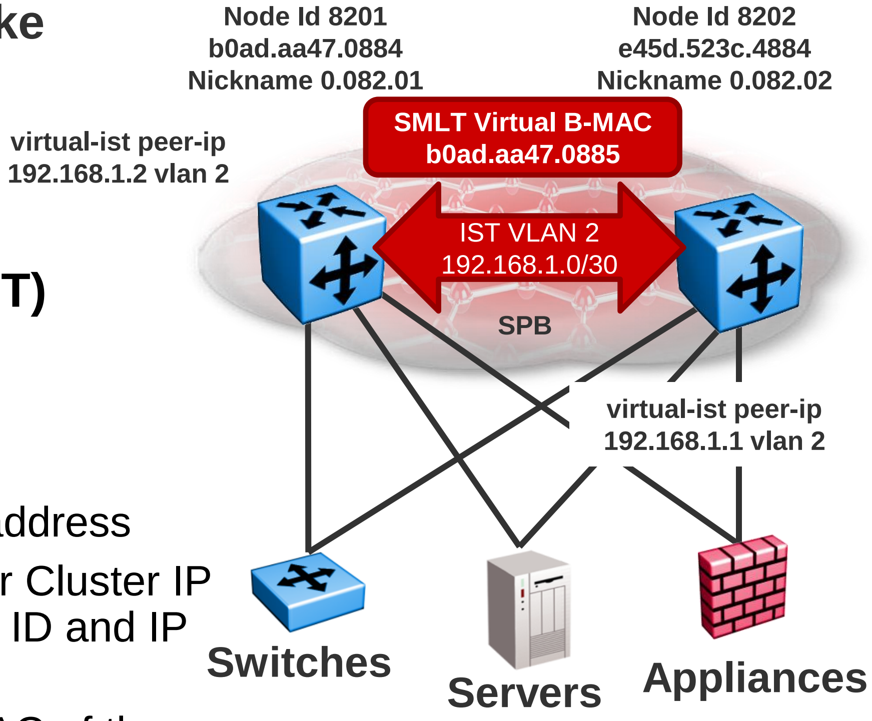
The “Simple” Aspect Includes the Configuration

How Do I Build a Switch Cluster ?

Virtual IST

- **Create a Switch Cluster core with two like Ethernet Routing Switches**
  - VSP 8000 (release 4.0)
  - VSP 4000 (release 4.1)
  - VSP7200 (release 4.2.1)
- **Create the Virtual Inter-Switch Trunk (vIST) between core switches**
  - SPB core required
  - Create vIST between Cluster switches
    - Create a L2VSN – VLAN with added IP address
    - vIST configuration - Peering with the peer Cluster IP address & VLAN using the L2VSN VLAN ID and IP address
  - SPB IST configuration - Peering with the B-MAC of the peer Cluster switch and use the same virtual B-MAC on both cluster switches
- **Connect edge devices**
  - SMLT, RSMLT on the Switch Cluster
  - Cannot assign edge VLANs to IST links as there are no physical ports
  - Have to assign an I-SID to edge VLAN

Switch Cluster



** Please note, the VSP 8000 VSP 7200, and VSP 4000 are vIST interoperable

How Do I Build a Switch Cluster ?

Simplified Virtual IST

- Both the VSP 4000, VSP 7200, and VSP 8000 support a simplified Virtual-IST configuration in order to enable seamless migration of legacy IST based SMLT to Virtual-IST based SMLT
  - **This feature can be enabled by first disabling SPBM and then enabling simplified virtual-IST at the Interface MLT level
    - VSPSwitch(config)#***no boot config flags spbm-config-mode***
    - VSPSwitch(config)#***interface mlt <mlt id>***
    - VSPSwitch(config-if)#***virtual-ist enable***

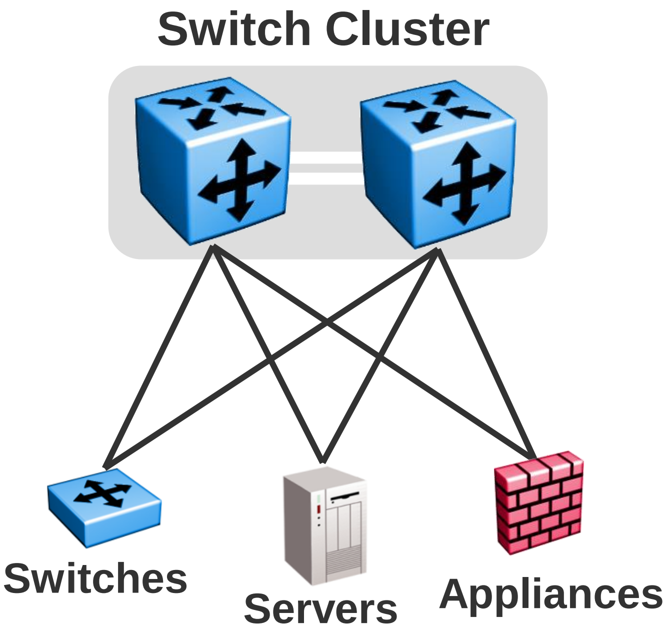
Feature Availability	**spbm-config-mode = ENABLED**	**spbm-config-mode = DISABLED**
SPBM Provisioning	☑	☒
CFM Provisioning (SPBM B-VLAN)	☑	☒
IGMP v1/v2/v3	☑	☑
Multicast over SPBm	☑	☒
PIM-SM, PIM-SSM	☒	☑
Simplified Virtual IST Configuration	☒	☑
All other features	☑	☑

**Note: You still have to configure an IP address for the IST and add the vIST peer

How Do I Build a Switch Cluster ?

Logical Design

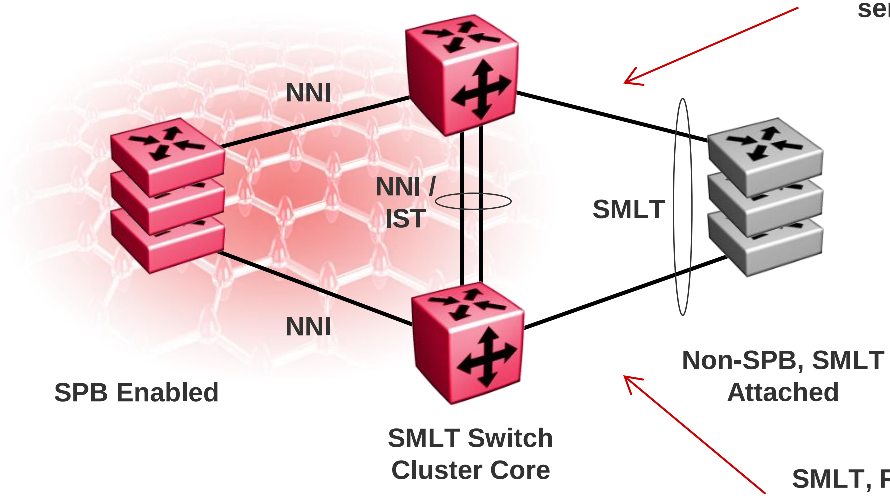
- **Default Gateway Redundancy**
  - Virtual Router Redundancy (VRRP)
    - Backup Master enhancement (active/active)
  - Routed Split MultiLink Trunking (RSMLT-Edge)
    - ERS 8800/8600 / ERS 8300 / VSP 9000 / VSP 8000 / VSP 4000 / VSP 7200
- **Simple Loop Prevention Protocol (SLPP)**
  - Prevents loops in Switch Cluster networks
  - Disables uplink port where loop is detected
  - Enabled on Access SMLT/SLT ports – disabled on IST or Simplified Virtual IST MLT
- **Virtual LACP (VLACP)**
  - Lightweight protocol for end-to-end health check
  - Detect end-to-end failure by propagating link status between ports that are either
    - Physically connected point-to-point
    - Logically connected point-to-point across an intermediate network
  - Does not perform link aggregation functions



The "Simple" Aspect Includes the Configuration

SMLT in Core

- SMLT in the core between SMLT cluster switches is only required for traditional protocols
- If SPB is used, SMLT is only required to connect to non-SPB switches, servers, or other appliances if active/active redundancy is required
- Only supported with the ERS 8000 and VSP 9000
  - SPB with SMLT/RSMLT in core is not supported with the VSP 4000 / 7200 / 8000
    - Traditional OSPF can be used with OSPF enabled on each link



Use SMLT to attach to non-SPB switches, servers, or other devices requiring active/active resiliency

SMLT, RSMLT, RSMLT Edge, and VRRP with Backup Master is supported

MLT Scaling and ID Recommendations

Switch Model	Links per MLT Group	MLT Groups per Switch /Stack	MLT-based SMLT Groups				Port-based SLT Groups			
			Copper	Fiber (1GbE)	Fiber (10GbE)	Fiber (40GbE)	Copper	Fiber (1GbE)	Fiber (10GbE)	Fiber (40GbE)
VSP 9000	16	512	511	511	511	511	N/A	N/A	N/A	N/A
VSP 8000	8	84	84	84	84	84	N/A	N/A	N/A	N/A
VSP 4000	8	24	24	24	24	N/A	N/A	N/A	N/A	N/A
VSP 7200	8	84	84	84	84	84	N/A	N/A	N/A	N/A
VSP 7000	8	64	63 (10.3.1)	63	63	63	256	256	256	16
ERS 8800	8	128	127	127	127	N/A	382	238	22	N/A
ERS 8300	4	31	30	30	30	N/A	382	398	67	N/A
ERS 5000	8	32	31	31	31	N/A	398	190	62	N/A

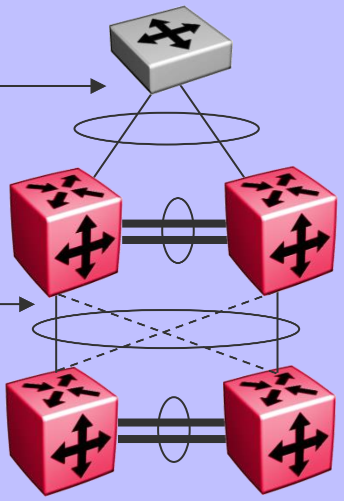
Switch Model	Software Version for SMLT	MLT-based SMLT ID's	Port-based SLT ID's
VSP 9000	3.0 and higher	1-512	MLT-ID's 1-512
VSP 8000	4.0 and higher	1-512	MLT-ID's 1-512
VSP 4000	4.1 and higher	1-512	MLT-ID's 1-512
VSP 7200	4.2.1 and higher	1-512	MLT-ID's 1-512
VSP 7000	10.2 and higher	1-64	65-512
ERS 8800	4.1 and higher	1-128	129-512
ERS 8300	3.0 and higher	1-31	32-512
ERS 5000	5.0 and higher	1-32	33-512

Note: Prior to release 10.3 for the VSP 7000, up to 32 MLT groups are supported

Switch Clustering

Terminology – RSMLT

RSMLT – Routed Split MultiLink Trunking
"Standard" Core Routing Layer 3 Design or Routing with Layer 3 Edge



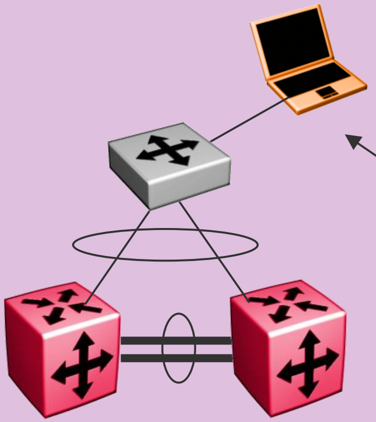
Routed Connection Between Edge and Core

Routed Connection Between Switch Clusters

Sub-second failover without modifying any layer 3 protocols or timers

Support for square or full mesh designs

RSMLT Edge – Routed Split MultiLink Trunking Replacement for VRRP with Layer 2 Edge



Provides Default Gateway for End Stations

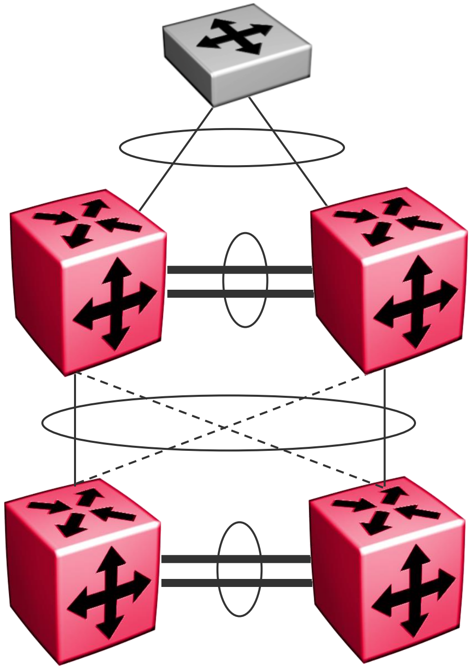
Scales beyond maximum number of VRRP instances

Do not run VRRP and RSMLT on the same edge VLAN simultaneously

Avaya Switch Cluster Core

Routed SMLT (RSMLT) Solution

- Unparalleled Layer 3 resiliency
- End to end sub-second failover for routed VLAN traffic
- RSMLT will take care of resiliency
- Perfect complement to SMLT/SLT from the edge



- All routers in the core VLAN run standard IGP such as RIP or OSPF
- No tuning of IGP necessary
- No VRRP, ECMP required in core VLANs
- Multiple routed RSMLT Core VLANs supported per switch

- Support for square and full mesh topologies in the core
- Support for triangles to the edge – when used as a VRRP replacement or Layer 3 edge

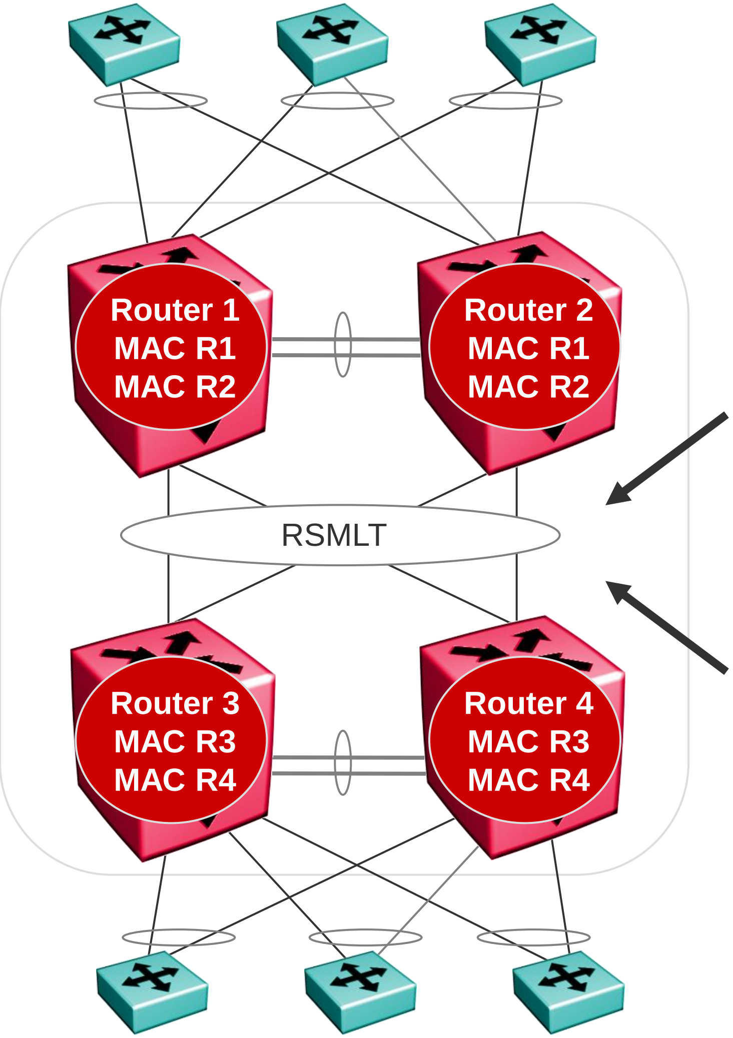
Sub-second Failover in Layer 3 Environments

LAN Solution Example

Routed SMLT (RSMLT) Operation

Normal Operation

- RSMLT aggregation switch pairs exchange local router MAC addresses
- Both router MAC addresses are routing packets on both RSMLT aggregation switches
- SMLT/RSMLT ensures that no packets are duplicated



Core VLAN between all four routers

- RSMLT enabled on core VLAN(s)
- All cluster switches in the same IP subnet
- Routing protocol enabled
- No ECMP / VRRP required

Any IGP

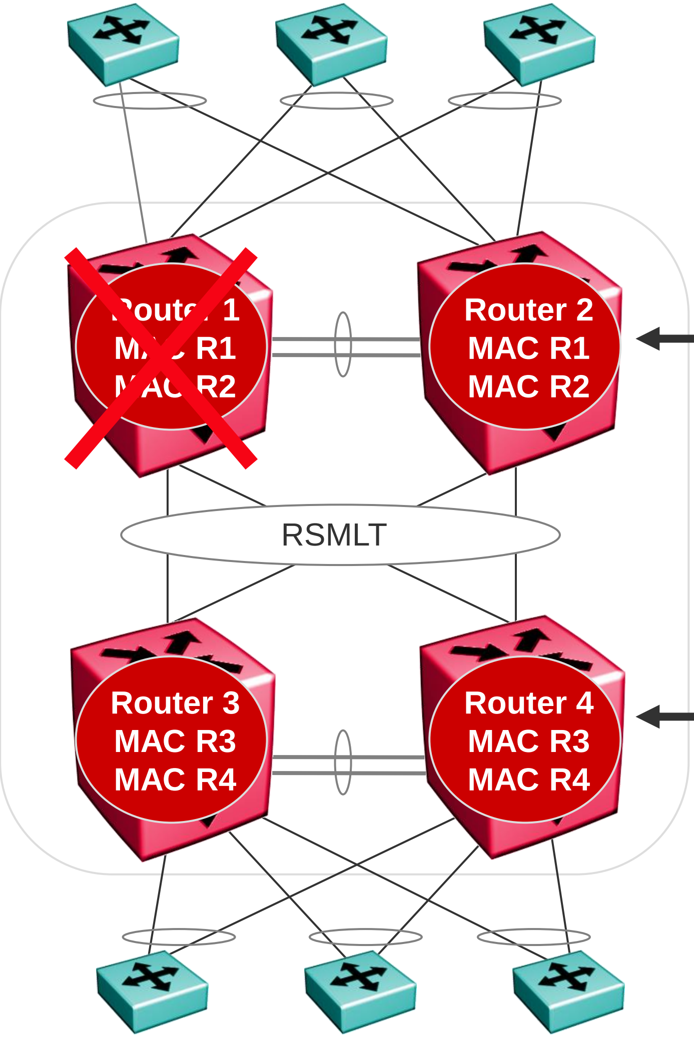
- OSPF or RIP
- IGP knows nothing of RSMLT
- Protocol timers unchanged

LAN Solution Example

Routed SMLT (RSMLT) Operation

Failure Operation

- Routing protocol converges just as normal
- Router 2 continues to forward for Router 1
- After routing protocol is converged MAC R1 is flushed



Hold-up timer kicks in

- Keeps MAC R1 alive for timer duration
- Timer should be greater than IGP convergence time

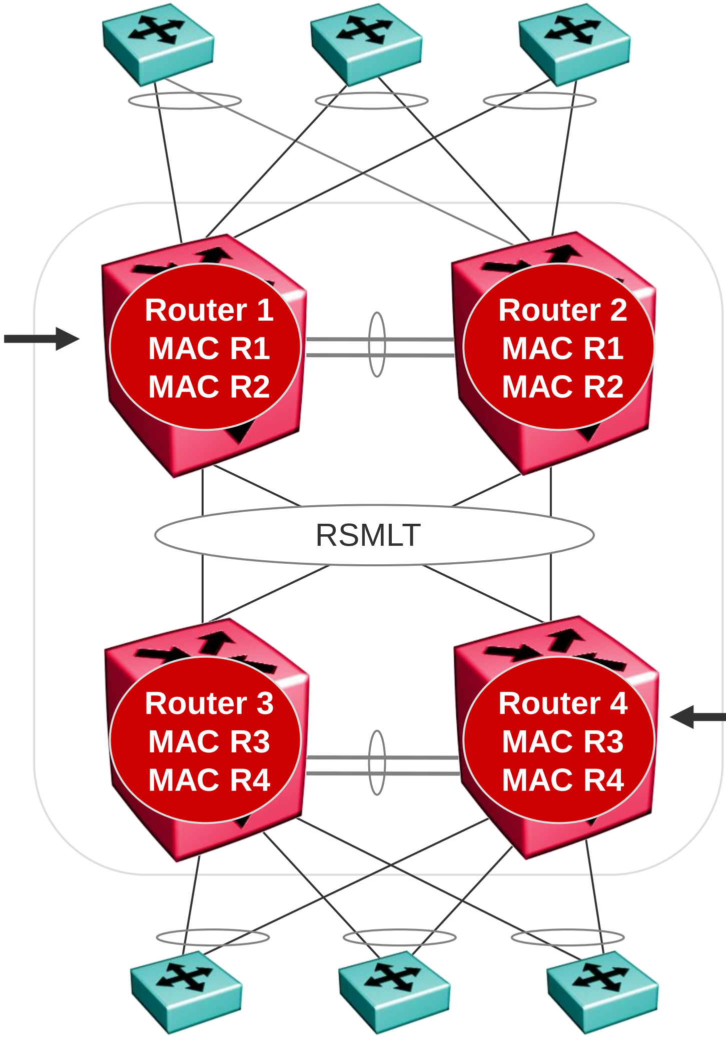
Router 3 & Router 4 continue to forward without any service disruption

LAN Solution Example

Routed SMLT (RSMLT) Operation

Hold-down timer kicks in

- Will not start forwarding traffic until timer expires
- Timer should be greater than IGP convergence time



Recovery of Router 1

- Routing protocol converges just as normal
- After routing protocol is converged Router 1 and Router 2 begins to forward for MAC R1 and MAC R2

Router 3 & Router 4 continue to forward without any service disruption

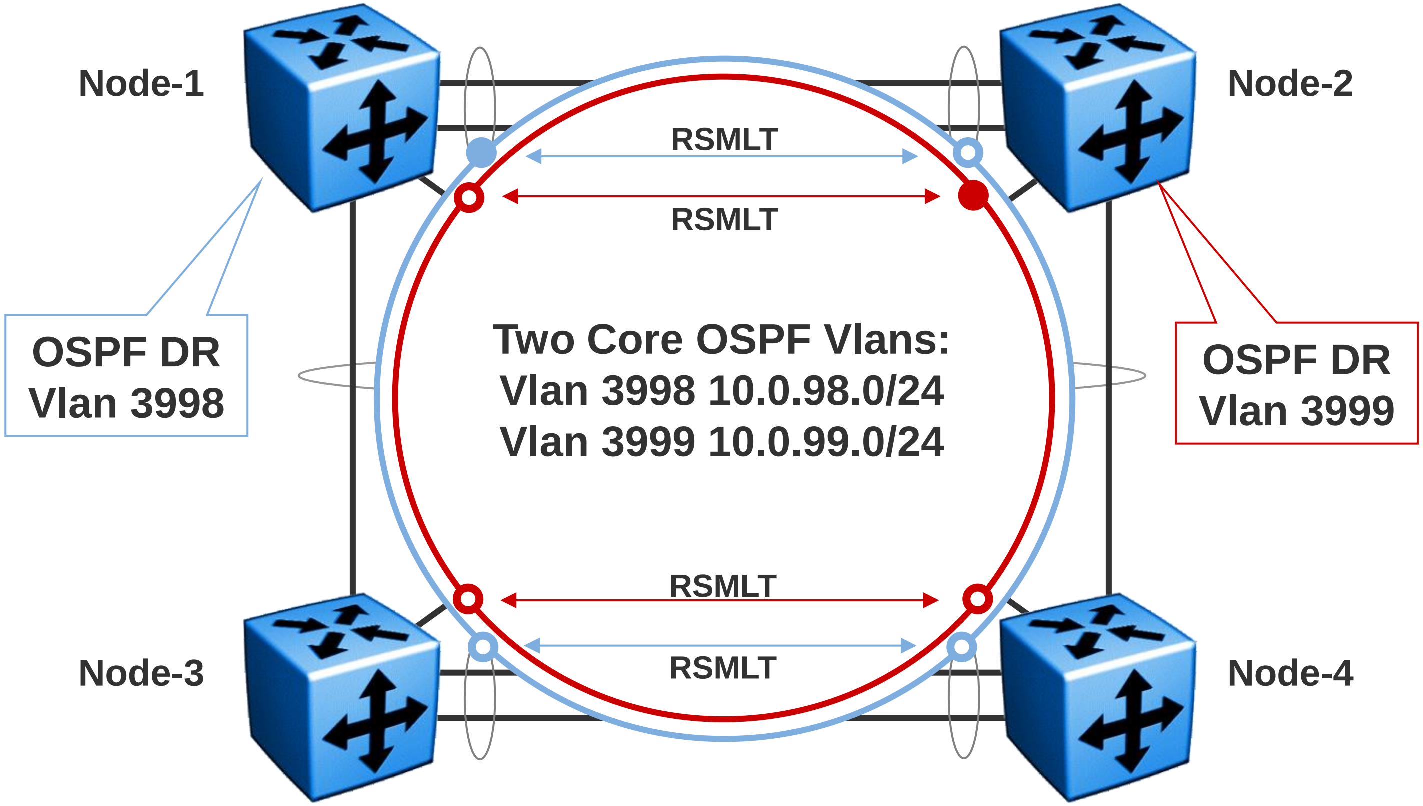
RSMLT with Dual Core OSPF VLANs

Enhancing Availability for ERS 8000

- Issue
  - In the event of losing the DR (designated router), two SPF runs are necessary to restore routes across the segment. However, the OSPF holddown timer will not allow 2 consecutive SPF runs to occur within the value of the timer. That timer defaults to 10secs on the ERS 8800/8600. So every time the DR is lost, a traffic interruption of up to 10secs can occur.
- Solution
  - Create a second OSPF Core VLAN and force different nodes to become the DR for each VLAN.
  - Each OSPF Core VLAN will have DR (set priority to 100) and no BDRs (set OSPF priority to 0 on all routers/switches not intended to become the DR).
  - No BDR is necessary since the two VLANs back each other up from a routing perspective.
- Note
  - This solution does not support multicast
  - If multicast is required, use only one OSPF enabled VLAN in the core

RSMLT with Dual Core OSPF VLANs

Enhancing Availability for ERS 8000



Node-1

Node-2

Node-3

Node-4

RSMLT

RSMLT

RSMLT

RSMLT

OSPF DR
Vlan 3998

OSPF DR
Vlan 3999

Two Core OSPF Vlans:
Vlan 3998 10.0.98.0/24
Vlan 3999 10.0.99.0/24

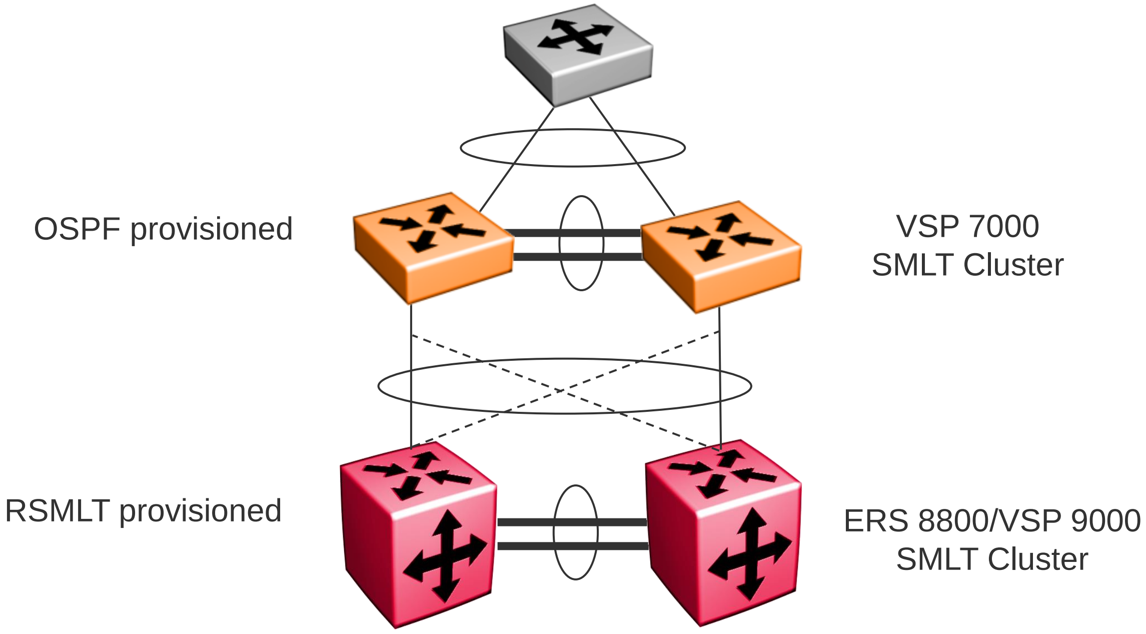
● ● IP interface on VLAN with **OSPF priority = 100**

○ ○ IP interface on VLAN with **OSPF priority = 0**

For the VSP 9000, only one VLAN is required. If Node-1 & Node-2 are VSP 9000 switches and Node-3 and Node-4 are ERS 8000 switches, use only one VLAN and set the OSPF priority on the ERS 8000 switches to 0 and 100 on the VSP 9000 switches.

VSP 7000 - OSPF Routing in SMLT Square

- The VSP 7000 supports OSPF routing in a SMLT square topology



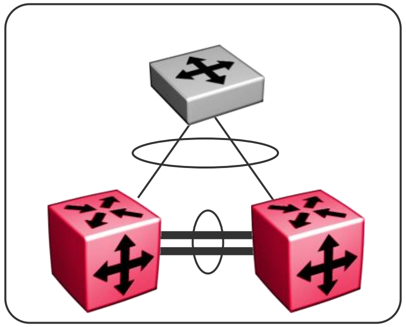
- Although RSMLT is not supported on the VSP 7000, OSPF can be provisioned in a SMLT square topology
  - RSMLT can only be provisioned if the cluster is either a ERS 8000 or VSP 9000/8000 SMLT cluster

Switch Clustering

Supported Topologies

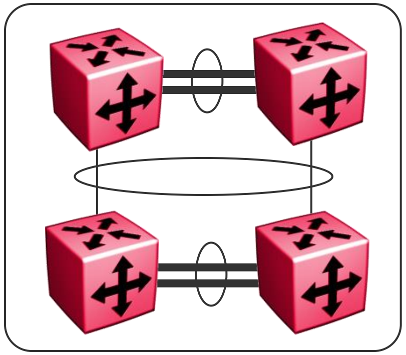
Triangle

- Single Switch Cluster at the core with edge directly connected



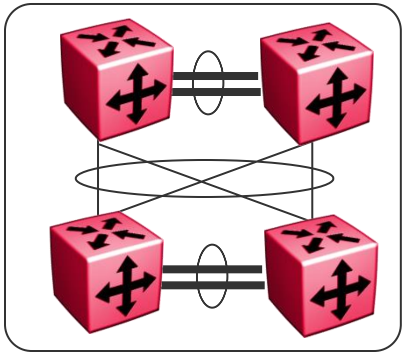
Square

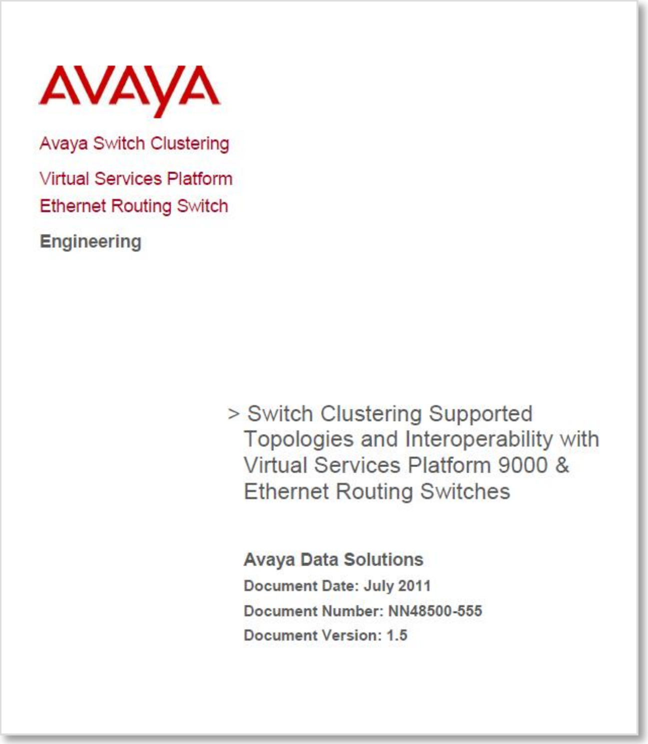
- Two pairs of Switch Clusters interconnected by SMLT. Squares can be scaled with additional pairs of Switch Clusters



Full Mesh

- Expanding on the Square topology, the full mesh adds additional connections between the pairs so that each switch has at least one connection to every other switch in the square





AVAYA

Avaya Switch Clustering

Virtual Services Platform
Ethernet Routing Switch

Engineering

> Switch Clustering Supported Topologies and Interoperability with Virtual Services Platform 9000 & Ethernet Routing Switches

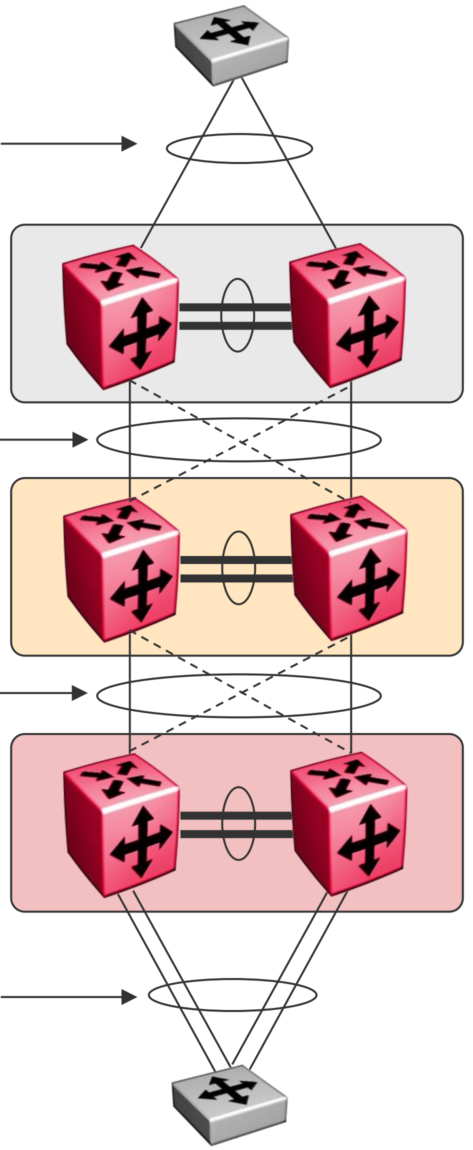
Avaya Data Solutions

Document Date: July 2011

Document Number: NN48500-555

Document Version: 1.5

Switch Clustering Solution



Access SLT
(Triangle)

Core SMLT Layer 2
Core RSMLT Layer 3
(Square or Full Mesh)

Core SMLT Layer 2
Core RSMLT Layer 3
(Square or Full Mesh)

Access SMLT
(Triangle)

Switch Cluster (SC1)

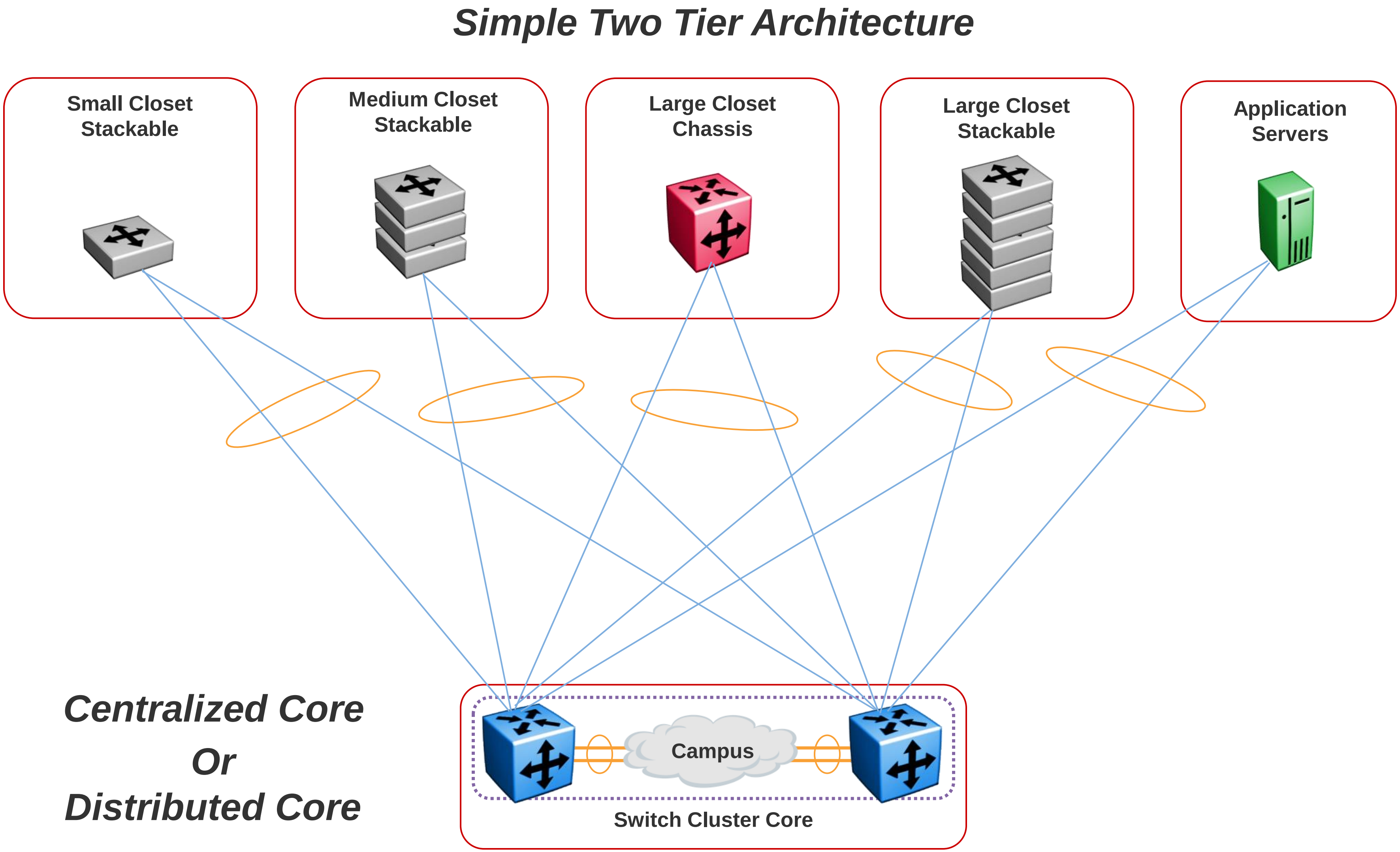
Switch Cluster (SC2)

Switch Cluster (SC3)

Cascaded Switch Clusters

Network Design Flexibility

Reducing OPEX



Centralized Core

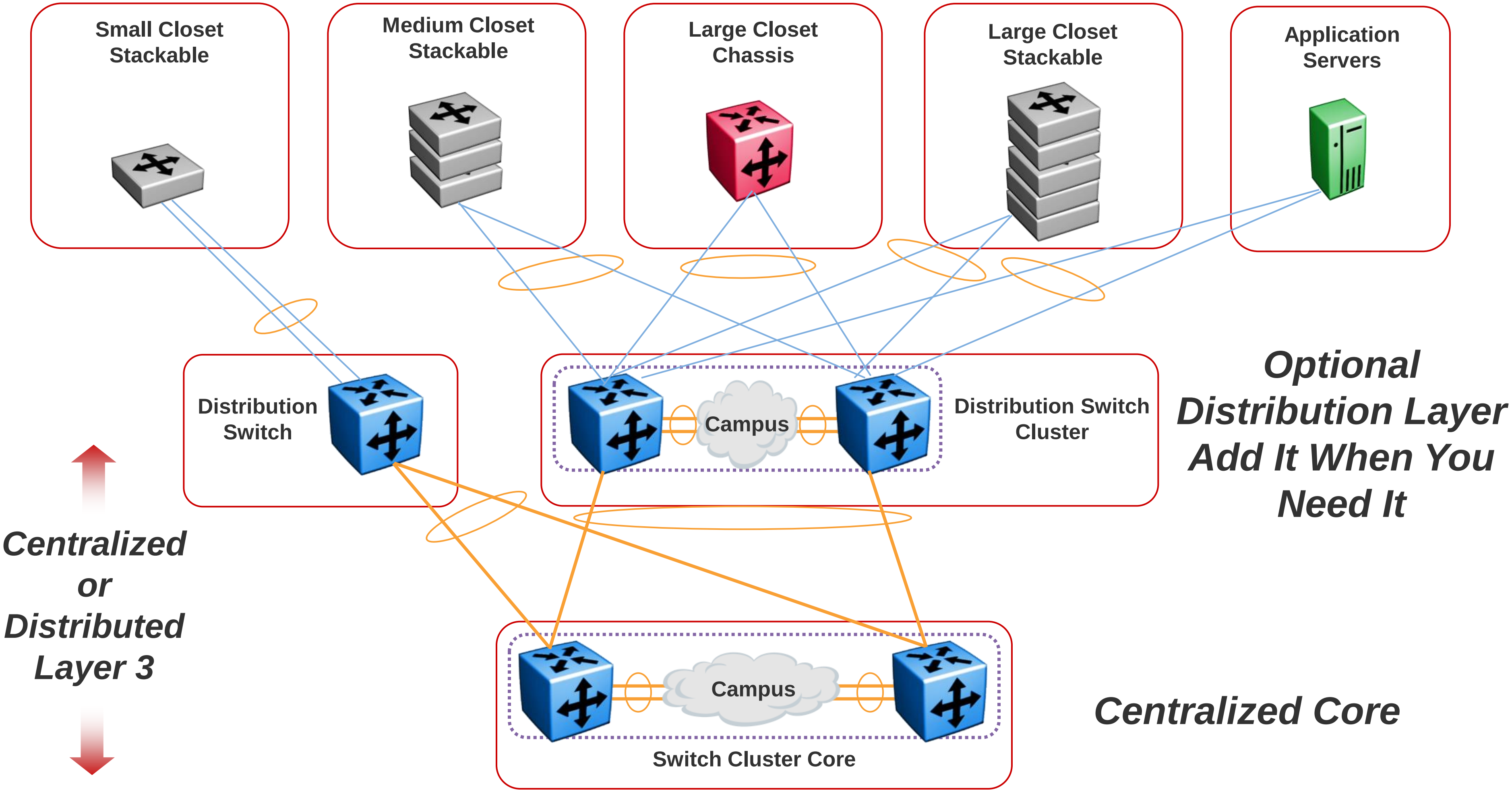
Or

Distributed Core

Network Design Flexibility

Reducing OPEX

Three Tier Architecture



Optional Distribution Layer Add It When You Need It

Centralized or Distributed Layer 3

Centralized Core

SMLT/RSMLT support for IPv6

SMLT/RSMLT support for IPv6 Overview

Std. Compliance

Avaya Differentiator

Value Statements

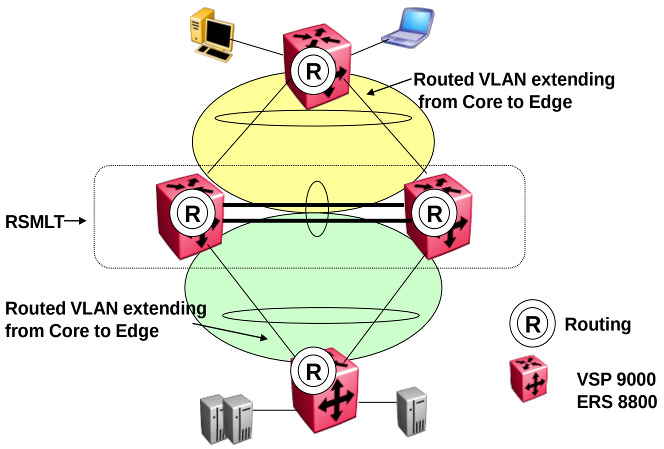
- L3 resiliency with sub 20 msec failover for IPv6
- Simple & proven technology
- Supports Dual stack (IPv4/IPV6) topologies
- Transition or overlay IPv6 network with existing network infrastructure

Requirements

- VSP Release 3.2
- ERS 8800 Release 7.1
- Advanced Software License

Description:

Layer 3 Edge and Layer 3 Core with RSMLTv6


Routed VLAN extending from Core to Edge
RSMLT
Routed VLAN extending from Core to Edge
R
Routing
VSP 9000
ERS 8800

SMLT/RSMLTv6:

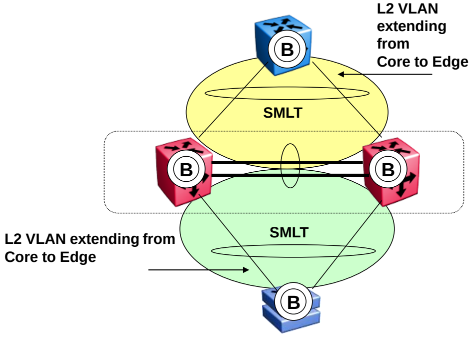
- SMLT (Split Multi-Link Trunking) & RSMLT (Routed SMLT) extended to support IPv6 over L2/L3 network
- With RSMLT enabled an SMLT switch performs IP forwarding on behalf of its SMLT peer – thus preventing IP traffic from being sent over the IST
- It supports IPv6 in Triangular, Square and Full mesh topologies
- RSMLT peers exchange their IPv6 configuration (MAC address, IPv6 addresses and prefixes) and track each other's state by means of IST messages
- When RSMLT node receives an IPv6 packet destined to peer's MAC address it will perform IPv6 forwarding on the packet
- RSMLT VLAN supports IPv4-only, IPv6-only or IPv4+IPv6 operation
- If a VLAN has both IPv4 and IPv6: enabling/disabling RSMLT, starting/stopping forwarding for peer, operation of HoldUp and HoldDown timers apply simultaneously to IPv4 and IPv6 forwarding

RSMLTv6 Configuration

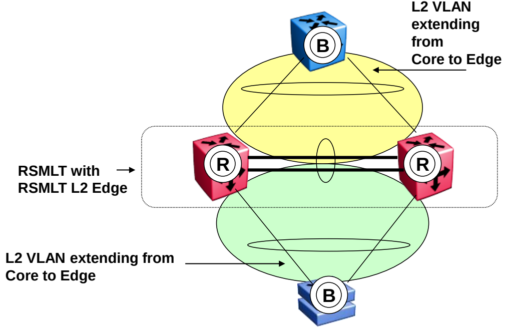
- As with IPv4, it is assumed that RSMLT peers configure the same set of IPv6 prefixes (equivalent of IPv4 subnet) on the RSMLT VLAN
- It is not recommended to enable both VRRP and RSMLT on the same VLAN
- It is not recommended to enable sending IPv6 redirect messages on RSMLT VLANs
- This feature does not introduce any new configuration or show commands
- Current (IPv4) configuration and show commands apply both to IPv4 and IPv6

SMLT/RSMLTv6 Network Topologies

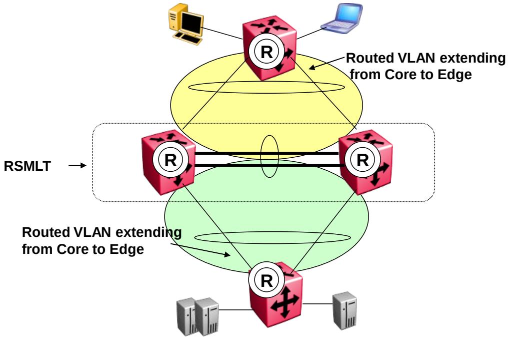
#1 IPv6 Triangle SMLT: Layer 2 Edge and Layer 2 Core


B
L2 VLAN extending from Core to Edge
SMLT
B
B
SMLT
L2 VLAN extending from Core to Edge
B

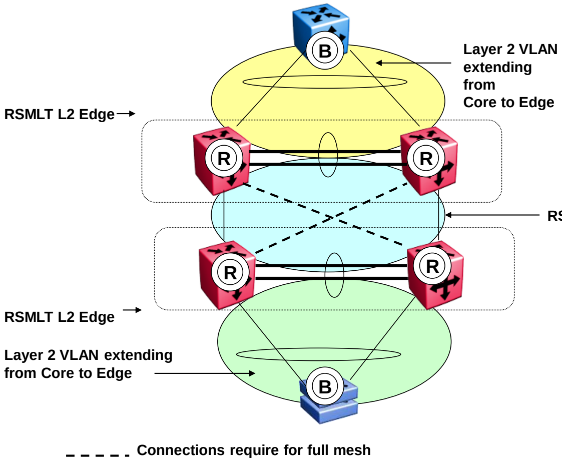
#2 IPv6 Triangle SMLT: Layer 2 Edge and Layer 3 Core with RSMLT L2 Edge


B
L2 VLAN extending from Core to Edge
RSMLT with RSMLT L2 Edge
R
R
L2 VLAN extending from Core to Edge
B

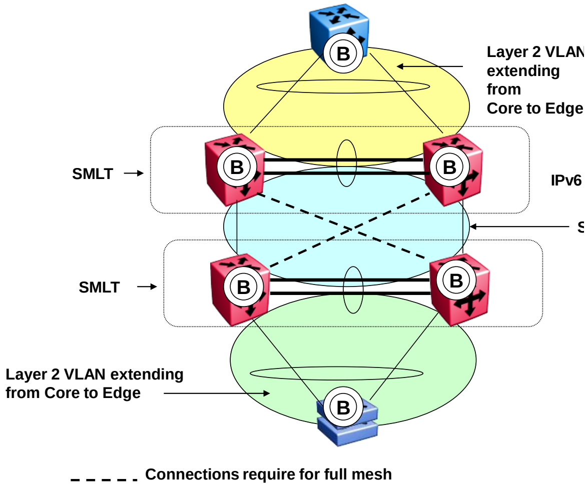
#3 IPv6 RSMLT: Layer 3 Edge and Layer 3 Core with RSMLT


R
Routed VLAN extending from Core to Edge
RSMLT
R
R
Routed VLAN extending from Core to Edge
R

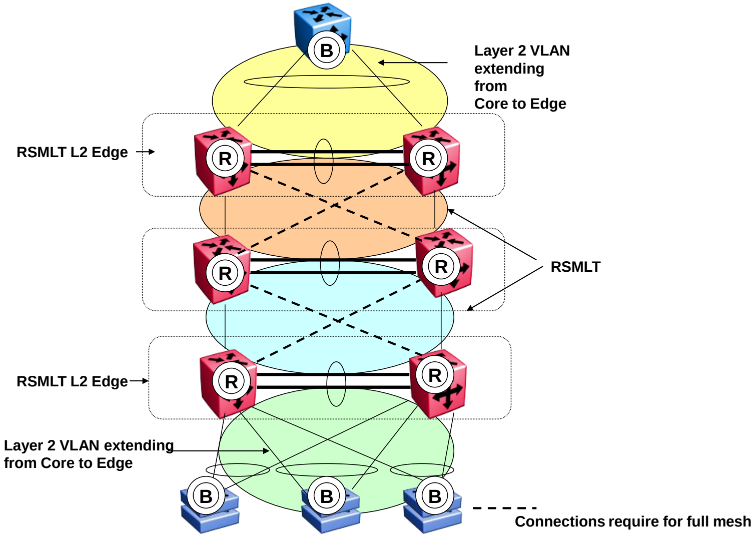
#4 IPv6 Square/Full Mesh RSMLT: Layer 2 Edge and Layer 3 Core


B
Layer 2 VLAN extending from Core to Edge
RSMLT L2 Edge
R
R
RS
R
R
RSMLT L2 Edge
Layer 2 VLAN extending from Core to Edge
B
Connections require for full mesh

#5 IPv6 Square/Full Mesh SMLT: Layer 2 Edge and Core


B
Layer 2 VLAN extending from Core to Edge
SMLT
B
B
IPv6
S
SMLT
B
B
Layer 2 VLAN extending from Core to Edge
B
Connections require for full mesh

#6 IPv6 Square/Full Mesh: Layer 2 Edge and Layer 3 RSMLT Core


B
Layer 2 VLAN extending from Core to Edge
RSMLT L2 Edge
R
R
R
R
RSMLT
RSMLT L2 Edge
R
R
Layer 2 VLAN extending from Core to Edge
B
B
B
Connections require for full mesh

Ⓑ Bridging

Ⓡ Routing

VSP9000
ERS8800
(only in pairs)

ERS8xxx

Stackable Switch

IPv6 VRRP Overview

Std. Compliance

draft-ietf-vrrp-ipv6-spec-08.txt

Value Statements

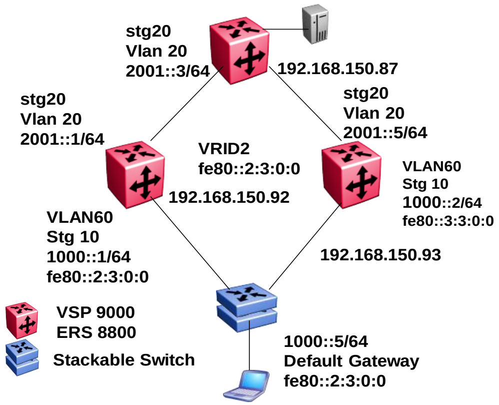
- Faster failover for IPv6 traffic
- Backup-Master VRRP will improve Layer 3 resiliency in SMLT environment

Requirements

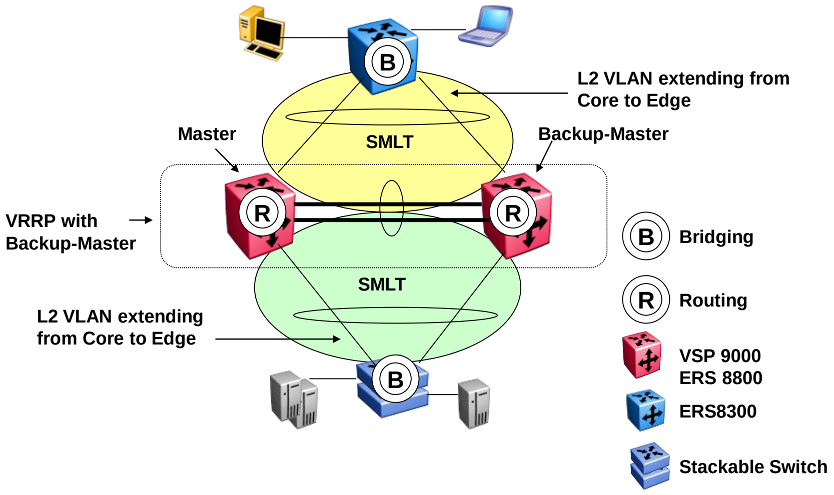
- VSP Release 3.2
- ERS 8800 Release 7.1
- Advanced Software License

Description:

IPv6 VRRPv3 in non-SMLT



IPv6 VRRP Backup-Master with SMLT



VRRP:

- VRRP provides faster failover of default router to the IPv6 LAN hosts
- Once default router is learned through neighbor discovery protocol, unicast neighbor solicitation messages are used to detect the failure of default router
- In addition, an extension to the VRRP VRRP backup master concept improves the layer 3 capabilities of VRRP in conjunction with SMLT
- VRRP for IPv4 and IPv6 can be configured on same VLAN

- Two different VRRP MAC's used for both the versions.
  - 00-00-5e-00-01-<vrId> is used for VRRPv2 &
  - 00-00-5e-00-02-<vrId> is used for IPv6 VRRP
- Using BackupMaster on the SMLT aggregation switch, the backup VRRP switch also routes traffic if it has a destination routing table entry

IPv6 VRRP Configuration

- Enabling VRRP and RSMLT on the same VLAN is restricted
- Assigning link-local address to the VRRP is mandatory to participate in VRRP election procedure
  - Assigning global address to VRRP is optional
- The VRRP address(es) between the VRRP peers should be in Sync (which includes both link-local & global address)
- VRRP can be configured only on Triangular-SMLT topology and it should be configured on aggregation switches
- Always recommended to enable VRRP in SMLT scenario with backup-master enabled

Switch Clustering Design Best Practices

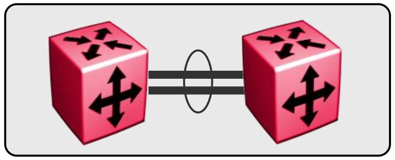
Disclaimer:

The recommendations provided here are based on large scale network testing validated by Avaya. Using values and timers outside of these recommendations are permitted at the user’s own risk based on specific network design requirements. If issues are encountered when running outside of these recommendations, the first step will be to move to recommended values before pursuing further.

Creating the Switch Cluster

Configuration Recommendations

- The IST should be a distributed MLT for added resiliency between the Switch Cluster Cores
  - The DMLT links that comprise the IST must be of the same speed
  - Number and speed of links for the IST is determined by the amount of traffic that would use the IST during a failed condition – during normal operation, the majority of traffic across the IST is from devices that are not dual-homed
  - Use a private address space with 30 bit mask for IST IP's
  - Do NOT use the IST IP addresses as the next hop address for any static routes
  - No routing protocol should be enabled on IST IP interfaces with the exception of PIM for IP Multicast routing which benefits in reconvergence times if enabled on the IST interfaces
  - The IST IP interfaces are only locally significant to the SMLT Cluster to which they belong. Hence the same IP addresses can be re-utilized across each and every SMLT Cluster

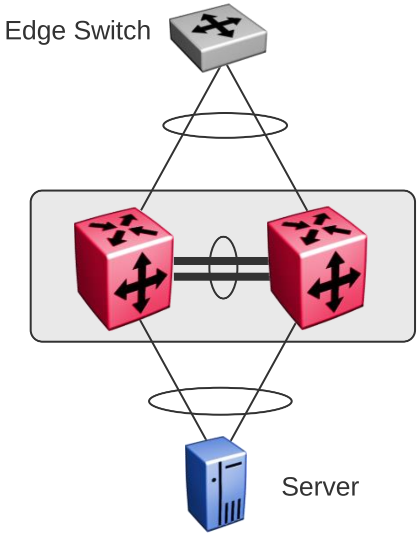


- All VLANs that span SMLT/SLT connections must be tagged across the IST
  - Except on ERS8600/8800 or VSP 7000 if those VLANs are L2VSNs (i.e. have an I-SID assigned); in this case the VLAN must not be tagged on the IST connection as it is already transported by SPB
    - This exception does NOT apply to the VSP9000

Link Aggregation

Connecting the Edge to the Switch Cluster

- Switch Cluster is agnostic of Edge devices
  - Must enable a form of link aggregation on the Edge
- VSP 9000, VSP 8000, VSP 7200, VSP 4000, ERS 8800/8600 or ERS 8300 Switch Cluster
  - Edge devices can use 802.1AX or static link aggregation
- ERS 5000 & VSP 7000 Switch Cluster
  - Edge devices can use static link aggregation
  - 802.1AX over SMLT/SLT is supported as of 6.2 for the ERS 5000 and 10.2 for the VSP 7000
    - Requires at least two ports per unit of an IST stack for SMLT
      - Supported with SLT via one port
    - 802.1AX is not supported on the IST
  - IGMP over SMLT/SLT is supported as of 6.3 for the ERS 5000 and 10.3 for the VSP 7000
- Server Connectivity
  - VSP 9000, VSP 8000, VSP 4000, ERS 8800/8600, ERS 8300, VSP 7000, and ERS 5000 support both 802.1AX and static link aggregation over SMLT/SLT
    - When using 802.1AX over SMLT, enable:
      - LACP SMLT-SYS-ID
  - For details on NIC teaming and Avaya Switch Clustering, refer to Data Center Server Access Solution Guide (NN48500-577)



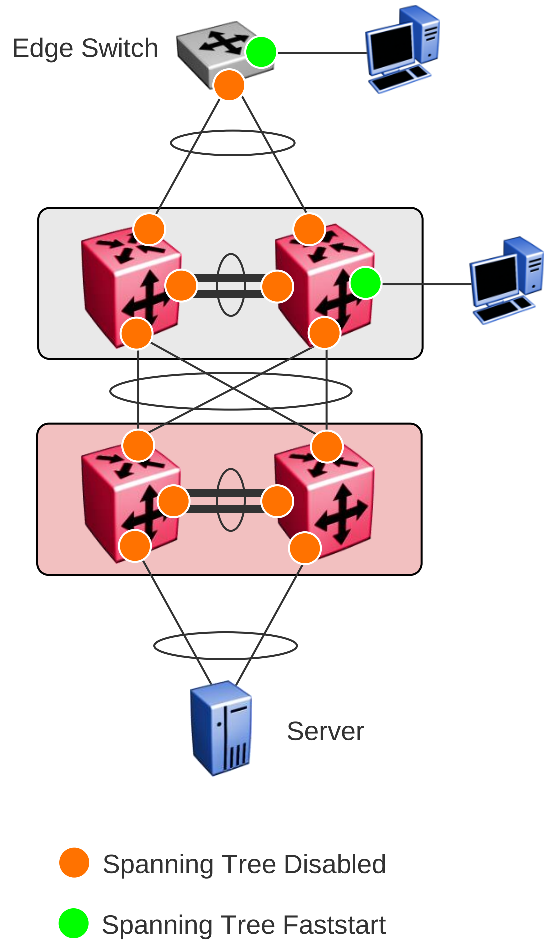
Switch Clustering Best Practices

Configuration Checklist

☑ **Spanning Tree**

Disabled on uplink and IST ports – Core
Disabled on uplink ports – Edge
FastStart / EdgePort enabled on all other ports

- Spanning Tree – Core
  - Globally enabled on all Switches (any mode: STPG, RSTP, MSTP)
  - Must be disabled on IST and Uplink ports
    - STP is automatically disabled on VSP 9000/7000/8000/4000 and ERS 8800/8600/8300/5000 uplink and IST ports
- Spanning Tree – Edge
  - Globally enabled on all Switches
  - Must be disabled on the uplink ports forming the MLT
  - Spanning Tree mode (STPG/RSTP/MSTP) should match Core's
    - i.e. if core is VSP 9000, VSP 8000, or VSP 4000 use MSTP to allow for ease of management via tools such as VLAN Manager in COM
- Spanning Tree FastStart (or RSTP/MSTP EdgePort)
  - Enable on all non-uplink and non-IST ports for added protection
  - Allows ports to come up in forwarding state immediately
  - Spanning Tree will still block if a loop is detected via BPDUs
  - Allows BPDU-Filtering protection to be effective against edge loops



BPDU Filtering

Feature Overview

- Allows the network administrator to achieve the following:
    - Block an unwanted root selection process when an edge device is added to the network. This prevents unknown devices from influencing an existing spanning tree topology.
    - Block the flooding of BPDUs from an unknown device
- When a port has BPDU-Filtering enabled and it receives an STP BPDU, the following actions take place:
    - The port is immediately put in the operational disabled state
    - A trap is generated and the following log message is written to the log:
        - BPDU received on port with BPDU-Filtering enabled Port <x> has been disabled
    - The port timer starts
    - The port stays in the operational disabled state until the port timer expires
    - If the timer is set to 0, then the port remains disabled

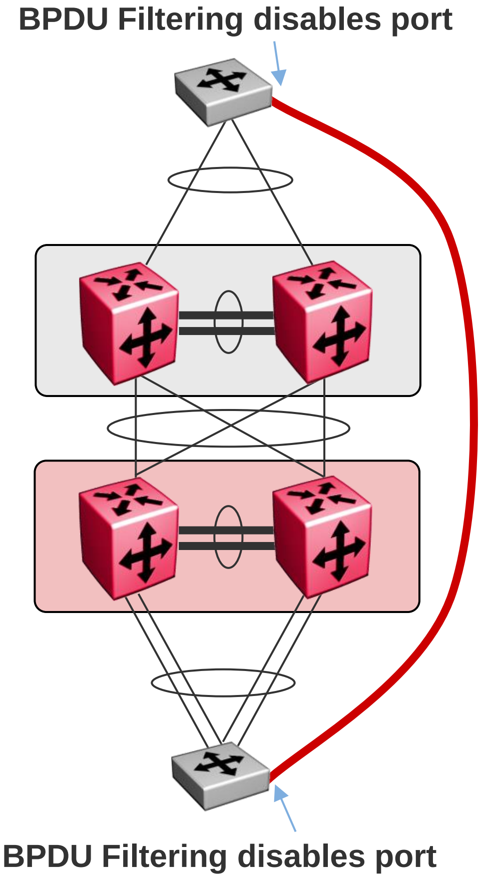
Switch Clustering Best Practices

Configuration Checklist

☑ **BPDU Filtering**

Enabled on Edge access ports

- Enable on all Edge access ports
- Set timeout to 0
  - Port remains disabled until manual intervention to re-enable it
- BPDU Filtering is not supported on MultiLink Trunk (MLT) ports
- Supported on
  - ERS 2500 with Release 4.2
  - ERS 3500 with Release 5.0
  - ERS 4000 with Release 5.1
  - ERS 5500 with Release 5.1
  - ERS 5600 with Release 6.0
  - ERS 8300 with Release 4.2
  - ERS 8600/8800 with Release 7.0
  - VSP 7000 with Release 10.1
  - VSP 9000 with Release 3.4

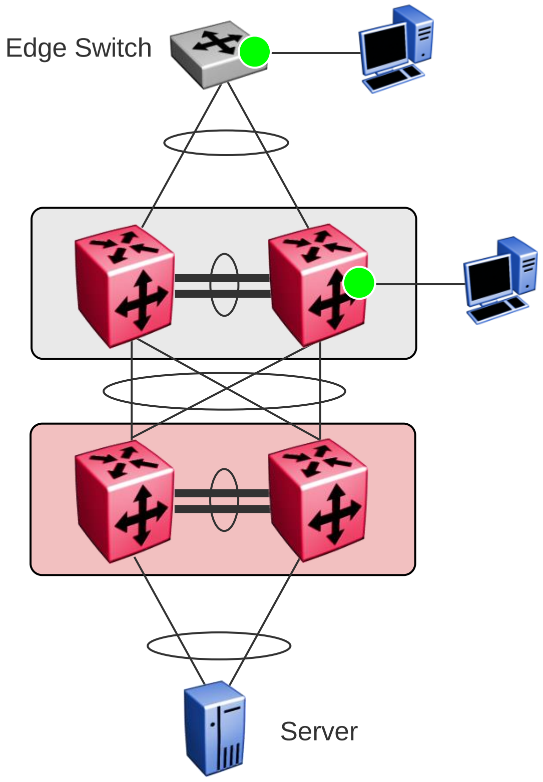


Switch Clustering Best Practices

Configuration Checklist

☑ **Autonegotiation enabled on all ports**

Can use CANA on access ports if required



● CANA Configured Ports

- Autonegotiation
  - Leave enabled on all GbE & 100FX ports
  - Ensures RFI (remote fault identification)
  - Autonegotiation does not exist for 10Gig or 40Gig – RFI is already built-in
- Custom Autonegotiation Advertisements (CANA)
  - Provides capability to set autonegotiation advertisements for speed and duplex
  - Useful to limit end station connectivity bandwidth
  - Supported on:
    - ERS 2500
    - ERS 3500
    - ERS 4000
    - ERS 5000
    - VSP 7200
    - ERS 8300
    - ERS 8800/8600 (R & RS modules only)
    - VSP 9000
    - VSP 7000
    - VSP 8000

Switch Clustering Best Practices

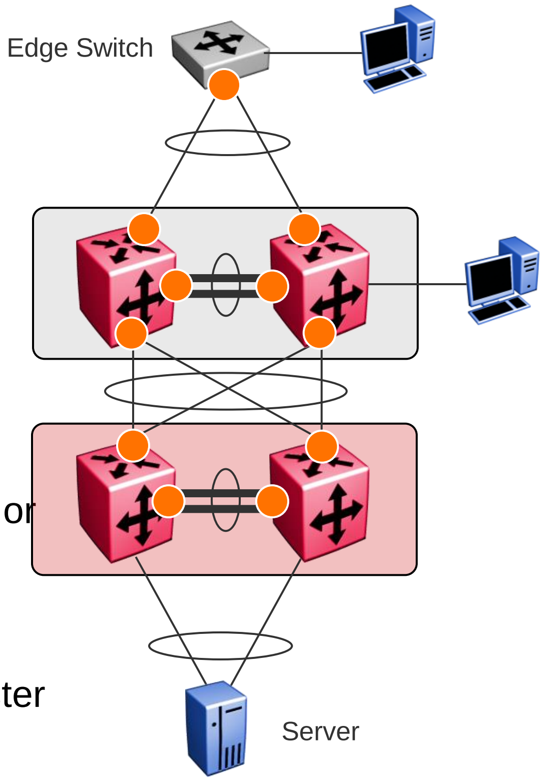
Configuration Checklist

☑ 802.1Q enabled on uplink / IST ports

☑ Discard Untagged Frames enabled on uplink / IST ports

☑ Increase FDB Timer on Switch Cluster Core VLANs

- 802.1Q VLAN Tagging
  - Enable 802.1Q on uplink ports even if only one VLAN at the Edge
  - Facilitates ease of adding VLANs to the uplinks in the future
  - Allows for use of Discard Untagged Frames on uplinks
- Discard Untagged Frames
  - Any untagged packets received on the port will be dropped at ingress
  - Protection mechanism against Edge switch reset to factory default or a new switch not configured properly being added to the network and possibly causing a loop in the network
  - Do not enable on ERS 5510 when using VLACP
  - Do not enable on the IST ports if using the VSP 7000 as the SMLT cluster
- Increase FDB Timer per VLAN (ERS 8800 / 8600 / ERS 8300)
  - Increase the FDB timer on all Switch Cluster Core VLANs from the default of 300 seconds to 21601 seconds (1 second greater than the ARP timer)
  - Reduces the amount of re-ARPs when the FDB timer for a given MAC ages out
  - Leave FDB timer at default of 300 seconds on ERS 5000 / VSP 7000 / VSP 9000 / VSP 8000 / VSP 4000 Switch Cluster

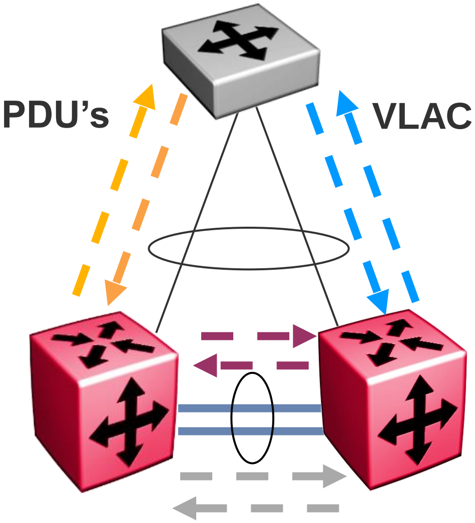


● 802.1Q VLAN Tagging and Discard Untagged Frames

Virtual LACP (VLACP)

Feature Overview

- Virtual LACP (VLACP) = Lightweight LACP
  - Detects end-to-end failure by propagating link status between ports that are:
    - Directly connected point-to-point
    - Logically connected point-to-point across an intermediate network
  - Can detect
    - Complete link failure
    - Receive or transmit link disruptions (one-way link failure)
  - Transmits VLACPDU every “x” milliseconds so both ends of the link maintain state
  - VLACP doesn’t perform Link Aggregation functions
  - Based on LACP but is Intellectual Property of Avaya
  - Supported on:
    - ERS 2500
    - ERS 3500
    - ERS 4000
    - ERS 5000
    - ERS 8300
    - VSP 7200
    - ERS 8600
    - ERS 8800
    - VSP 7000
    - VSP 4000
    - VSP 8000
    - VSP 9000



VLACP-PDU’s

VLACP-PDU’s

VLACP-PDU’s

Default VLACP Settings						
Switch	**VLACP State**	**Slow Timer**	**Fast Timer**	**Default Timer**	Timeout Scale	Default MAC
Stackables VSP 7000	Disable	30000	500	long	3	01:80:c2:00:11:00 (Global)
ERS 8000 VSP 4000 / 7200 / 8000 / 9000	Disable	30000	200	long	3	01:80:c2:00:11:00 (per interface)

Switch Clustering Best Practices

Configuration Checklist

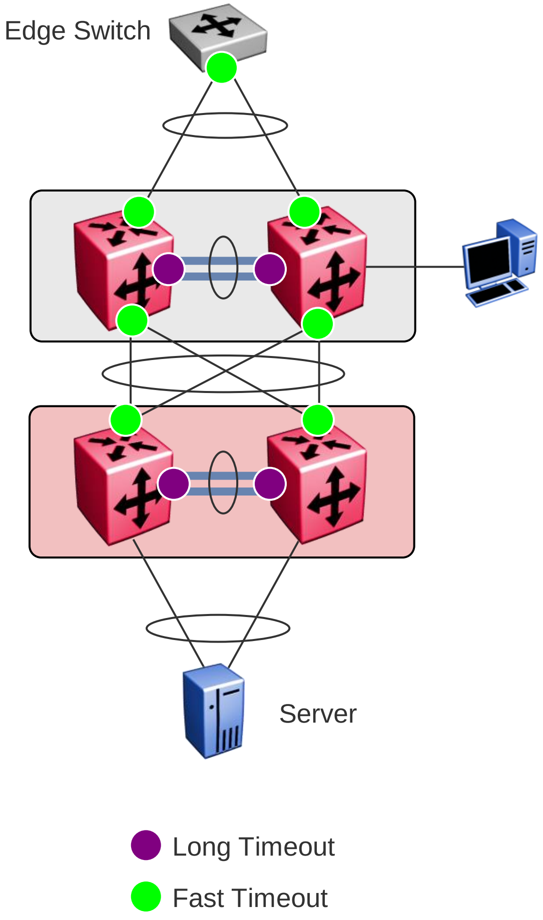
☑ **VLACP enabled on uplink and IST ports**

- Enable VLACP
  - Globally and on each individual uplink and IST port
  - Both ends must have matching Multicast MAC, Ethertype, and Timers
  - Do not enable VLACP and LACP on the same links
  - No not enable VLACP on IST ports members on the VSP 7000
- For directly connected point-to-point links
  - Use reserved multicast MAC 01-80-c2-00-00-0f
  - Ensures packet is not flooded across a defaulted switch
- For end to end connections traversing intermediate networks
  - Use default MAC 01:80:c2:00:11:00

Short Timeout = Timeout Scale * Fast Periodic Timer

Long Timeout = Timeout Scale * Slow Periodic Timer

Connection Type	Fast Timer	Slow Timer	Timeout	Timeout Scale
Uplink	500ms	N/A	Short	5
IST	N/A	10000	Long	3



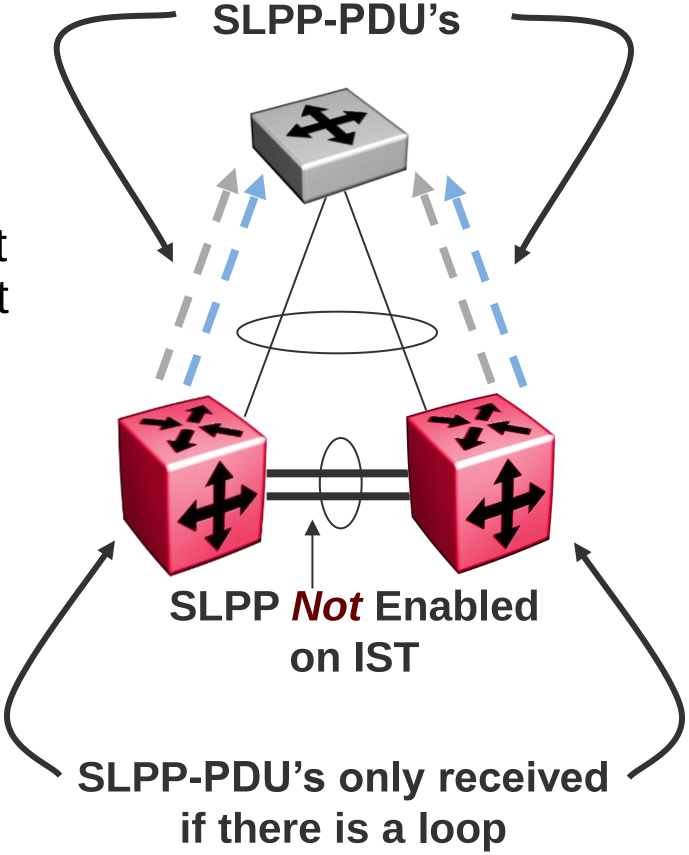
Simple Loop Prevention Protocol (SLPP)

Feature Overview

- Prevents loops in a Switch Cluster network
  - Loops can occur when:
    - MLT at the edge is misconfigured
    - MLT not created at the edge but links are plugged in anyway
    - MLT configuration is lost (switch set back to factory default)
- SLPP uses an SLPP-PDU which is generated by the Switch Cluster cores
  - Loop detection is achieved by detecting whether the SLPP-PDU is received on the IST peer switch port or on the same switch where it originated
  - If a self or SMLT peer originated SLPP PDU packet is received
    - The port is taken down (if the packet is received on the same VLAN it originated on)
    - A log file entry is generated
    - An SNMP trap is sent
  - Once the port is down, it will stay in the down state and need manual intervention to be enabled
  - SLPP PDU transmission is enabled on a per VLAN basis
  - SLPP PDU reception/detection is configured on a per port basis
  - SLPP can be used with LACP between Avaya switches
    - SLPP with LACP between Avaya switches and some servers may have issues where some sever drivers may reflect packet when aggregation is on yet active causing SLPP frames to be detected and therefore taking down a port

Simple Loop Prevention Protocol (SLPP) *Feature Overview*

- Enabling SLPP on a VLAN causes the switch to transmit the multicast SLPP-PDU – the packet is constrained to the VLAN on which it was sent
- The SLPP-PDU receiving and processing works only on ports where SLPP-Rx is enabled
- When SLPP-PDU receiving process works on the port which is a member of an MLT, all port members in that MLT will be taken down
- The SLPP-PDU can be received by the originated Switch or the IST peer Switch. All other switches treat the SLPP-PDU as normal multicast packet and will forward it on the VLAN
- SLPP threshold based on the sum of all packets received
- Port-based VLANs only
- Supported on:
  - VSP 9000/8000/4000/7200/7000(10.2), ERS 8800/8600/8300/5000



SLPP-PDU's

SLPP *Not* Enabled on IST

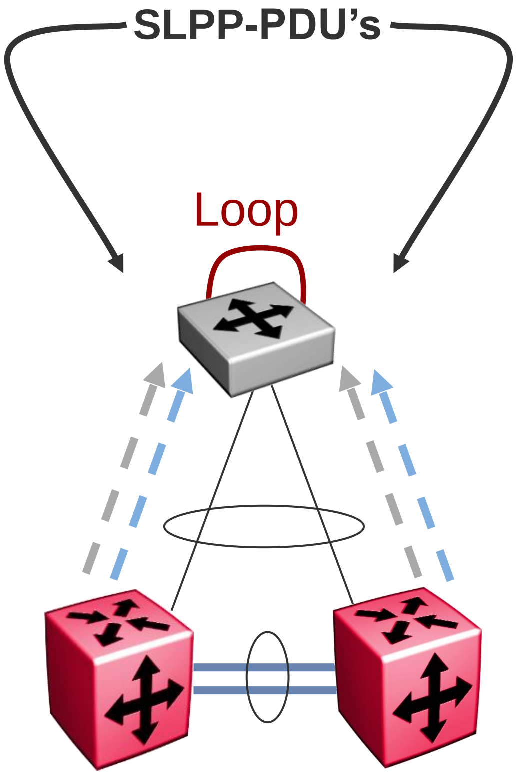
SLPP-PDU's only received if there is a loop

SLPP Guard *Feature Overview*

- The Switch Clustering implementations on the VSP9000/8000/4000/7200/7000, ERS8800/8600 and ERS5000 provide a Simple Loop Prevention Protocol (SLPP) packet which operated to help prevent loops occurring when Switch Clustering is implemented
- In some customer environments there is a need to provide additional loop protection when used in combination with Avaya’s Switch Clustering (SMLT).
- SLPP-guard helps prevent loops in customer’s networks by administratively disabling a edge port if they received a SLPP packet
- Loop prevention for edge ports can be provided by enabling STP on edge ports, which provides protection for customers looping back edge ports to the same switch or stack.
  - Note that STP loop prevention will not work if the attached device (hub, switch, IP phone) does not support Spanning Tree (does not pass STP BPDU’s) and is connected to two or more ports on the Avaya edge switch
- In some networks due to moves, adds or changes, it could be possible to create a loop within the customers networks by connecting an edge port back to a port of the switch cluster
- When operational, SLPP-guard will immediately administratively disable a port when a SLPP packet is received on a port and generate a local log message, syslog message (if the syslog server(s) are configured) and SNMP traps (if SNMP tap receivers are configured)

SLPP Guard *Feature Overview*

- Each port has its own administrative hold-down timer
  - when the port is shutdown due to reception of a SLPP packet the timer should start for that port.
  - When the timer reaches the configured interval, the port is re-enabled and a local log message, syslog message (if the syslog server(s) are configured) and SNMP traps (if SNMP tap receivers are configured).
  - This timer is user configurable between 10 seconds and 65535 seconds, with 60 seconds set as the default. The port timer is disabled if it is configured as 0, which means the port will be disabled until an administrator re-enabled the port.
- The default SLPP Ethertype is (hex): 0x8102, though on some switches it has used an old value of 0x8104.
  - You can globally configure the EtherType for SLPP guard.
- Starting with release 10.4 for the VSP 7000, SLPP-Guard can be enabled on MLT/LACP links



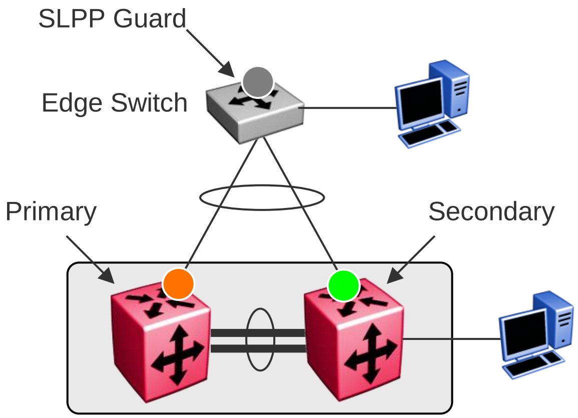
Switch Clustering Best Practices

Configuration Checklist

☑ **SLPP enabled per VLAN and on uplink and access ports**

☑ **SLPP Guard on Edge switch if supported**

- Enable SLPP
  - Per VLAN
  - Per port by setting Rx Threshold
- Identify one IST peer as Primary and the other as Secondary
  - Not a configurable option, strictly from a design standpoint
  - Enable RxThreshold per table below on uplink ports
- Do not enable auto recovery – Once the port is down, it will stay in the down state and need manual intervention to be enabled
- Do not enable SLPP-Rx on IST ports
  - Never want to take these ports down



● RxThreshold = 5

● RxThreshold = 50

● SLPP Guard

Note: A SLPP Rx threshold of 5 and 50 applies only if one or perhaps two VLANs are configured to the edge switch where the secondary is multiplied by 10. These number should be increased if additional VLANs are add and SLPP is enabled for each VLAN.

SMLT Cluster Switch	Ethertype	Packet Rx Threshold	Transmission Interval
Primary	Default	5	Default (.5 seconds)
Secondary		50	

Edge Switch	SLPP Holdown timer
All non core ports (edge only ports)	0

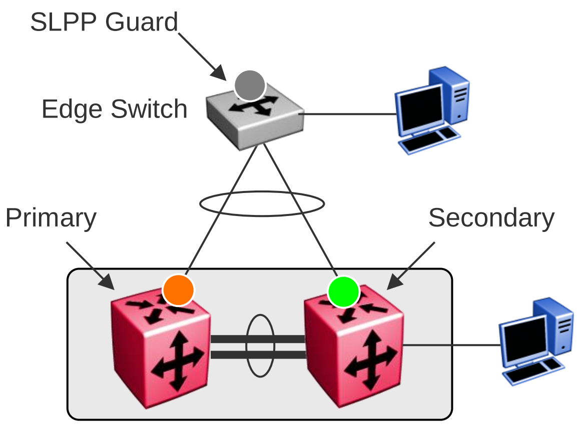
Switch Clustering Best Practices

Configuration Checklist

☑ **SLPP enabled per VLAN and on uplink and access ports**

☑ **SLPP Guard on Edge switch if supported**

SLPP Rx-threshold is a cumulative count. This may cause a situation depending on the duration in which multiple different loop events and may lead to an event where both primary and secondary links have their threshold reached and both links bring their ports down, and edge isolation could occur. A disable/enable of SLPP, which does not impact the network, should be performed after any SLPP event to clear the counter . Note that as of release 5.1.4 and 7.1, the ERS 8800 will automatically re-arm the SLPP counters after every 24 hours. The VSP 9000/7000/8000 will automatically re-arm the SLPP counters every 6 hours.



🟠 RxThreshold = 5

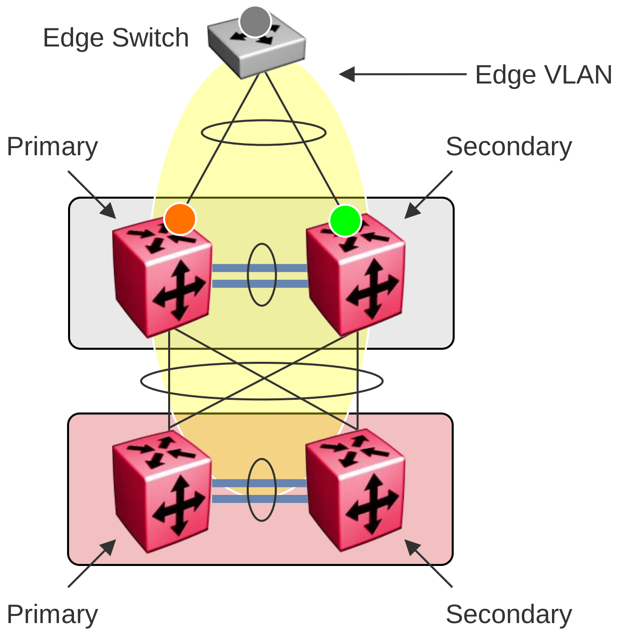
🟢 RxThreshold = 50

● SLPP Guard

As the number of VLANs running SLPP scale off of a specific uplink port, the Rx-threshold value may need to be increased to prevent complete isolation of the offending edge. Critical to note is that the primary goal of SLPP is to protect the core at all costs. In certain loop conditions, what may occur is the secondary switch also detecting the loop and its SLPP Rx-threshold is reached before the primary can stop the loop by taking its port down. Therefore, both switches eventually take their ports down and the edge becomes isolated. The larger the number of VLANs associated with the port, the more likely this could occur, especially for loop conditions that affect all VLANs. The recommended step here is to increase the Rx-threshold. As a guideline, increase the Rx Threshold on the primary switch using a multiplier of 5 for each VLAN and increase this number on the secondary switch using a multiplier of 10, i.e. for 5 VLANs, use SLPP RxThreshold values of 25 and 250.

Switch Clustering Best Practices

Bridged Core



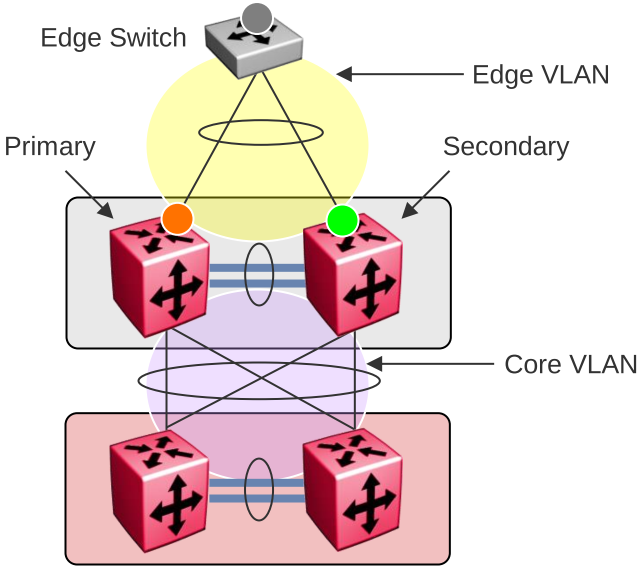
🟠 RxThreshold = 5

🟢 RxThreshold = 50

⚫ SLPP Guard

- Increase RxThreshold on the Switch Cluster core ports
- Loops at the edge should be caught at the edge and not shut down core ports

Routed Core



🟠 RxThreshold = 5

🟢 RxThreshold = 50

⚫ SLPP Guard

- Use/Create a VLAN that spans ONLY the routed core between Switch Clusters
- Enable SLPP on that VLAN to catch loops that occur within the core
- Loops at the edge will not cause ports to be shut down in the core

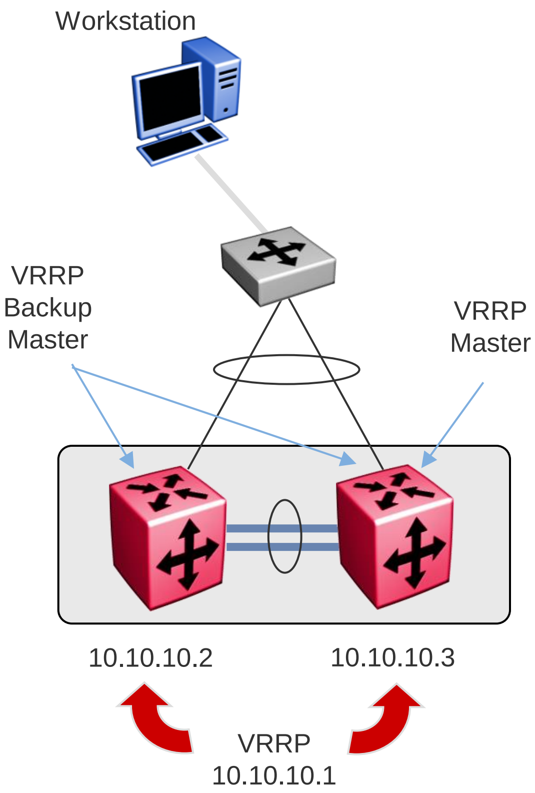
Switch Clustering Best Practices

Configuration Checklist

☑ Default Gateway Redundancy

VRRP with Backup Master

- Always use three addresses, two physical VLAN addresses and one virtual address – Do not use the physical IP address of the VLAN as the VRRP address
- Enable Backup Master on core switches
- Define VRRP Master by increasing VRRP Priority to 200 (100 is default – any value greater than 100 is acceptable)
- Balance VRRP Master between core switches across VLANs
- On the ERS 8000 only
  - Set Advertise Interval to 10 seconds
  - Set Hold-Down Timer to 60 seconds
- Supports up to 255 instances
- Do not use the same VRID across multiple VLANs
- See ERS 5000 Design Requirements section for other details on VRRP recommendations with ERS 5000 Switch Cluster



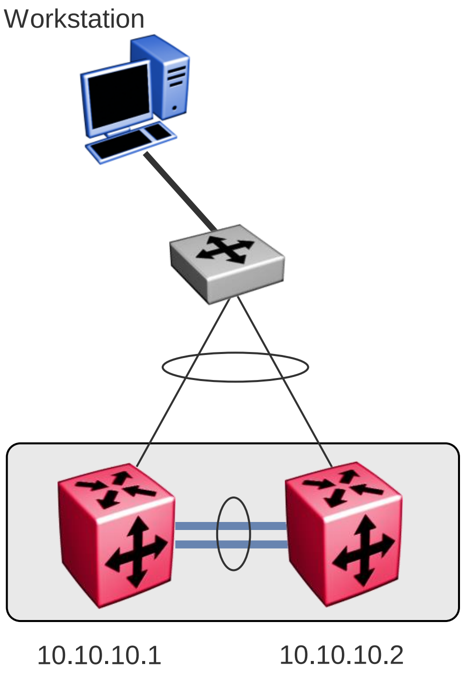
Switch Clustering Best Practices

Configuration Checklist

Default Gateway Redundancy

☑ RSMLT Layer 2 Edge

- Both IST peers can forward on behalf of each other
- Much less overhead than VRRP
- Scales beyond 255 instances
- Set Hold-Up Timer to 9999 seconds (infinity)
- Globally enable rsmlt-edge support
- Once both IST peers are up and running, must save the configuration file on each switch to ensure the peer's MAC address is saved
- Do not use VRRP and RSMLT Layer 2 Edge on the same VLAN simultaneously
- Supported on:
  - VSP 9000/8000/7200/4000
  - ERS 8800/8600
  - ERS 8300



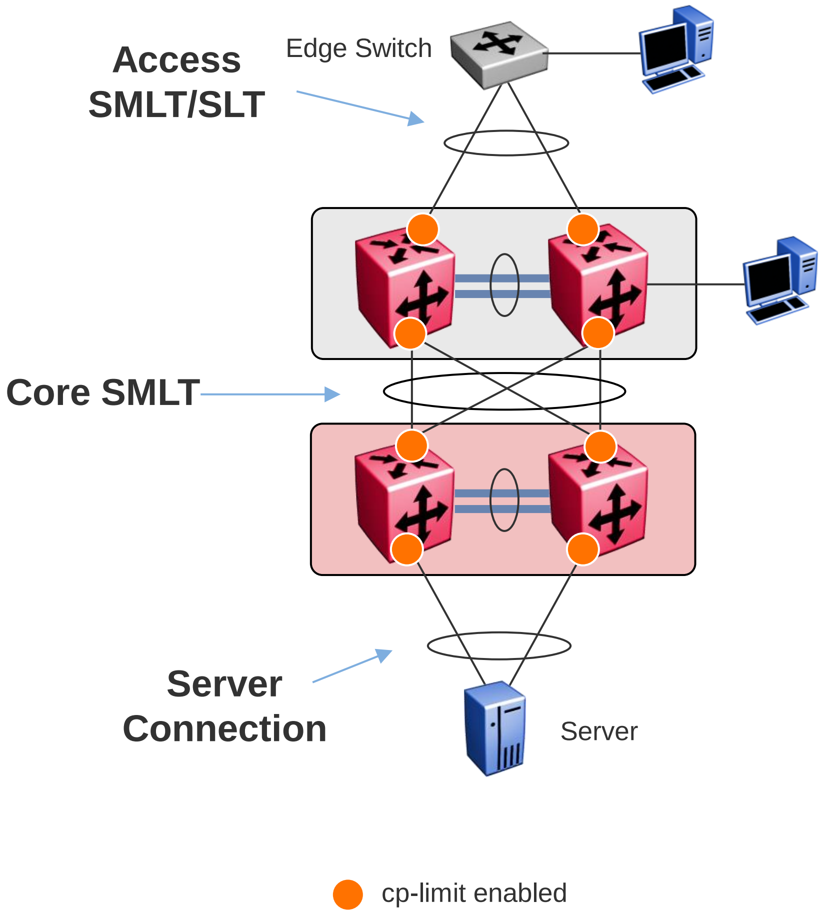
Control Plane Rate Limit (CP-Limit)

Feature Overview

- Protects CPU from broadcast and multicast storms
  - Looks at "**Control**" multicast traffic and broadcast traffic and also independently QoS preference level 7 traffic hitting the CPU
    - Only packets destined to the CPU
    - Should be used with Rate Limit (bandwidth limit) feature where CP Limit should be lower that the Rate Limit
  - If the defined packet rate per second is exceeded, the port is shut down
    - Need to disable/enable port to recover
  - Does NOT look at data packets (session/user traffic)
  - Does NOT protect against traffic exception traffic such as: SNMP, telnet, ICMP, IP with TTL1, Unknown SA, etc.
  - Enabled on all ports by default
    - Automatically disabled on IST ports during IST creation
- Supported on:
  - VSP 9000
  - ERS 8800/8600
  - ERS 8300
- Please note that with Release 6.0 for the ERS 5000 series, a CPU limiting feature was implemented, however, this is not a user configurable feature

CP-Limit Guidelines

	CP-Limit Values	
	Broadcast	Multicast
Aggressive		
Access SMLT/SLT	1000	1000
Server	2500	2500
Core SMLT	7500	7500
Moderate		
Access SMLT/SLT	2500	2500
Server	5000	5000
Core SMLT	9000	9000
Relaxed		
Access SMLT/SLT	4000	4000
Server	7000	7000
Core SMLT	10000	10000



Extended CP-Limit

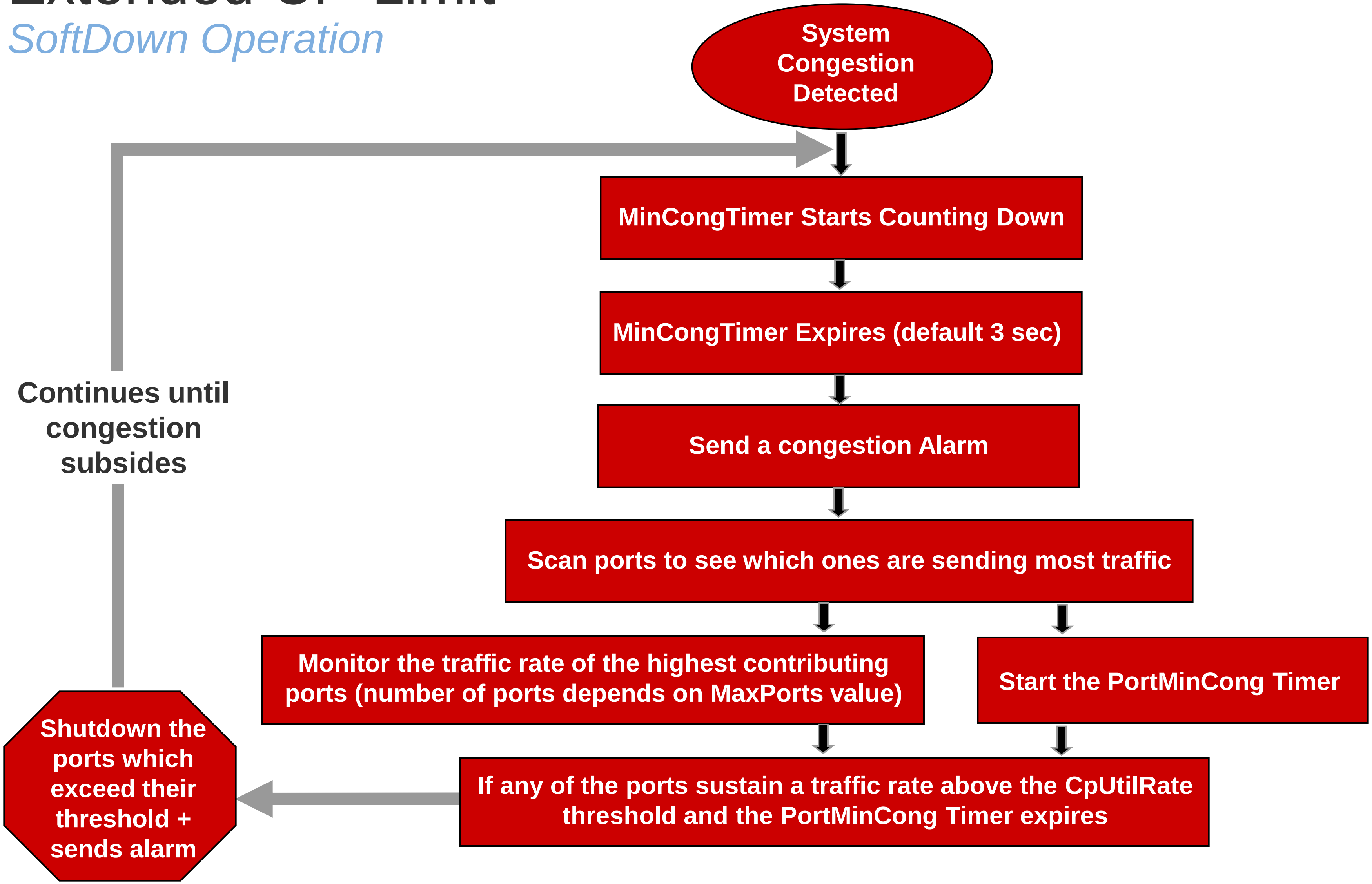
Feature Overview – ERS 8600-8800 Only

- Can be used in conjunction with cp-limit and expands on the ability of cp-limit by monitoring
  - Buffer congestion on the CPU
  - Port level congestion on the I/O modules
- Does NOT look at data packets (session/user traffic)
- SoftDown monitors port for “x” duration, if congestion remains, port is disabled
- Can be enabled on all ports of the 8800/8600 – the max ports value indicates the number of ports monitored during a time of congestion (i.e., With maxports = 5, the 5 highest ports in terms of utilization are monitored)
- Enable SoftDown with the following values:
  - Maxports = 5
  - MinCongTime = 3 seconds (default)
  - PortCongTime = 5 seconds (default)
  - CPLimitUtilRate = Dependent on network traffic *

* Network must be baselined to understand the average utilization rate. Once the average rate is known, the CPLimitUtilRate should be set to a value of (3 * average utilization rate), but not higher than 70% under normal, average network conditions. This is a basic guideline for the “normal” enterprise network – for networks with normally high utilization, these values may differ.

Extended CP-Limit

SoftDown Operation



System Congestion Detected

MinCongTimer Starts Counting Down

MinCongTimer Expires (default 3 sec)

Send a congestion Alarm

Scan ports to see which ones are sending most traffic

Monitor the traffic rate of the highest contributing ports (number of ports depends on MaxPorts value)

Start the PortMinCong Timer

If any of the ports sustain a traffic rate above the CpUtilRate threshold and the PortMinCong Timer expires

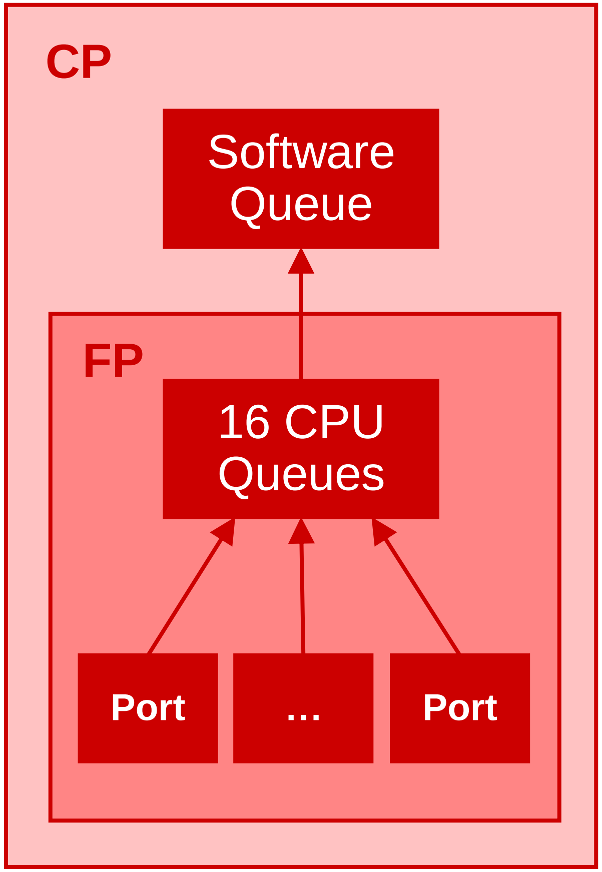
Shutdown the ports which exceed their threshold + sends alarm

Continues until congestion subsides

CP Overload Protection

Feature Overview – VSP 4000/7200/8000

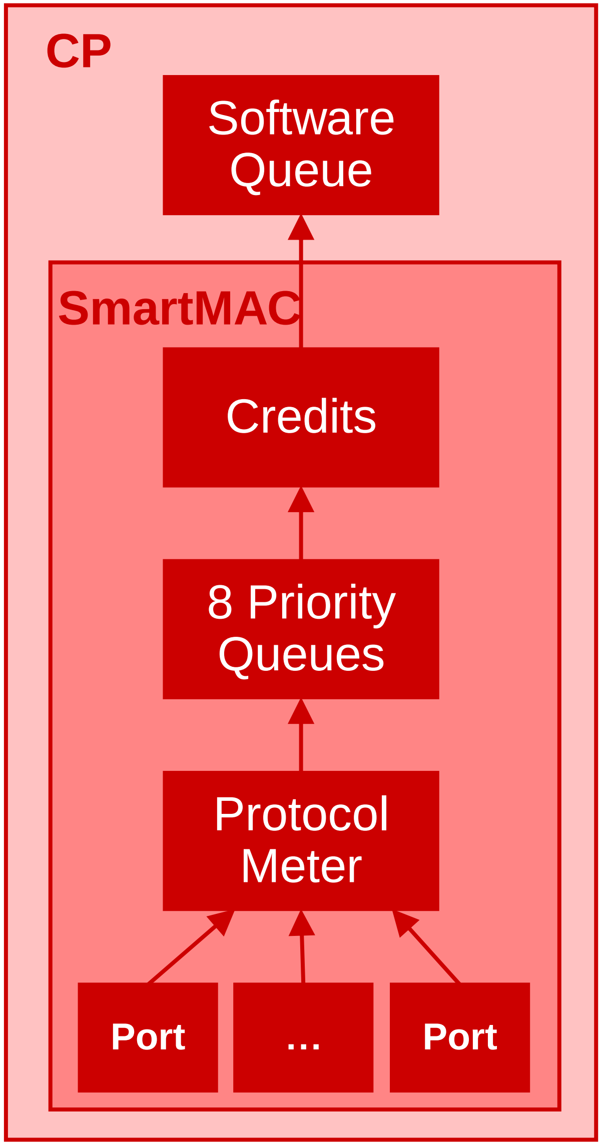
- Hardware Assist to protect CP from getting overloaded with Packet Processing
- CPU Queue Meters : Limit traffic to CP based on classified protocols
  - Prevents one protocol from storming the CP impacting all other protocols
  - non-IP traffic (ISIS, VLACP, CFM, IST,…)
    - 9 CPU queues (Q15 - Q8 & Q0)
  - IP traffic (VRRP,ICMP,OSPF,BGP,…)
    - 7 cos queues (Q7 - Q1)
- Control traffic queues cannot be configured, modified to displayed.
- CP-limit cannot be configured on port



CP Overload Protection

Feature Overview – VSP 9000

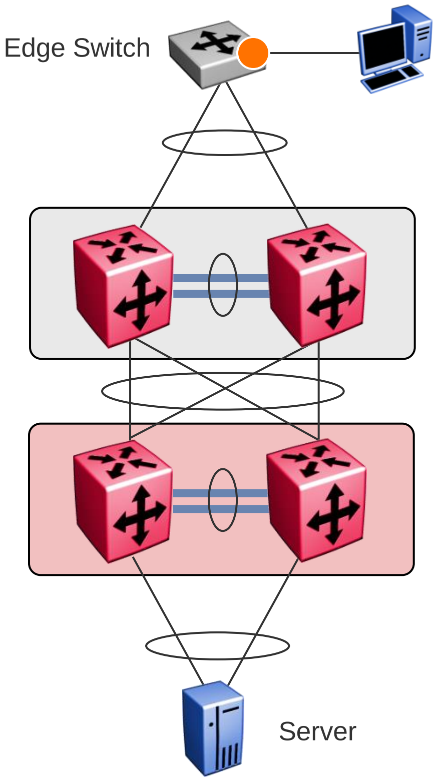
- Hardware Assist to protect CP from getting overloaded with Packet Processing and DoS attacks
- Port meters limit the number of packets going to CPU from a particular port. Port meter values can be modified using the cp-limit command.
- Smart Mac classifies packets to an internal protocol value (ISIS, VLACP, OSPF, BGP, …). Protocol Meters limit the number of packets from a particular protocol. These are not user configurable and are based on scaling numbers. Out of profile packets are dropped
- Priority Queues have a packet per second threshold. If exceeded, it flow controls all the way to the Fabric Access Device on the IO card.
- Hardware Credit mechanism provides bounded latency to service real-time control traffic under congestion and also provides weighted-fair bandwidth allocation among control packet classes of service.



Broadcast / Multicast Port Rate Limiting

Feature Overview

- Enable Rate Limiting on the edge access ports to protect from broadcast/multicast storms
  - Protects against non-CPU bound traffic
  - Must understand multicast and broadcast traffic in the network before enabling rate limiting
- ERS Implementation
  - Broadcast / Multicast Rate limiting allows the user to configure the allowed amount of bcast/mcast traffic on a port. When traffic exceeds this threshold, it is dropped.
  - ERS 2500 / 3500 / 4000 / 5000, VSP 7000
    - 1 – 10% of port speed
    - Recommendation → 10%
  - ERS 8300
    - 1-100% of port speed
    - Recommendation → 10%
  - VSP 9000 / ERS 8800/8600 (legacy, E-series, M-series modules)
    - Broadcast / multicast rate limiting
    - Allowed rate is in packets per second (pps)
    - Recommendation → 3 times normal pps
  - ERS 8800/8600 (R-series, RS-series modules)
    - Broadcast / multicast bandwidth limiting
    - Allowed rate is in kbps
    - Recommendation → 3 times normal kbps

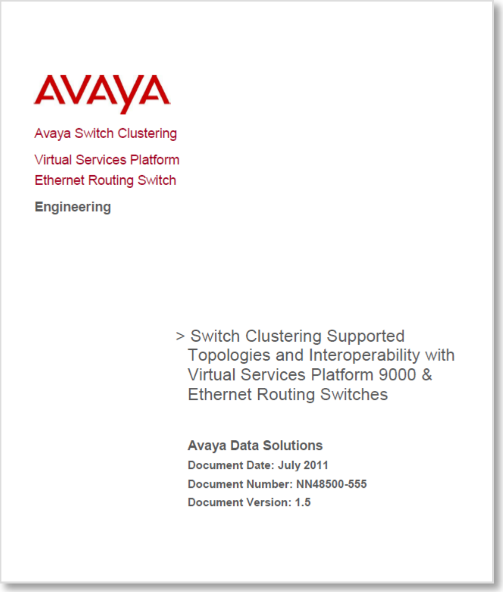


● Rate Limiting Enabled

Multicast and Switch Clustering

Supported Configurations & Features

- PIM-SM with Switch Clustering is supported on:
  - ERS 8800/8600/8300 (SMLT/SLT/RSMLT)
  - VSP 9000 (SMLT/SLT/RSMLT)
- PIM-SM is NOT supported on the ERS 5000 Switch Cluster. IGMP is supported on the ERS 5000 Switch Cluster as of 6.3 and on the VSP 7000 as of 10.3
- Enable PIM-SM on the IST VLAN (no unicast routing protocol is required) for fast recovery of multicast
- Enable IGMP Snooping and Proxy on the Edge switches when running multicast on the network
- When running PIM-SM over ERS 8800/8600 Square or Full Mesh, enable mcast-smlt square-smlt flag
- For details on all supported topologies refer to:
  - Switch Clustering Supported Topologies and Interoperability (NN48500-555)
- For details on Multicast and Switch Clustering configurations refer to:
  - Resilient Multicast Routing Using Split-Multilink Trunking for the ERS 8800/8600 (NN48500-544)



AVAYA

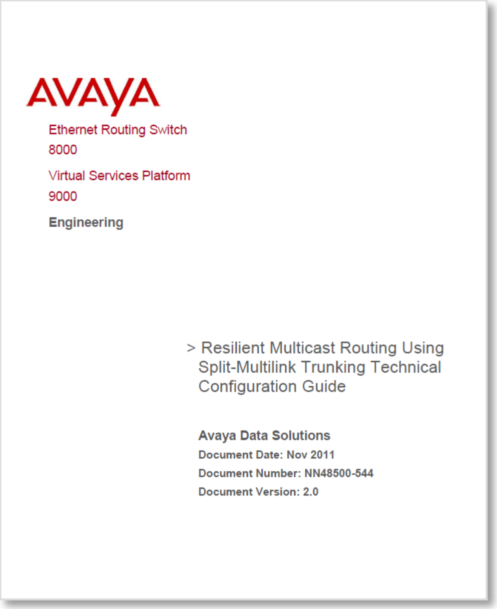
Avaya Switch Clustering

Virtual Services Platform
Ethernet Routing Switch

Engineering

> Switch Clustering Supported Topologies and Interoperability with Virtual Services Platform 9000 & Ethernet Routing Switches

Avaya Data Solutions
Document Date: July 2011
Document Number: NN48500-555
Document Version: 1.5



AVAYA

Ethernet Routing Switch
8000

Virtual Services Platform
9000

Engineering

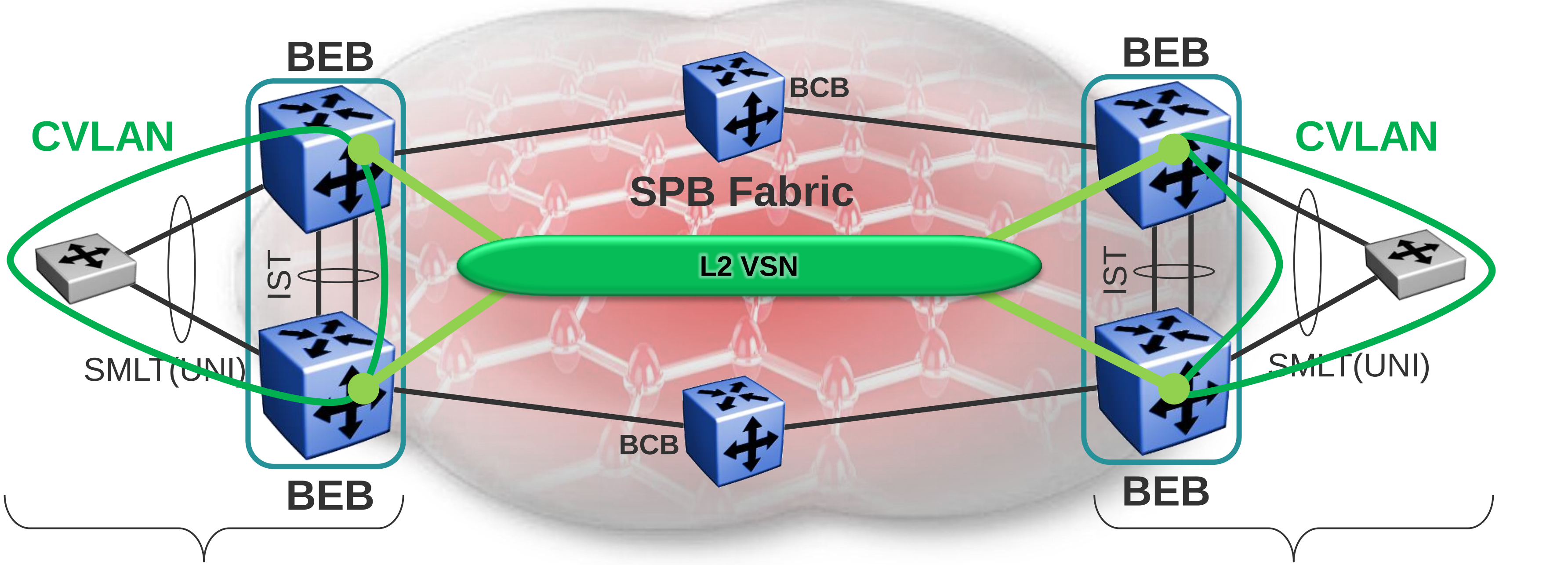
> Resilient Multicast Routing Using Split-Multilink Trunking Technical Configuration Guide

Avaya Data Solutions
Document Date: Nov 2011
Document Number: NN48500-544
Document Version: 2.0

SPB SMLT BEB Design Best Practices

All the recommendations covered in the previous sections, also apply to SMLT when used at the edge of an SPBM Fabric (SMLT-BEB).
This section covers some additional best practices specific to an SMLT-BEB configuration.

SMLT BEB – CVLAN Guidelines for L2VSN

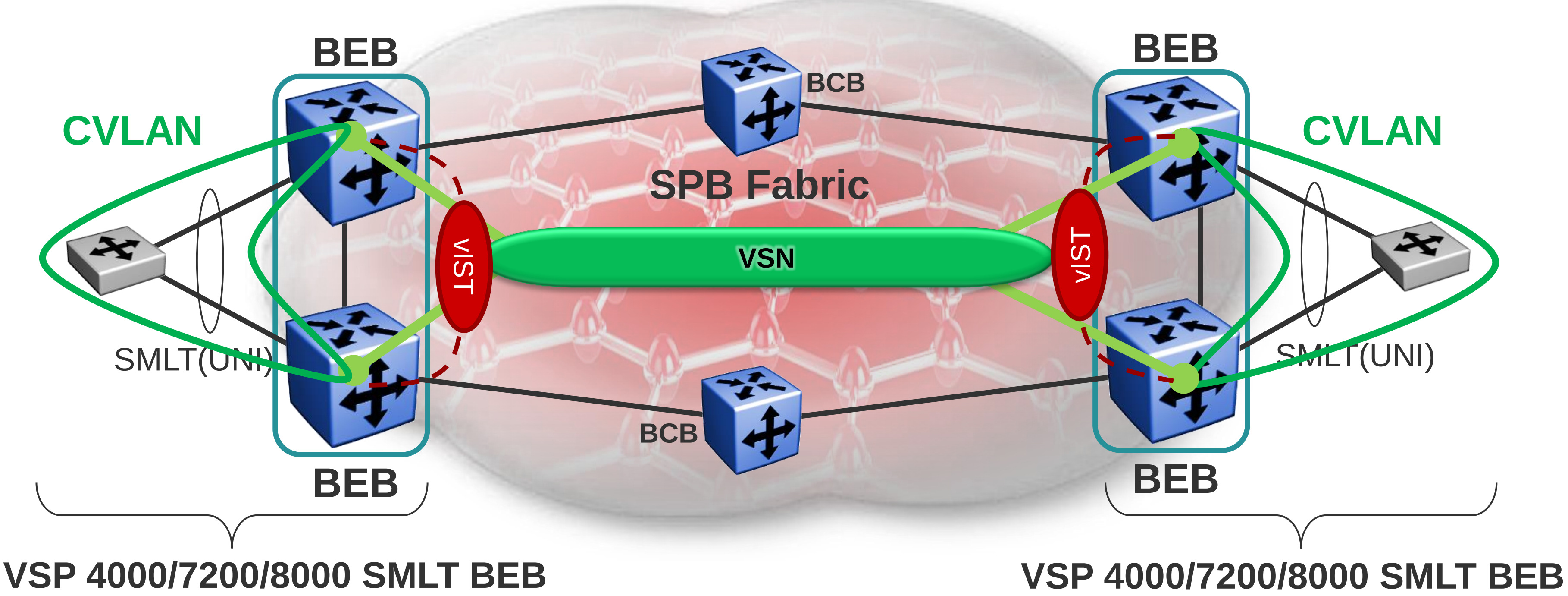


VSP 9000
SMLT BEB

ERS 8800 and VSP 7000 SMLT
BEB

- Customer VLAN (CVLAN) has I-SID assigned and is thus L2 extended with L2VSN
- On the ERS 8800, VSP 7200, VSP 4000, VSP 8000, and VSP 7000 the CVLAN cannot be configured on any NNI interface (including the IST)
- On the VSP 9000 the CVLAN cannot be configured on any NNI interface (except on the IST where it MUST be configured)

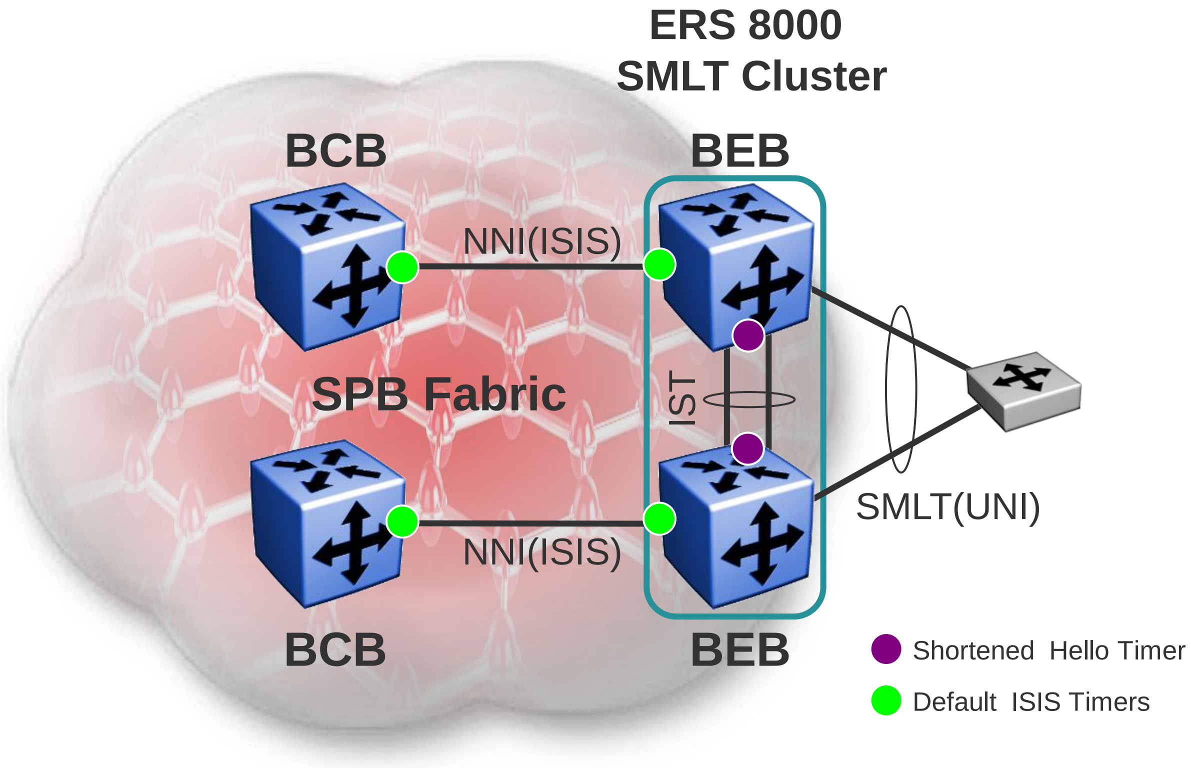
SMLT BEB – Guidelines for vIST



- For each C-VLAN provisioned, an SPBM I-SID must be assigned on both cluster switches regardless of the service (L2VSN, IP Shortcuts, or L3VSN)
- For an L2VSN service, this I-SID can be used on other BEB switches in the network to form an L2VSN service
- For an IP Shortcuts or L3VSN service, this I-SID value must be unique and used only on the two vIST cluster switches
  - In addition, for an L3VSN service, a separate I-SID value is used to identify the L3VSN service

SMLT BEB – ERS 8800 ISIS Hello Timer Guidelines

- The recommended timers on this slide only applies to the ERS 8800
- On the IST, ISIS is enabled on the MLT bundle
- Upon node restart, we need the ISIS adjacency over the IST MLT to come up before the IST comes up, therefore the ISIS Hello timer is reduced to 1 sec
- The hello multiplier is increased by the same factor to ensure the same time delay for an ISIS adjacency to transition in the down state
  - 1x27 = 9x3 = 27



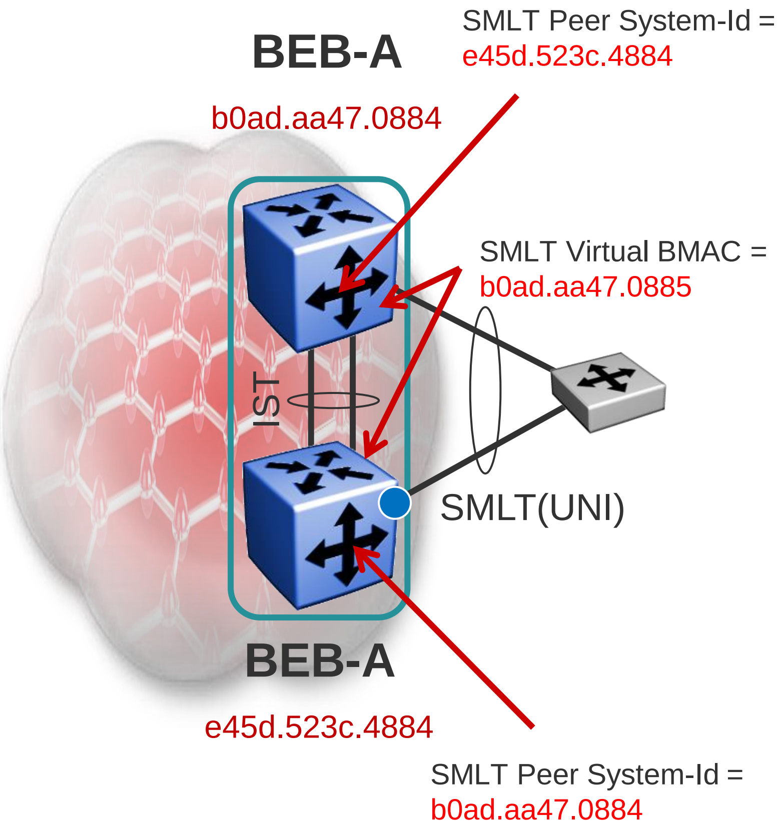
Connection Type	l1-hello-interval	l1-hello-multiplier
● IST -NNI ISIS	1 sec	27
● NNI ISIS	9 secs (default)	3 (default)

System ID and SMLT Peer System ID

- By default, the SPB System-Id is a MAC address using the MAC address range reserved for the switch
- To ensure there will be no de-stabilizing System-Id conflicts in the network, Avaya recommends to use of the default System-Id value
  - Please note that starting in release 5.0 for the VSP 4000/7200/8000, support for duplicate System-Id and SMLT peer System-Id detection is supported
- If the System-Id value is configured, it is very critical to ensure that each SPB enabled switch in the network uses a unique ISIS System-Id value
  - If you do change the System-Id, it is recommended to set the locally administered bit
    - This is the second least significant bit of the most significant byte of the MAC address that should be set to 1 to indicate the MAC address is locally administered
    - For more details, please go to http://en.wikipedia.org/wiki/MAC_address

System ID and SMLT Peer System ID, con’t

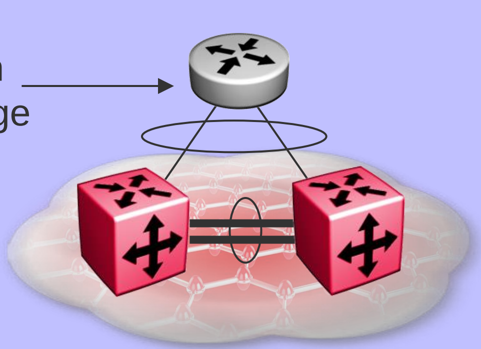
- In an SMLT cluster, each cluster switch must peer with it’s neighbors ISIS System-Id
- By default, an SMLT virtual BMAC will created using the lowest System-Id in the cluster plus one to allow an remote node to reach either SMLT cluster switch
  - If the System-Id is configured, please ensure that the System-Id is separate by at least two octets to avoid the System-Id being the same value as the SMLT Virtual value



Switch Clustering

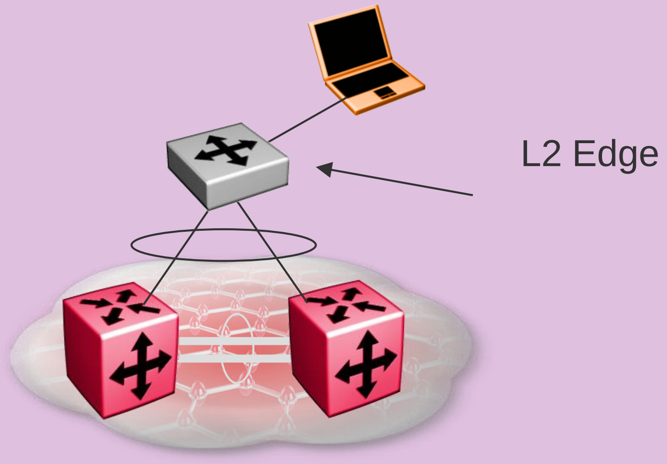
Terminology – RSMLT

RSMLT – Routed Layer 3 Edge



** Need to enable route redistribution between ISIS and protocol used on edge router, i.e. OSPF to ISIS and vise versa.

RSMLT Edge – Routed Split MultiLink Trunking Replacement for VRRP with Layer 2 Edge



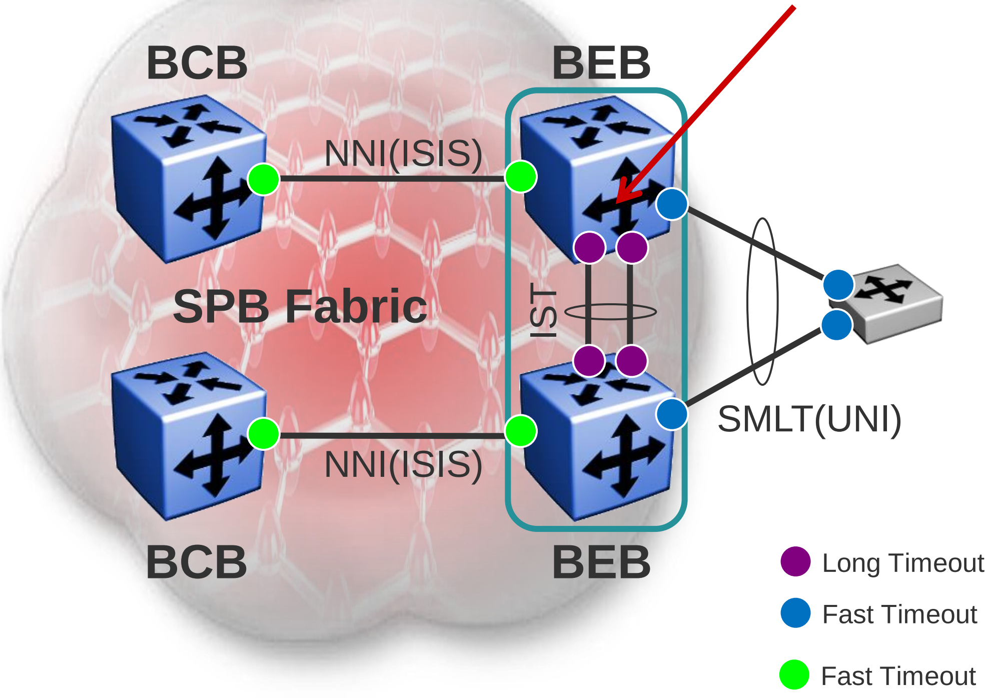
Scales beyond maximum number of VRRP instances

Do not run VRRP and RSMLT on the same edge VLAN simultaneously

** Note that if the OSPF network has multiple entry points via multiple SPB nodes, OSPF route policies or ISIS Accept Policy must be configured on the SPB BEB switches to deny OSPF routes from each remote BEB entry point to prevent routing loops.

SMLT BEB – VLACP Guidelines

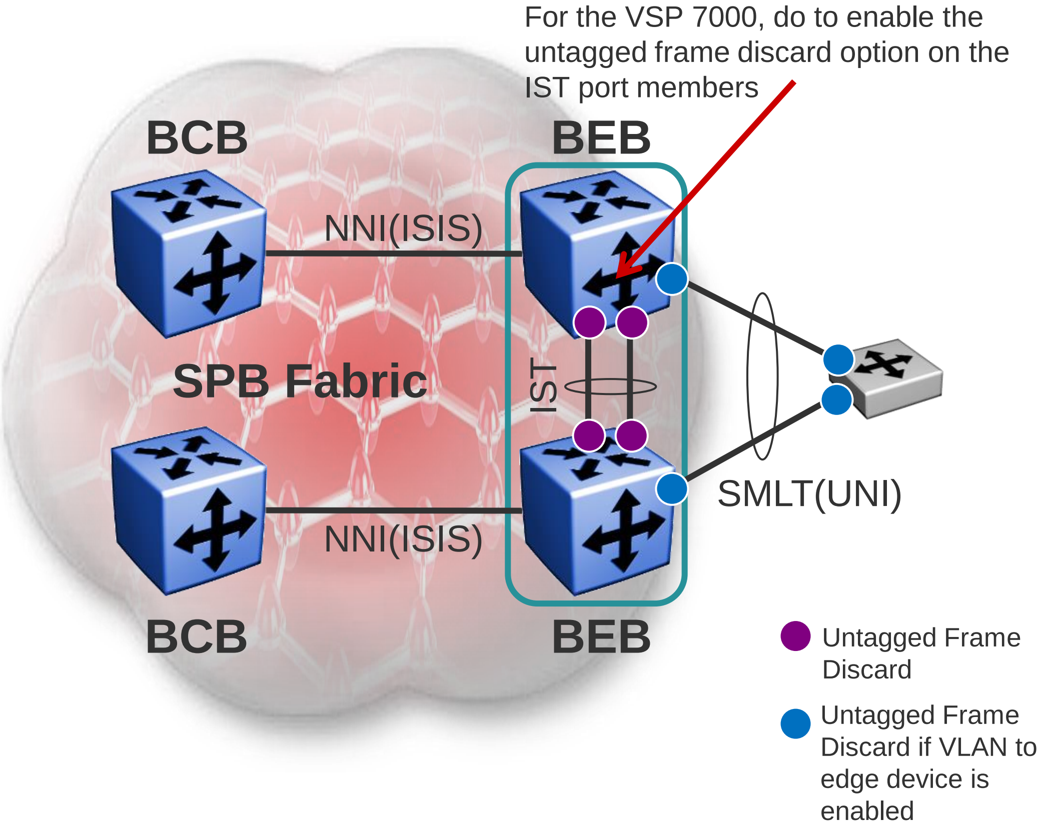
- Enable VLACP on all NNI ISIS enabled interfaces
- IST (which is now also an NNI connection) uses same VLACP slow timers
  - This does not apply to the VSP 7000 where VLACP should not be enabled on the IST port members
- Core facing NNI interfaces use same VLACP timers as SMLT UNI connections


For the VSP 7000, VLACP should not be used on the IST port members

Connection Type	Fast Timer	Slow Timer	Timeout	Timeout Scale	ERS 8000 VSP 9000	VSP4000 VSP 7200 VSP8000	VSP 7000
● IST (+ NNI ISIS)	N/A	10000	Long	3	☑	N/A	☒
● SMLT (UNI)	500ms	N/A	Short	5	☑	☑	☑
● NNI (ISIS)	500ms	N/A	Short	5	☑	☑	☑

IST – Untagged Frame Discard Option

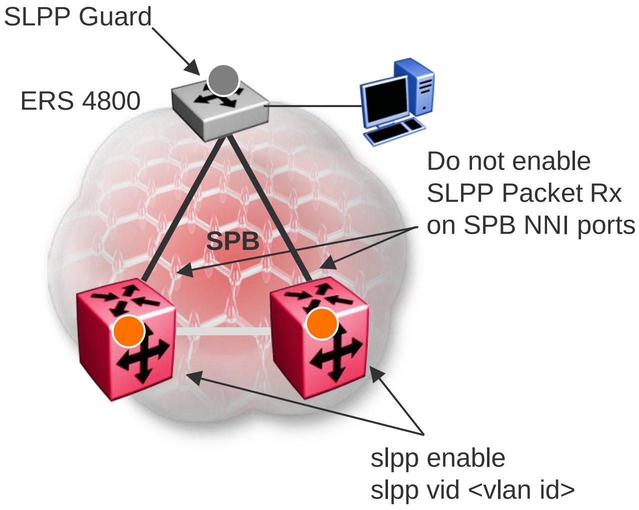
- Enable the untagged frame discard option is a good option to enabled on all tagged links
  - On the VSP4000 and VSP8000 with vIST, this only applies to the SMLT UNI ports
- On the VSP 7000 SMLT cluster only, the untagged frame option should not be enabled on the IST port members
  - You can still enable on SMLT link proving the trunk port is tagged


For the VSP 7000, do to enable the untagged frame discard option on the IST port members

SLPP Guard

VSP 7000, ERS 3500/5500/5600/4500/4800

- SLPP can enabled on the core bridges and in turn SLPP Guard can be enabled on the edge switch for local port loop detection
  - The setting of the overload bit on the edge switch allows it to operate as a stub node on the SPB network
    - This prevents traffic from one NNI port to be forwarded to another NNI port
  - Because of this feature, SLPP can be enabled on the core SPB bridges and in turn allowing SLPP Guard to be enabled on the edge switch
- Only enable SLPP on the C-VLAN on the core SPB bridges
  - Do not enable SLPP Packet Rx on core NNI ports
    - Never want to take these ports down



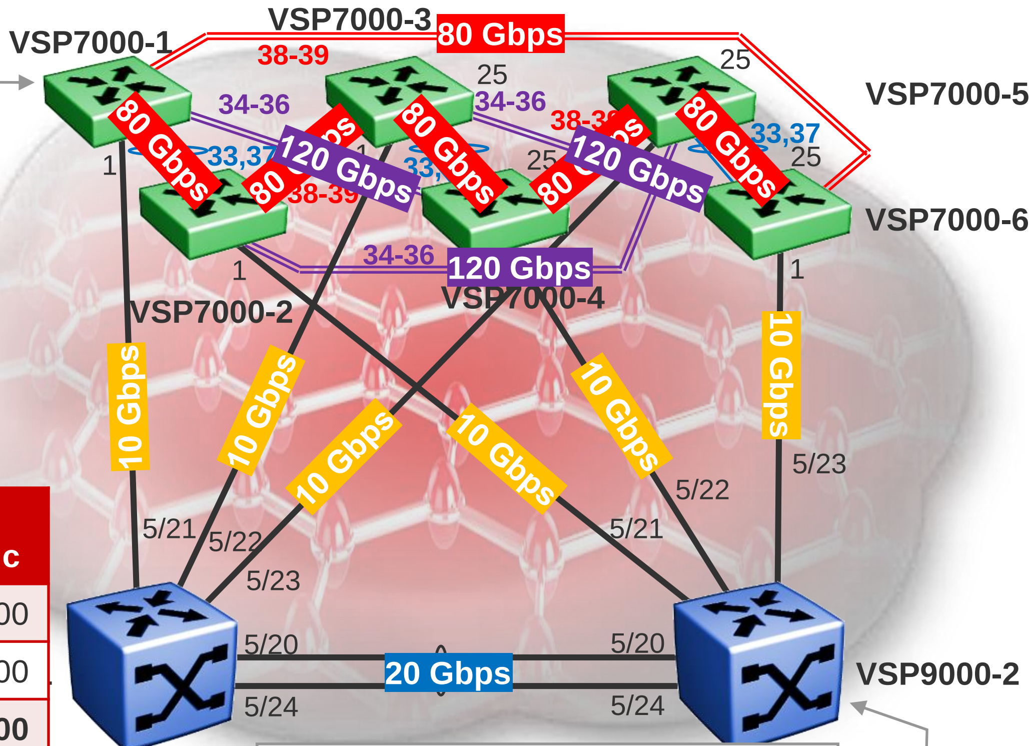
● Enable SLPP on C-VLAN

● SLPP Guard

Tuning ISIS metrics (optional)

```
interface Ethernet 1
   isis spbm 1 l1-metric 200
exit
interface Ethernet 33,37
   isis spbm 1 l1-metric 25
exit
interface Ethernet 34,35,36
   lacp key 3436
   isis spbm 1 l1-metric 17
exit
interface Ethernet 38,39
   lacp key 3839
   isis spbm 1 l1-metric 25
exit
```

Link Speed	Interface Type	ISIS L1-metric
1 Gbit/s	Native ethernet	2000
2 Gbit/s	MLT bundle	1000
10 Gbit/s	**Native ethernet**	**200**
20 Gbit/s	**MLT bundle**	**100**
40 Gbit/s	Native ethernet	50
80 Gbit/s	**MLT bundle / FI**	**25**
100 Gbit/s	Native ethernet	20
120 Gbit/s	**MLT bundle / FI**	**17**
160 Gbit/s	MLT bundle	13
1 Tbit/s	Future!	2



```
interface GigabitEthernet 5/21-5/23
   isis spbm 1 l1-metric 200
exit
interface mlt 512
   isis spbm 1 l1-metric 100
exit
```

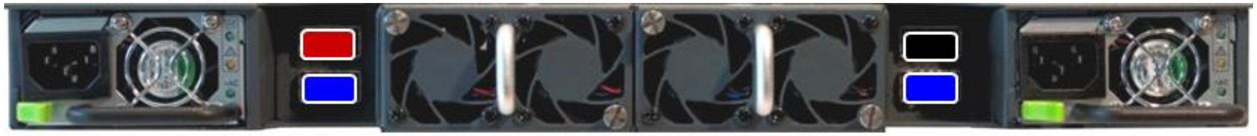
- Table has suggested metrics; default metric is always 10 no matter the port speed
- On VSP7000 FI, to have a different metric on different FI rear ports (with LACP enabled) you have to have different LACP keys
  - Note: changing the LACP key will make the ISIS adjacency bounce

VSP 7000 SMLT and Rear Port Mode

Fabric Interconnect

Desired Deployment Model	Needed Rear-port Mode	SMLT (IST) Needed	SPB enabled	Virtual Servers (e.g. ESX) NIC teaming	Server NIC teaming (LACP)	Minimum Required Software
vToR FI Stacking	Disabled (= Stacking enabled)	No	Can be	Yes – Vport hashing on non-SLT ports	Yes – On DMLT ports (with or without LACP)	10.1.0 (10.3.0 if need to run SPB on uplinks)
dToR FI Stacking with SMLT	Disabled (= Stacking enabled)	Yes	Can be	Yes – Vport hashing on non-SLT ports - OR - IP hashing on SLT ports	Yes – On SLT ports (with or without LACP)	10.2.0 (10.3.0 if need to run SPB on uplinks & IST)
FI Mesh with SMLT	Enabled in raw mode	Yes	No	Yes – Vport hashing on non-SLT ports - OR - IP hashing on SLT ports	Yes – On SLT ports (with or without LACP)	10.2.0
SPB Mesh	Enabled in SPBM mode	No	Yes	Yes – Vport hashing on non-SLT ports	No – Use Active Standby NICs	10.2.0
SPB Mesh with SMLT	Enabled in SPBM mode	Yes	Yes	Yes – Vport hashing on non-SLT ports - OR - IP hashing on SLT ports	Yes – On SLT ports (with or without LACP)	10.3.0

VSP 7000 Rear Ports



Color	Physical Fabric Interconnect Port	Rear Port Mode	Throughput	Ports
Black	R1 Up (right) Top	Standard	240Gpbs (120 FDX)	34, 35, 36
		SPB	240Gpbs (120 FDX)	
Red	FI Down (left) Top	Standard	240Gbps (120 FDX)	38, 39, 40
		SPB	160Gbps (80 FDX)	38, 39
Blue	FI Up (right) Bottom	Standard	80Gbps (40 FDX)	33
		SPB	80Gbps (40 FDX)	
Blue	FI Down (left) Bottom	Standard	80Gbps (40 FDX)	37
		SPB	80Gbps (40 FDX)	

VSP 7000 Rear Ports – Default Settings

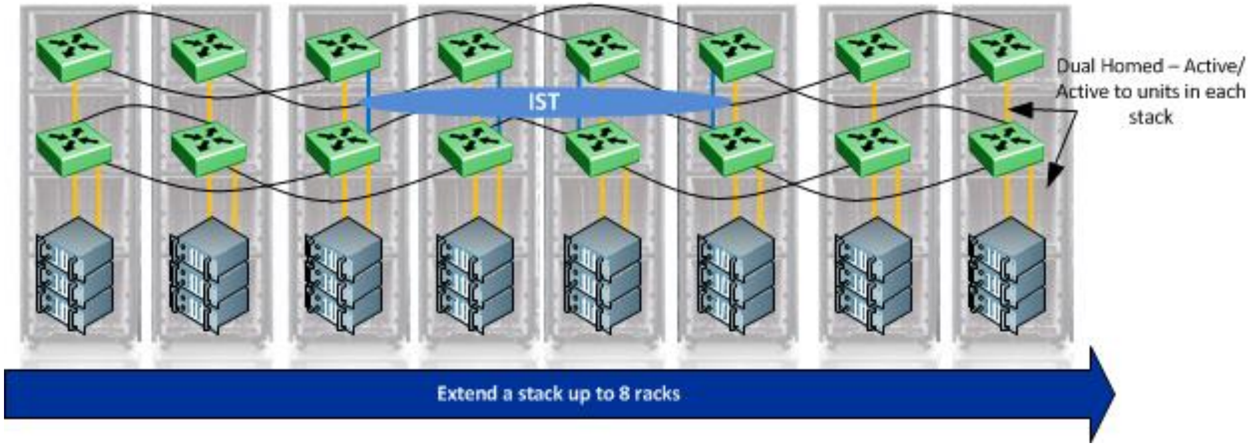
- When you enable rear port mode, the switch applies the following default settings to all FI ports on the rear of the chassis:
  - VLAN tagging for rear ports is set to tagAll
  - The LACP administration key is set to 4095
  - The LACP operating mode for rear ports is set to active
  - The LACP rear ports time-out value is set to short
  - LACP for rear ports is set to enable

VSP 7000 Rear Ports – SMLT Rules

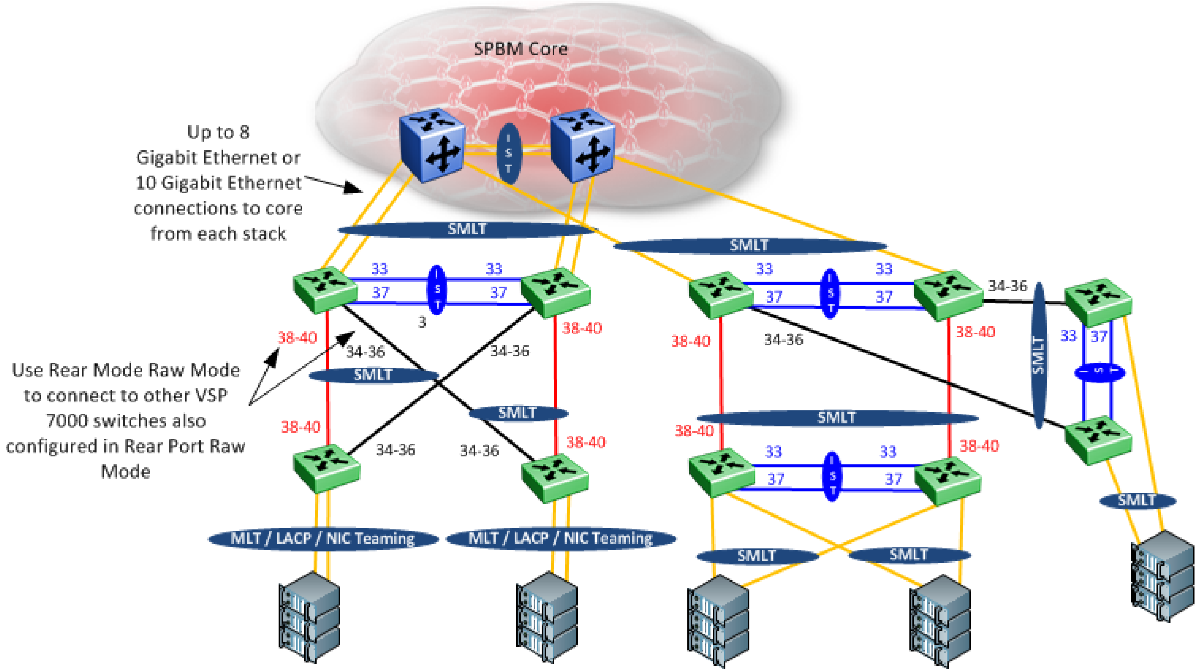
- The VSP 7000 requires at minimum two ports members for an IST
- LACP is not supported on the IST and must be disabled on the rear ports prior to enabling the IST
- Do not enable the *filter-untagged-frame* option on the IST port members
- If enabling SPB on the IST, in software releases prior to 10.4.0, the default PVID of all IST port members must be the primary B-VLAN ID on both IST peers
  - This will happen automatically providing SPB is enable first prior to enabling the IST
  - Starting in software release 10.4.0, the PVID on the primary SMLT SPBM switch will be the secondary BVLAN while the PVID on the secondary SMLT SPBM switch will be the primary BVLAN
    - This will occur automatically when an IST peer is upgraded from 10.3.x to 10.4.0
- VLACP should not be enabled on the IST port members
- IP forwarding is not supported if SPB is enable
  - You will be prompted with an error message if you try to enable IP forwarding when SPB is enabled
  - An IST can still be created in SPB mode even though IP forwarding is disabled

dToR – Distributed Top of Rack

- dToR
  - Up to 16 switches can be deployed in a dToR (2 SMLT stacks of 8 switches each)
  - VSP 7000 front ports are used for the IST port members
  - SPB is supported
    - 10.2.1 supports a stack of two switches
    - 10.3 supports a stack of eight switches



Rear Port Mode – Standard (Raw)

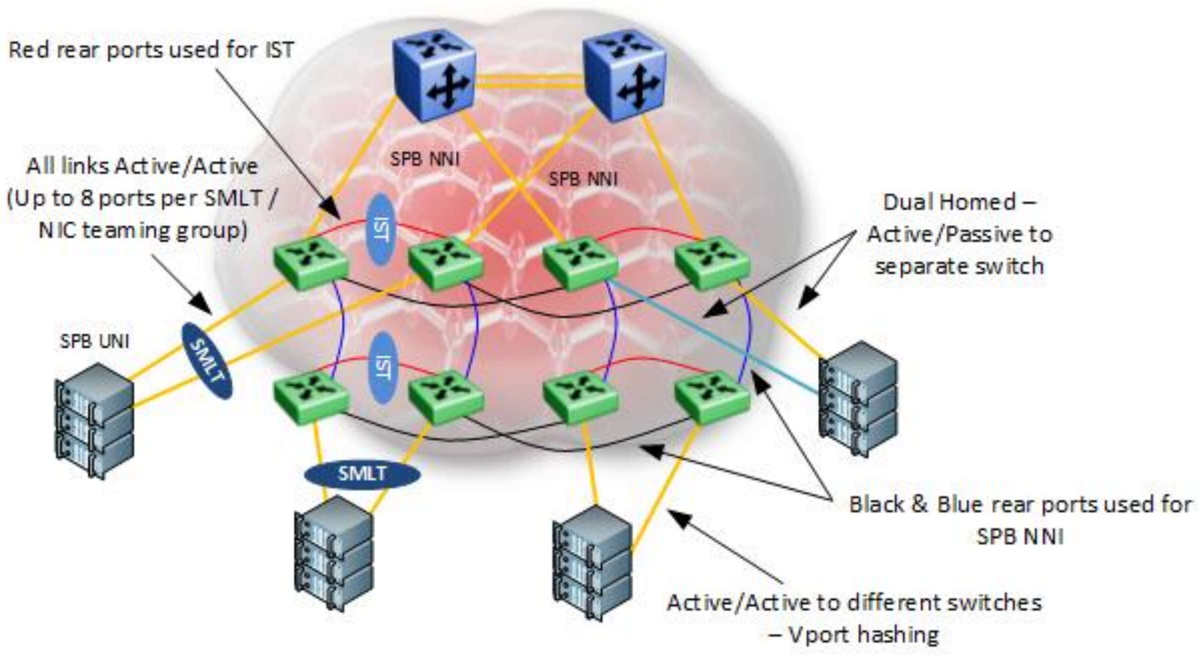


Up to 8 Gigabit Ethernet or 10 Gigabit Ethernet connections to core from each stack

Use Rear Mode Raw Mode to connect to other VSP 7000 switches also configured in Rear Port Raw Mode

- If the rear port standard mode is enabled, the rear port members can be used for high-speed IST/SMLT ports
  - All protocols are supported with the exception of Shortest Path Bridging
  - The IST must be made up of at minimum two ports
    - If you use the blue rear ports (ports 33 & 37) between two cluster switches, two stacking cables are required
    - If you use the red ports (ports 38-40) between two cluster switches, only one stacking cable is required
    - If you use the black ports (ports 34-26) between two cluster switches, only one stacking cable is required
  - LACP must be disabled prior to enabling IST or SMLT using the rear ports
  - SMLT square and triangle topologies are supported in rear port raw mode; SMTL full mesh is not supported

Rear Port Mode – SPB (Fabric Interconnect Mesh)



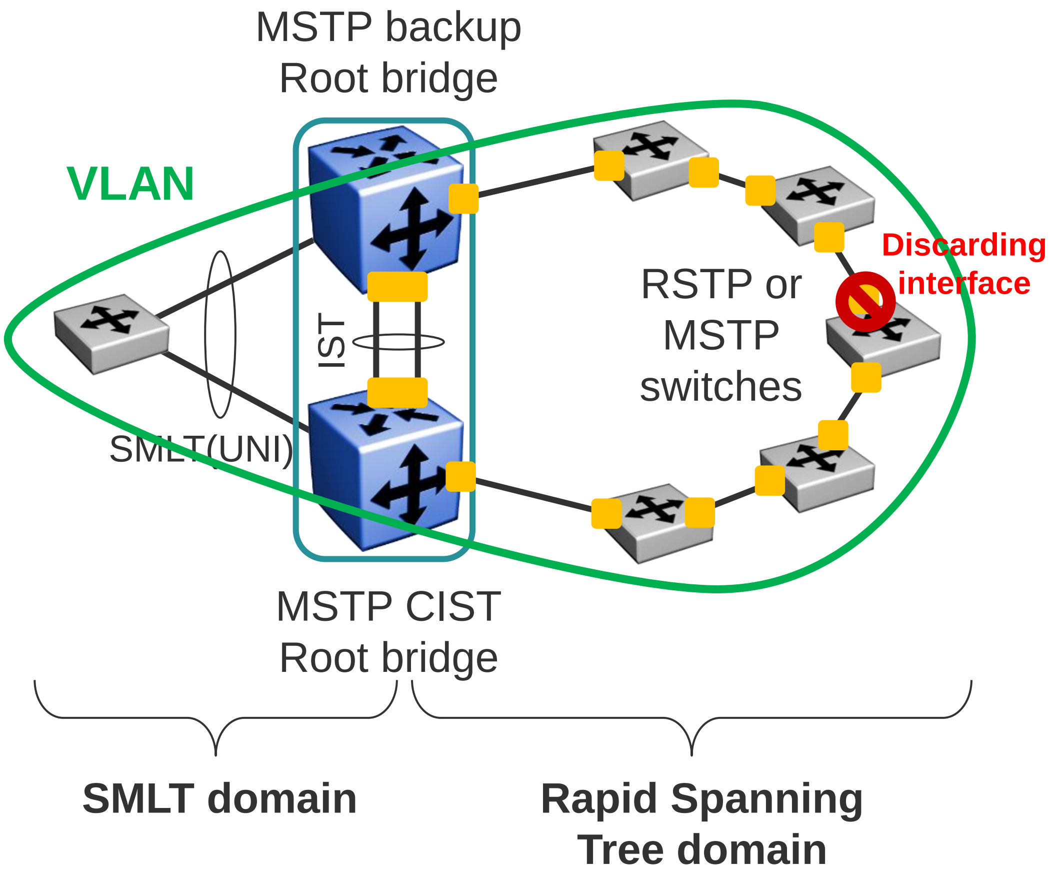
- If rear port mode SPB is enable, the rear ports can be used for high-speed SPB IS-IS ports with or without an IST
  - In rear port SPB mode, layer 3 is not supported and cannot be enabled, i.e. no Layer 3 protocol are supported
  - The IST must be made up of at minimum two ports
    - If you use the blue rear ports (ports 33 & 37) between two cluster switches, two stacking cables are required
    - If you use the red ports (ports 38-40) between two cluster switches, only one stacking cable is required
    - If you use the black ports (ports 34-26) between two cluster switches, only one stacking cable is required
  - LACP must be disabled prior to enabling IST using the rear ports
  - Release 10.3 or higher is required to also support rear port SPB mode with IST

STP over SMLT Design Best Practices

All the recommendations covered in the previous sections, also apply to SMLT when used with Spanning Tree enabled on the IST.
This section covers some additional best practices specific to an SMLT with STP enabled on the IST.

Spanning Tree over SMLT

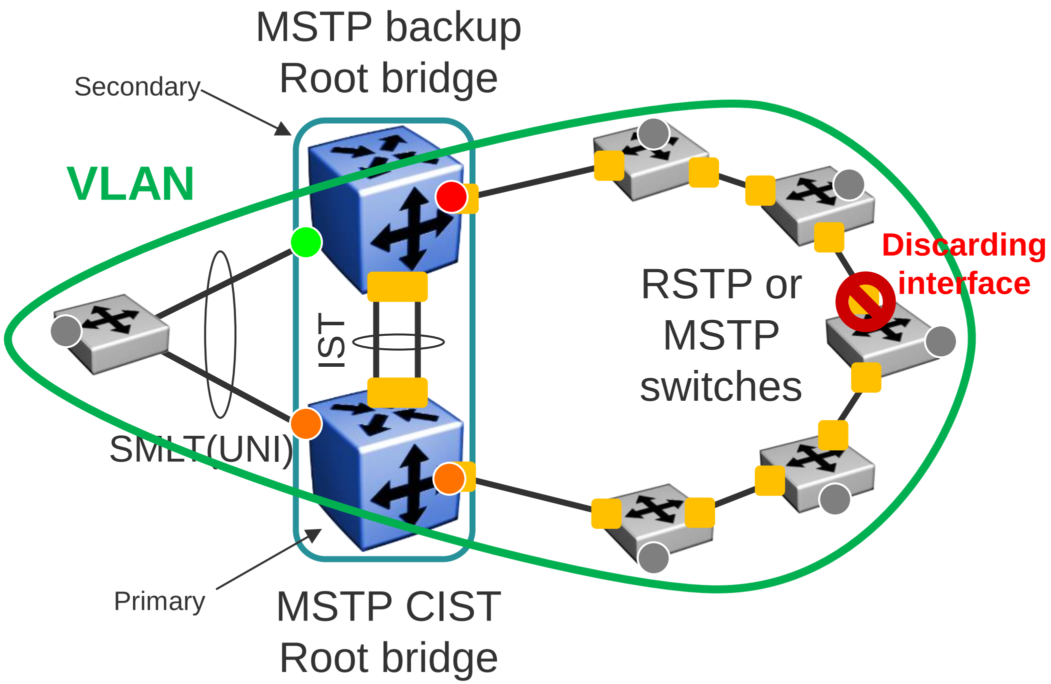
- ERS 8600/8800 feature introduced in software 7.1
- Ability to enable STP on the IST
- ERS 8600/8800 must be in MSTP mode
  - Not supported in legacy STP mode or RSTP mode
- Spanning Tree domain switches can be either in RSTP or MSTP mode
- Feature not supported in conjunction with SPB



🟨 Spanning Tree enabled interface

Spanning Tree over SMLT – SLPP Guidelines

- Leverage SLPP as a last resort defense in case of a Spanning Tree failure
- Use SLPP RxThreshold=1 on MSTP backup Root bridge Spanning Tree connection
- Follow same SLPP guidelines for all other interfaces



🔴 SLPP with RxThreshold = 1

🟠 SLPP with RxThreshold = 5

🟢 SLPP with RxThreshold = 50

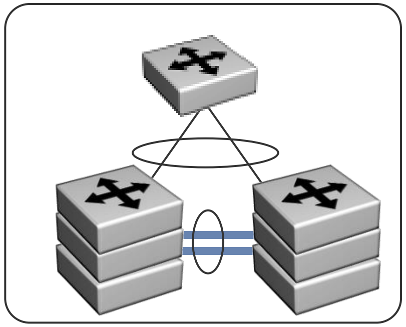
⚫ SLPP Guard

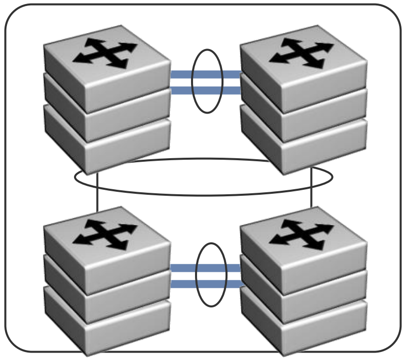
🟨 Spanning Tree enabled interface

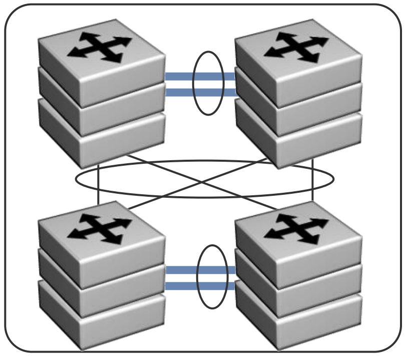
ERS 5000 Switch Clustering Design Requirements

ERS 5000 Switch Cluster

- No single point of failure in the core network
- Fast recovery when a link or a switch goes down
  - Sub-second recovery for L2 traffic in most cases
  - Sub-second recovery can be achieved in some cases for L3 traffic
- All redundant links are active – no Spanning Tree
- Switch Cluster supports
  - 1 IST
  - 31 SMLT
  - 512 SLT
- In 6.0 release, triangle, square, and full mesh are supported on both stand-alone and stack





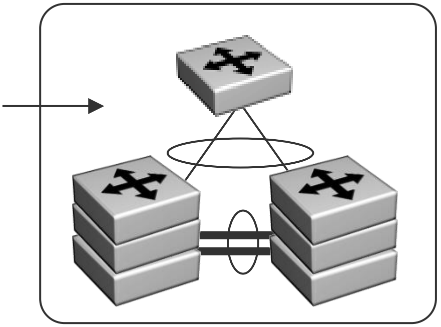


ERS 5000 Switch Cluster

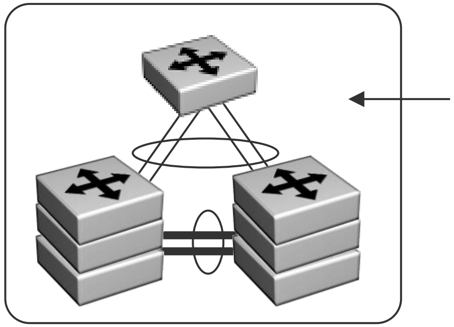
- IP Forwarding must be enabled (as of 6.2 it is automatically enabled)
- IST peers should use identical hardware, *but not an absolute requirement*
- LACP supported over SMLT/SLT starting in release 6.2
  - Requires at least two ports per unit in a IST cluster for SMLT
  - For SLT, one port is supported
- IGMP is over SMLT/SLT is supported starting in release 6.3
- PIM-SM is not supported with SMLT/SLT
- SMLT must have at least two port members on the switch/stack, thus, you cannot create an SMLT with just one port on each pair, must be SLT

SLT

Maximum Two Uplinks per Edge Connection



SMLT

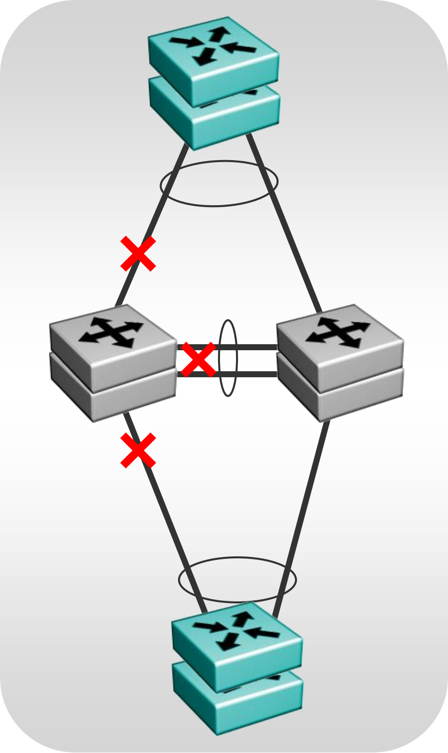


Scaling Uplinks Between the Edge and the Core

ERS 5000 Switch Cluster – Stack of 2

- Two behaviors in this case
- Behavior A
  - Software versions prior to 6.2.5
  - Software version 6.2.5 or later if stack force-mode is not enabled
  - Pros: Delivers faster convergence times upon unit failure
  - Cons: non-SLT ports on surviving unit remain isolated after unit failure
- Behavior B
  - Software version 6.2.5 or later when stack force-mode is enabled
  - Pros: behavior inline with expectations & surviving unit is not isolated
  - Cons: Delivers slightly slower convergence times upon unit failure

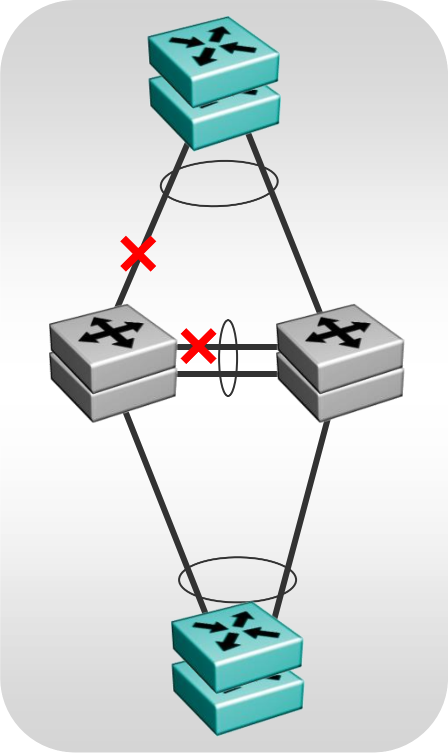
Behavior A



When a unit fails

- All SLT/SMLT/IST → Down

Behavior B

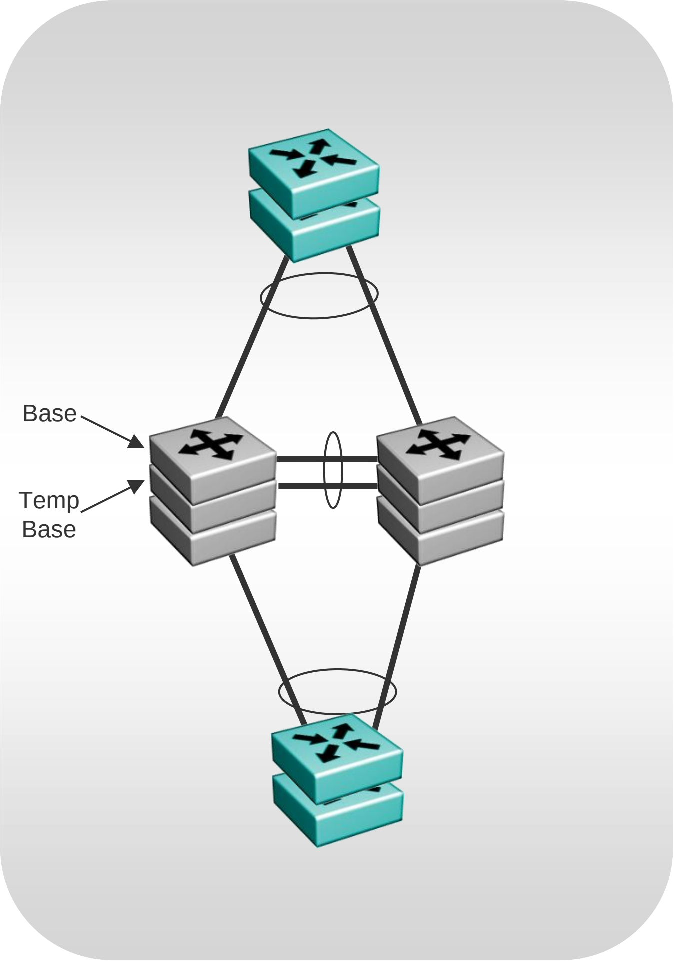


When a unit fails

- Only SLT/SMLT/IST on the failed unit → Down

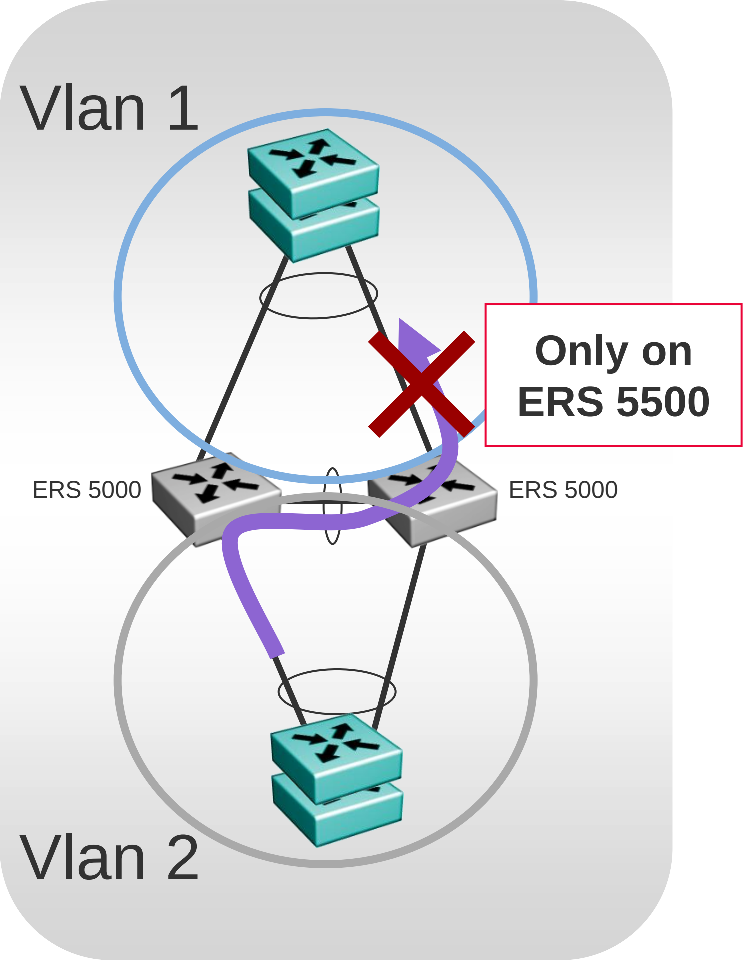
ERS 5000 Switch Cluster

- For the best resiliency, use at least three units when stacking the Switch Cluster Core
- For 2-unit stacks, enable stack forced mode
- When in stack configuration, at least one IST link should be on base unit (unit #1) and temporary base unit (unit #2). If more links are used in the IST, they can be spread across remaining units in the stack
- IST VLAN cannot be assigned as the management VLAN
- Cannot ping VRRP IP address from local CPU
- In a square or full mesh, aggregation pairs cannot use same VRID on the same VLAN



ERS 5500 Switch Cluster

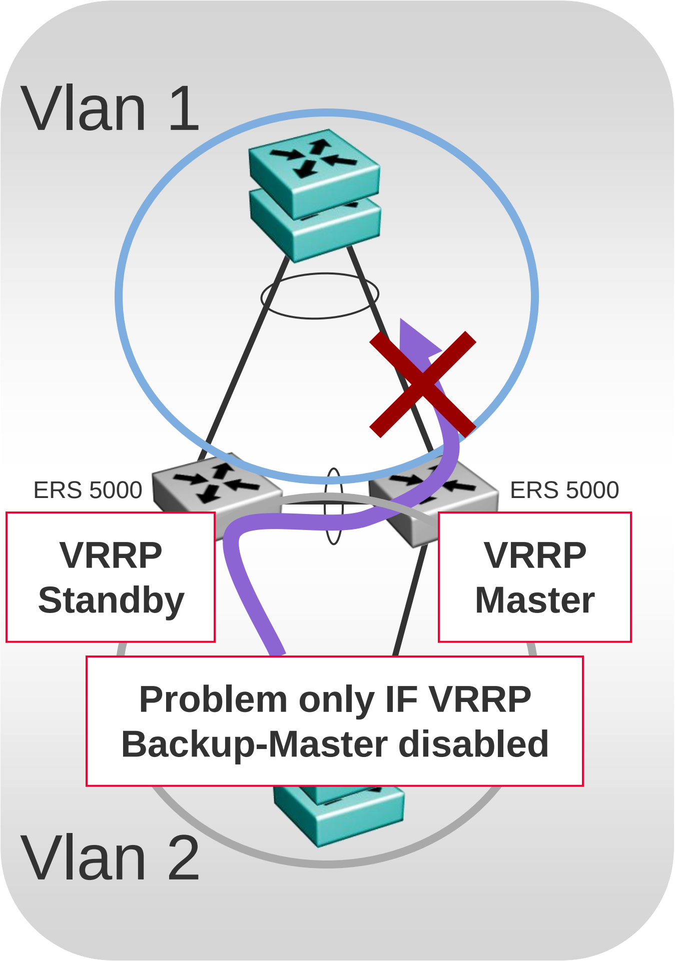
- SMLT Loop prevention
  - The simple rule of SMLT in the 5500 is that any packet received on the IST CANNOT be L2 switched out of an SMLT or SLT port which is active; where active means that the corresponding SMLT link on the IST-peer switch is up and running
  - The ERS 5500 will also not forward L3 packets that traverse the IST out an SMLT or SLT port
  - The ERS 5600 with 6.2 software does not have this limitation
    - With software 6.0 & 6.1 behaves like ERS 5500



Only on ERS 5500

ERS 5500 Switch Cluster

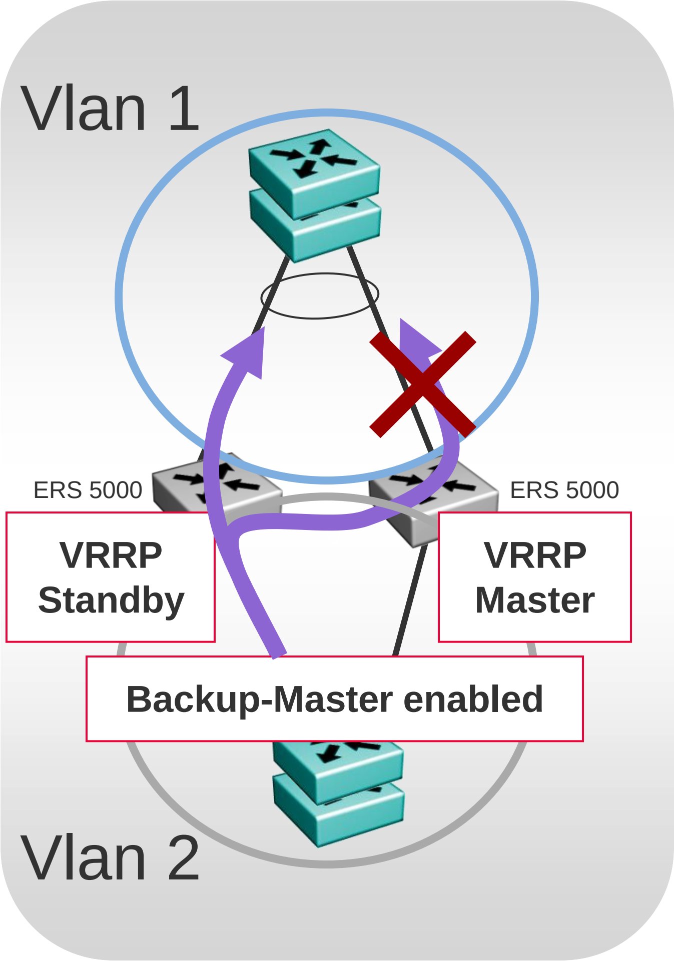
- Where is this a problem?
- VRRP
  - If VRRP Backup-Master is not enabled (or VRRP is disabled on one of the switches)



Problem only IF VRRP Backup-Master disabled

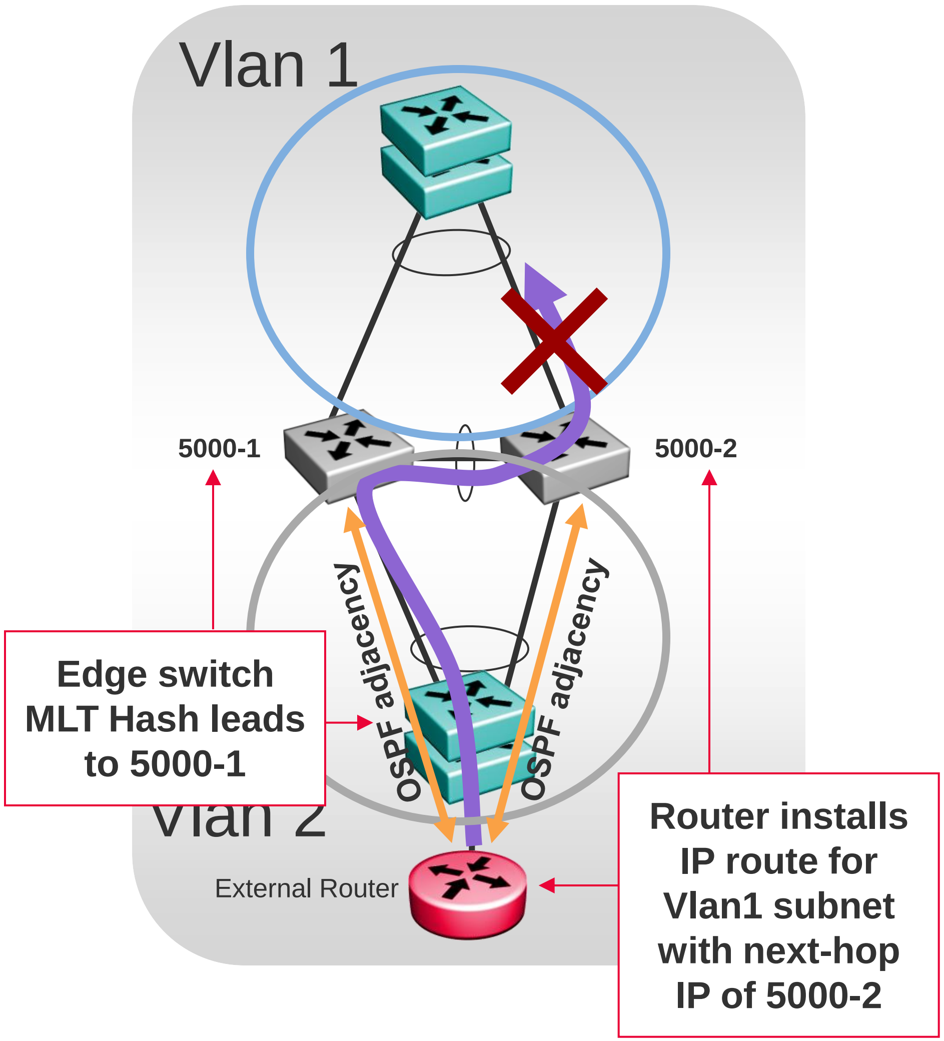
ERS 5500 Switch Cluster

- Where is this a problem?
- VRRP
  - If VRRP Backup-Master is not enabled (or VRRP is disabled on one of the switches)
- Solution
  - Enable Backup-Master
  - Do not activate any VRRP functionality which prevents Backup-Master from being active
    - Set VRRP-Hold-down to 0
    - No Critical VRRP interfaces



ERS 5500 Switch Cluster

- Where is this a problem?
- OSPF (also RIP)
  - If the edge router installs 5500-2 as next-hop in it’s routing tables and the edge switch MLT hashes routed traffic to 5500-1

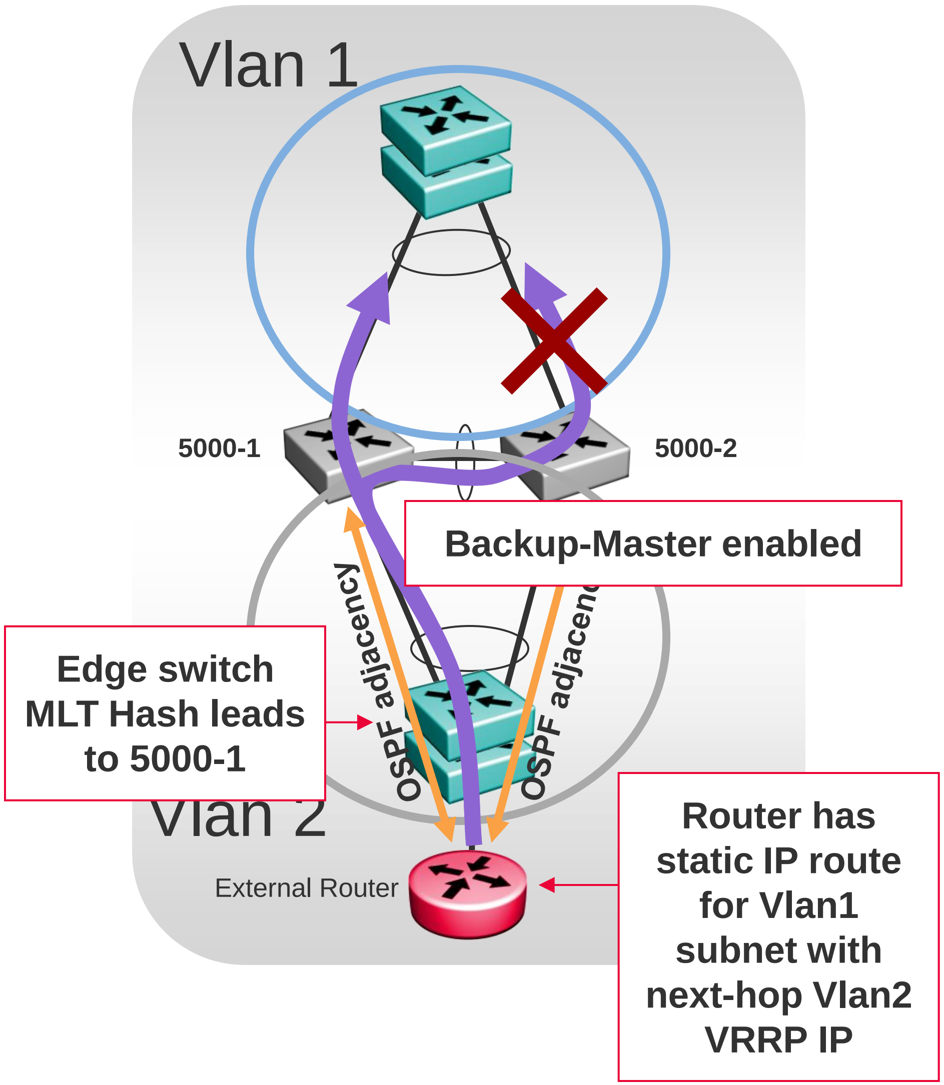


Edge switch MLT Hash leads to 5000-1

Router installs IP route for Vlan1 subnet with next-hop IP of 5000-2

ERS 5500 Switch Cluster

- Where is this a problem?
- OSPF (also RIP)
  - If the edge router installs 5000-2 as next-hop in it’s routing tables and the edge switch MLT hashes routed traffic to 5000-1
- Solution Today – 1
  - Use Static Routes on External Router and point them to ERS 5500 VRRP IP
  - VRRP Backup-Master required



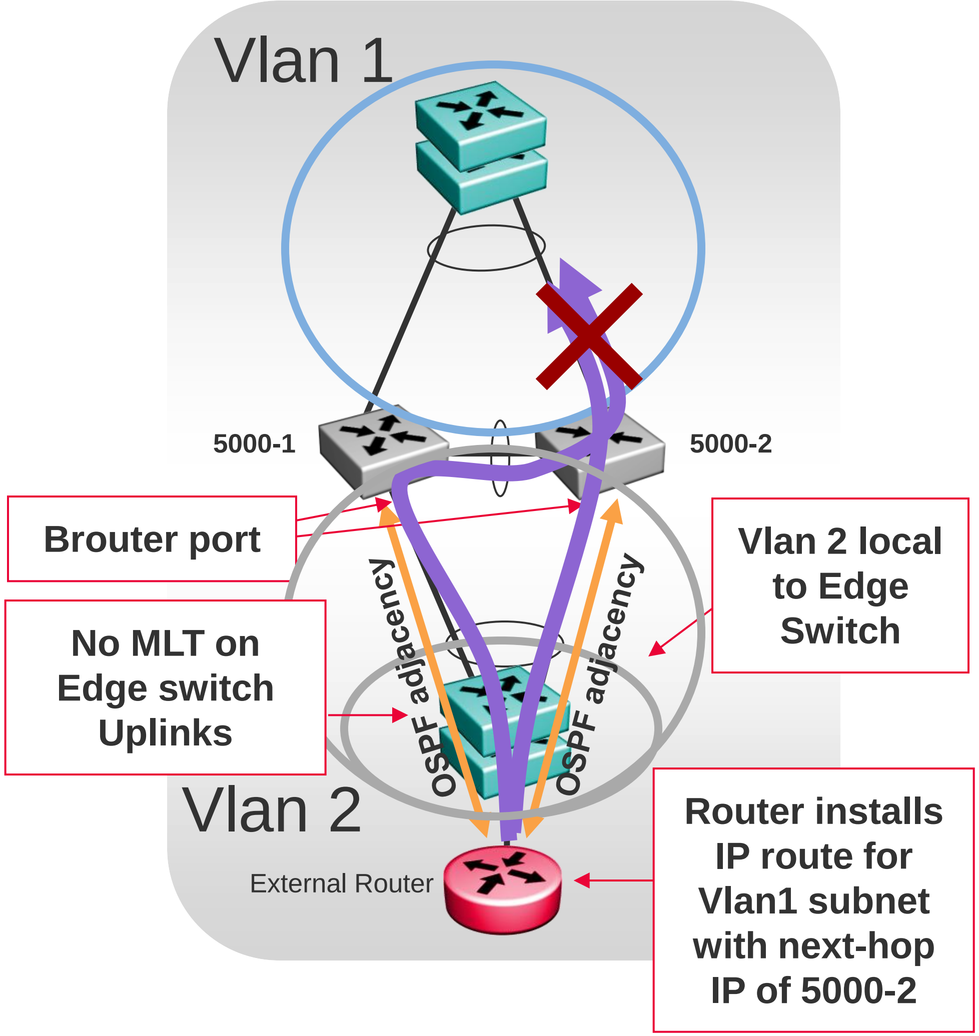
Backup-Master enabled

Edge switch MLT Hash leads to 5000-1

Router has static IP route for Vlan1 subnet with next-hop Vlan2 VRRP IP

ERS 5500 Switch Cluster

- Where is this a problem?
- OSPF (also RIP)
  - If the edge router installs 5000-2 as next-hop in it's routing tables and the edge switch MLT hashes routed traffic to 5500-1
- Solution Today – 2
  - If Static Routes not desired...
  - Connect edge switch with dedicated VLAN (no MLT/SMLT connection)



Brouter port

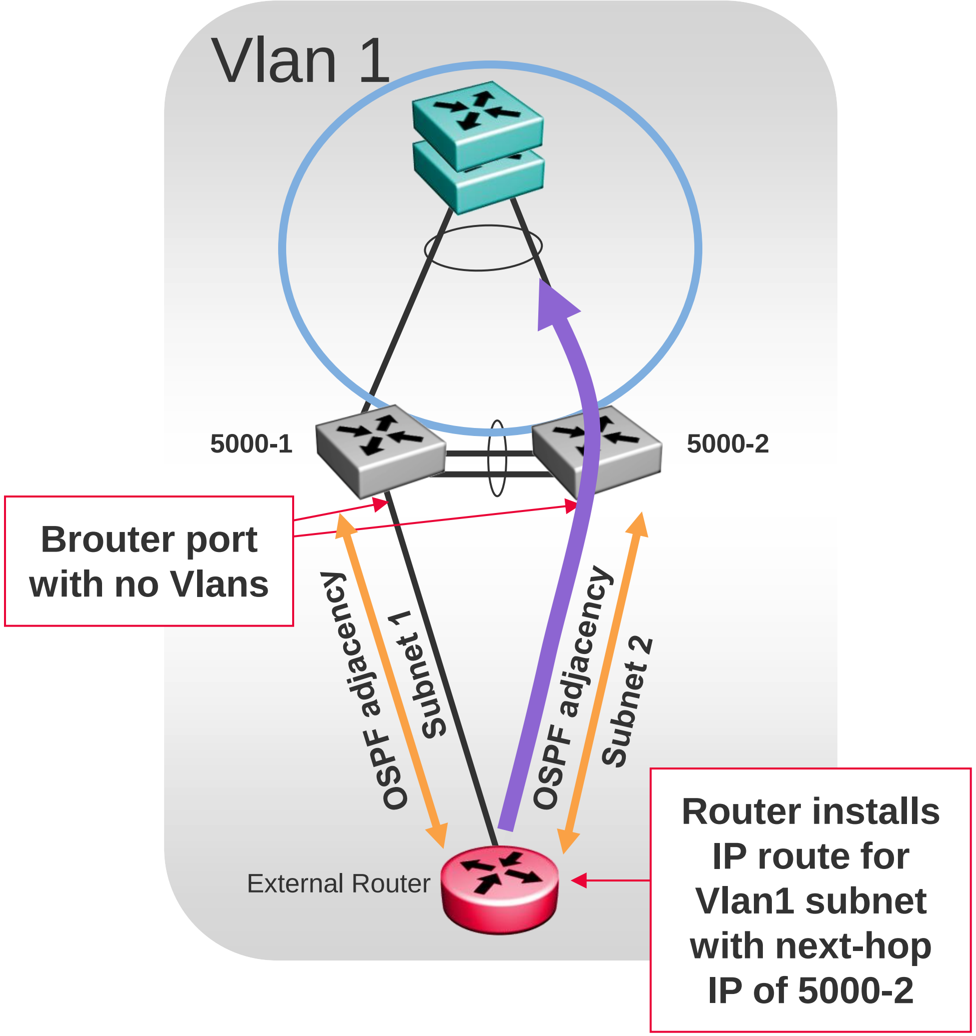
No MLT on Edge switch Uplinks

Vlan 2 local to Edge Switch

Router installs IP route for Vlan1 subnet with next-hop IP of 5000-2

ERS 5500 Switch Cluster

- Where is this a problem?
- OSPF (also RIP)
  - If the edge router installs 5000-2 as next-hop in it’s routing tables and the edge switch MLT hashes routed traffic to 5000-1
- Solution Today – 3
  - If Static Routes not desired…
  - Connect External Router via dedicated routed links

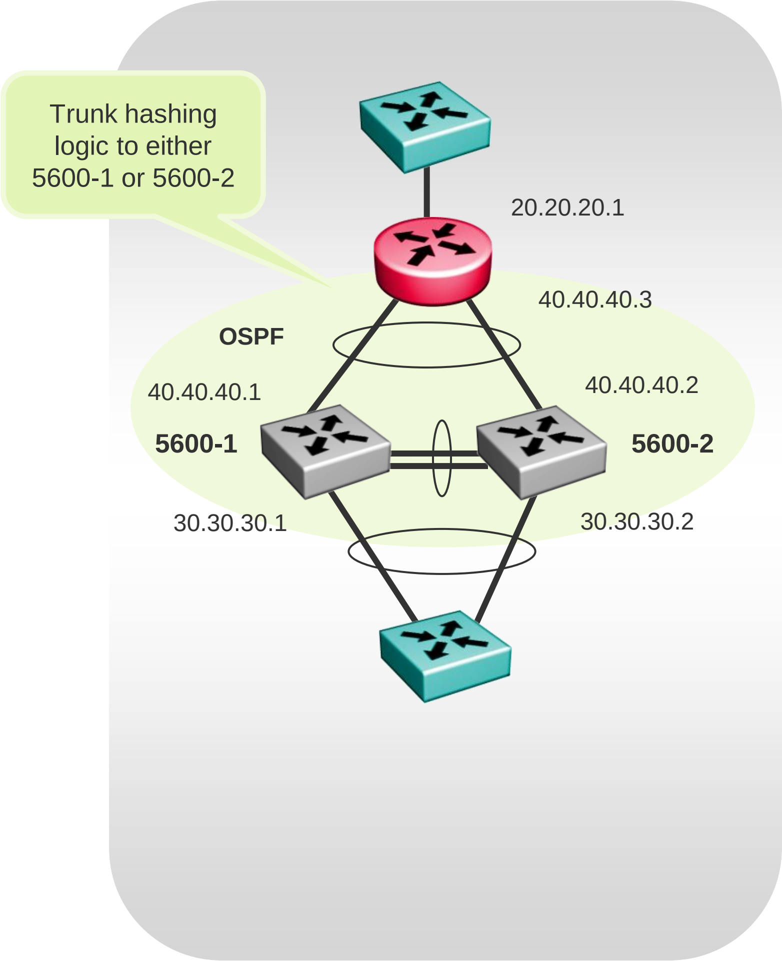


Brouter port with no Vlans

Router installs IP route for Vlan1 subnet with next-hop IP of 5000-2

Routing Protocol Support over SMLT – ERS 5600 cluster

- Prior to release 6.2, no traffic ever gets forwarded over Split trunk through the IST
  - Layer 3 traffic required VRRP backup master and static routes at the edge
  - No dynamic routing protocols are supported
- In the 6.2 release and for the ERS **5600** only, L3 traffic received by an IST port will be routed over SMLT trunk ports
  - Routing protocols such as OSPF across SMLT/SLT is supported eliminating the need for VRRP and static routes for L3 traffic


Trunk hashing logic to either 5600-1 or 5600-2

AVAYA

Engage The Power of We™